bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Flexibility in motor timing constrains the topology and dynamics of pattern generator circuits

Cengiz Pehlevan[1], Farhan Ali[2,3] and Bence P. Ölveczky[2,3#]

1. Center for Computational Biology, Simons Foundation

2. Department of Organismic and Evolutionary Biology, Harvard University

3. Center for Brain Science, Harvard University

#Corresponding Author

Summary

Temporally precise motor sequences underlie many motor skills and innate actions, yet the flexibility with which the timing of such stereotyped behaviors can be modified is not known. To probe this, we induced adaptive changes to the temporal structure of birdsong. We find that the duration of specific song segments can be modified without affecting the timing in other parts of the song. We derive formal prescriptions for how neural networks can implement such flexible motor timing. We find that randomly connected recurrent networks, a common approximation for how neocortex is wired, do not generally conform to these, though certain implementations can approximate them. We show that feedforward networks, by virtue of their one-to-one mapping between network activity and time, are better suited. Our study provides general prescriptions for how pattern generator networks can implement flexible motor timing, an important aspect of many motor skills, including birdsong and human speech.

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Introduction

The brain's ability to generate spatiotemporally precise motor patterns underlies much of what we do, from the expert performances of athletes and musicians to the daily acts of speaking or walking. Despite the ubiquity and importance of stereotyped motor sequences, our understanding of how the brain controls and modifies their timing remains poor (Buonomano and Laje, 2010; Viviani and Laissard, 1991).

Many different neural network architectures are capable of generating temporally reproducible dynamics, a minimum requirement for producing precisely timed motor output (Abeles, 1991; Goel and Buonomano, 2014; Kleinfeld and Sompolinsky, 1988; Laje and Buonomano, 2013; Sompolinsky and Kanter, 1986; Sussillo and Abbott, 2009). Randomly connected recurrent neural networks (RNN), a common approximation for how neocortex is wired (Sompolinsky et al., 1988) can be trained to generate prescribed temporal patterns (Jaeger and Haas, 2004; Laje and Buonomano, 2013; Sussillo and Abbott, 2009). Feedforward networks, such as synfire chains (Abeles, 1991), naturally map network activity to elapsed time, and hence are also suitable for generating timing signals underlying stereotyped motor output (Buonomano and Laje, 2010)

However, demands on pattern generator networks often go beyond simply producing the same dynamics every time an action is executed. A characteristic of skilled performance is the flexibility with which established motor patterns can be modified, including changes to their timing (Fortune et al., 2011; Gentner, 1987; Sternberg et al., 1990). Modifying the overall tempo of a continuous motor sequence is an obvious and often discussed example (Carter and Shapiro, 1984; Long and Fee, 2008) (Figure 1A). However, careful analysis of stereotyped action sequences suggests that the nervous system may be capable of altering motor timing in a far more specific and flexible manner (Figure 1A). In speech production or typing, for example, the

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

duration of discrete segments (e.g. phonemes or elements of key strokes) can change independently of other parts of a continuous action sequence (Sternberg et al., 1990, 1988). Such flexibility in motor timing can be essential for improving motor skills and adapting them to new contingencies or task demands. Increasing the power of your tennis serve, for example, may require selectively speeding up the power-generating stroke, while leaving the timing of the rest of the serve unchanged. Similarly, learning the speech patterns of a new language can benefit from being able to alter the duration of distinct phonemes and syllables in context-specific and flexible ways (Trofimovich and Baker, 2007).

To better characterize the flexibility with which the brain can modify the timing of stereotyped motor sequences, we took advantage of an experimental paradigm that allows adaptive changes to the temporal structure of zebra finch song to be induced (Ali et al., 2013). The complex learned vocalizations of adult songbirds are in many ways similar to human speech (Brainard and Doupe, 2013; Doupe and Kuhl, 1999), being temporally precise but, under certain experimental conditions, also quite malleable (Ali et al., 2013; Cai et al., 2011). Moreover, the temporal precision and reproducibility of birdsong are of ethological importance (Woolley and Doupe, 2008) as is the capacity to change the song's temporal structure (Fortune et al., 2011). These attributes make the adult zebra finch a good model for interrogating the flexibility with which the timing of complex and reproducible motor sequences can be modified.

Here we show that changes to the duration of specific song segments do not interfere with the temporal structure of other segments, and that two segments in the same song can be modified simultaneously and independently. Such flexibility in motor timing constrains the functional architecture of the underlying pattern generator circuits. We derive formal prescriptions for how such specific and adaptive changes to motor timing can be implemented in neural circuits. We show that vectors in synaptic weight space corresponding to maximal

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

changes in the duration of two intervals (i.e. the 'gradient' vectors) must be orthogonal for those intervals to change independently. We probe the extent to which neural networks capable of generating precise temporal patterns conform to this prescription. While we find that RNNs, in general, do not fulfill the strict conditions of orthogonality, certain implementations can achieve reasonable approximations. Feedforward networks, however, are more naturally suited as flexible time-keepers since they map network activity to time in an unambiguous manner.

Consistent with the theoretical analysis, simulations of temporal learning using biologically plausible plasticity rules in a synfire chain network reproduced the phenomenology observed in songbirds, while similar simulations in RNNs failed to attain the required degree of temporal flexibility. The different network simulations also generated predictions for how variability in the duration of song segment should change with changes to the segment's mean duration. Here too, the experimental data conformed to predictions from the synfire chain network, but not the RRNs. These results constrain the topology and dynamics of neural networks underlying flexible motor timing, and suggest that the synfire chain architecture may be ideally suited for generating the temporal structure of flexible motor behaviors.

Results

Flexible modifications to the temporal structure of birdsong

Perhaps the best understood example of how the brain acquires and executes complex learned motor sequences is the courtship song of the zebra finch (Immelmann, 1969). To characterize the flexibility with which the pattern generator circuit underlying stereotyped adult song can modify motor timing (Figure 1A), we challenged adult birds to produce changes to the duration of specific song segments (Figure 1B), using Conditional Auditory Feedback (CAF), a

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

reinforcement learning-based paradigm (Ali et al., 2013; Tumer and Brainard, 2007) (Methods). By playing loud (~80-90 dB) aversive sound bursts contingent on the duration of a targeted song segment, significant changes to the song's temporal structure can be induced in a matter of hours (Ali et al., 2013) (Figure 1C). To explore the specificity of these modifications, we compared changes in targeted segments to changes in other parts of the song.

Across the population of birds (n = 18 birds), the duration of targeted segments changed by, on average, 2.9 ± 1.7 ms (mean ± SD, n = 18 targets) per day relative to baseline drift (Figure 1D, Methods). In contrast, we saw no significant changes in non-targeted segments (n = 120) (Figure 1D). This shows that the timing of stereotyped learned motor patterns can change in temporally very *specific* ways.

We next tested whether learning-induced changes to the duration of one song segment interferes with adaptive modifications to other parts of the song by targeting two segments (>100 ms apart) for CAF simultaneously (Figure 1E, Methods). Interestingly, we found no difference in learning rates whether a segment was targeted for modification alone (2.8 ± 2.4 ms per day) or in conjunction with another segment (3.0 ± 2.4 ms per day, n = 12 segments in 6 birds, Figures 1F, G), suggesting that individual segments of song are, in terms of changes to their durations, *independent*.

Constraints imposed on flexible time-keeper circuits

Flexibility in motor timing can be characterized in terms of *specificity* and *independence*. *Specificity* means that changes to the duration of one part of a sequence leaves the temporal structure of other parts unaffected. *Independence* means that timing in different parts of a sequence can be modified simultaneously without any effect on learning rates. While many

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

different neural network architectures have been proposed to generate reproducible temporal patterns (Buonomano and Laje, 2010; Laje and Buonomano, 2013; Sussillo and Abbott, 2009), the extent to which they support such flexibility in motor timing has not been considered.

Here we use theoretical and computational approaches to examine how specificity and independence in motor timing constrain the topology and dynamics of the underlying pattern generator networks. We assume that network output is modified by changing connection strengths between its neurons, and seek to understand the network properties that enable flexible motor timing.

We show that in order to modify the duration of a discrete segment in a longer action sequence with independence and specificity (Figure 1A), there must exist a path in synaptic weight space along which *only* the duration of that segment (i.e. its interval) changes (specificity). Moreover, changes along such a path should not affect the network's capacity to change along paths associated with other intervals (independence). The analytical proof that this condition can be met in a generic neural network is presented in Supplementary Note I.

However, the existence of specific and independent paths in synaptic weight space is, on its own, not sufficient to ensure flexible motor timing. Such paths must be found and followed by biologically plausible learning algorithms. For real-world pattern generators, exemplified by the vocal control pathway of the zebra finch, modifications to motor timing often happens through trial-and-error learning, i.e. variations in temporal structure are generated and the certain variants reinforced (Ali et al., 2013). Our experiments with zebra finches showed that reinforcement targeted to specific song segments can produce specific and independent changes to the temporal structure of zebra finch song (Figure 1).

Synaptic learning rules that implement reinforcement learning typically find the direction (or gradient) in synaptic weight space that maximizes reward (or minimizes

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

punishment) and updates synaptic weights along that direction (Doya and Sejnowski, 1995; I. R. Fiete et al., 2007; Fiete and Seung, 2006; Williams, 1992). Restricting our discussion to this scenario, we show that specificity requires the direction in synaptic weight space corresponding to maximal changes in an interval (its 'gradient vector') to be orthogonal to all others gradients, or formally:

$$\sum_{\{ij\}} \frac{\partial I_\alpha}{\partial W_{ij}} \frac{\partial I_\beta}{\partial W_{ij}} = 0, \qquad \text{if } \alpha \neq \beta, \tag{1}$$

where I_α denotes the duration for interval α, W_{ij} the strength of the connection from neuron *j* to *i*, and $\frac{\partial I_\alpha}{\partial \mathbf{W}}$ the gradient for interval α, with the summation being over all synapses (see also Supplementary Note I). Importantly, the orthogonality constraint should hold throughout learning, not only for the original network. Otherwise, a network that initially exhibits specificity, could lose that property as the target segment (and the network) changes. See Supplementary Note I for a comprehensive discussion of this point. Importantly, equation (1) also ensures independent learning, assuming that the effect of separately delivered reinforcers sum linearly (see Supplementary Note I).

We next asked whether neural network architectures proposed to underlie the temporal patterning of complex action sequences conform to the prescriptions we derived. We consider two broad classes of networks: randomly connected recurrent networks (RNNs) (Laje and Buonomano, 2013; Sompolinsky et al., 1988; Sussillo and Abbott, 2009) and feedforward networks (Abeles, 1991; Goldman, 2009), and investigate the condition under which instantiations of these networks satisfy or approximate the orthogonality condition in equation (1). To quantify the extent to which a network deviates from the specificity and independence criterion (eq. 1), we numerically calculate the gradient vector associated with each interval and

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

then compute the "learning matrix", i.e. a symmetric matrix whose elements are given by the inner products of the gradients associated with pairs of target segments:

$$M_{\alpha\beta} \equiv \sum_{\{ij\}} \frac{\partial I_\alpha}{\partial W_{ij}} \frac{\partial I_\beta}{\partial W_{ij}}, \tag{2}$$

This matrix has non-zero off-diagonal elements when gradients are non-orthogonal.

The learning matrix of a pattern generator network quantifies the degree to which the network is flexible. For example, in an 'experiment' akin to the one we performed in songbirds (Figure 1), where the duration of a segment α (i.e. its interval) is targeted for modification, the expected change to the duration of another segment β, normalized by the change in the target interval α, is given by:

$$\frac{\Delta I^\beta}{\Delta I^\alpha} = \frac{M_{\beta\alpha}}{M_{\alpha\alpha}}, \tag{3}$$

Equation (3) shows how the learning matrix relates to *specificity*. Suppose $\alpha \neq \beta$. If the gradients with respect to intervals α and β are non-orthogonal, then $M_{\beta\alpha}$ is non-zero and there is a change in interval β, even though the reward was targeted at α. Hence, timing modifications are not specific to the targeted interval. In an 'experiment' where two intervals (α and β) are targeted for modification simultaneously, the expected change in interval α, compared to the expected change in an 'experiment' where only interval α is targeted, is given by:

$$\frac{\Delta I^\alpha_{two-tar.}}{\Delta I^\alpha_{single-tar.}} = 1 + \frac{M_{\beta\alpha}}{M_{\alpha\alpha}} \frac{\frac{dR^\beta}{dI^\beta}}{\frac{dR^\alpha}{dI^\alpha}}, \tag{4}$$

where R^α denotes the reinforcement targeted at interval α. Equation (4) shows how the learning matrix relates to independence. If the gradients with respect to intervals α and β are non-orthogonal, then $M_{\beta\alpha}$ is non-zero, meaning that learning directed at interval β will affect the learning rate of interval α. Hence, temporal modifications to different segments are not

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

independent. We note that the ratio $M_{\beta\alpha}/M_{\alpha\alpha}$ appears in both equations (3) and (4) and quantifies the deviation from specificity and independence. Below, we use it, or rather its absolute value, as a measure of non-flexibility, and call it the *interference* between intervals α and β. Therefore, with the knowledge of the learning matrix, we can quantify the degree to which a time-keeper network allows for flexible motor timing.

Probing random recurrent networks as flexible time-keepers

RNNs in their chaotic regime (Sompolinsky et al., 1988) produce spatiotemporal patterns suitable for generating complex motor output. However, chaotic network dynamics cannot produce the same output repeatedly and reliably because even small perturbations to the network can cause large changes in its output. Recently, algorithms that suppress chaos in RNNs (Laje and Buonomano, 2013; Sussillo and Abbott, 2009) have demonstrated the capacity of recurrent networks to learn and reliably produce complex stereotyped output, making them suitable as pattern generators. Next, we examine the extent to which such networks, specifically the feedback-stabilized RNNs of Sussillo and Abbott (2009) and the dynamic attractor networks of Laje and Bounomano (2013), support flexibility in motor timing.

FORCE training of feedback-stabilized RNNs exhibits a robustness-flexibility trade-off

Using the FORCE algorithm (Sussillo and Abbott, 2009), we trained the read-out units of an RNN with 500 'rate' neurons to produce an output pattern with readily identifiable 'intervals' (Figure 2A). In this network, chaos in the network is suppressed by having the linear read-out neuron feed back to neurons in the RNN (Rajan et al., 2010; Sussillo and Abbott, 2009), and hence we refer to it as "feedback-stabilized RNN" or fsRNN. For simplicity, we trained the network to produce a waveform with 10 equally spaced peaks, 50ms apart (Figure 2A, Supplementary

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 1A, Methods). This allowed us to define the start and end of 'segments' as the times the output crossed a preset threshold.

In fsRNNs, feedback from the read-out unit to the neurons in the RNN should be comparable in strength to the recurrent inputs to those neurons in order to suppress chaos and allow for successful training (Sussillo and Abbott, 2009). In our implementation, the strength of the feedback is governed by the parameter g_{FB}, which is typically set to be around 1 in applications of FORCE training RNNs (Boström et al., 2013; Hoerzer et al., 2014; Sussillo and Abbott, 2012, 2009). We varied this number from 0 to 5, trained 20 networks for each value, and found that increased feedback improves training success as reflected by a decrease in the 'test' error, i.e. the normalized error between network output and desired output (Figure 2B, Methods). To assess whether such networks can produce precise timing signals, we quantified the fraction of simulations (out of 400) in which the network failed to produce all 10 intervals within 3ms (i.e. 6%) of the target duration (the 'timing failure rate', Figure 2B). This metric was close to zero even when the strength of the feedback was low, meaning that even networks with outputs further from the specified target can produce activity modulations useful for marking the starts and ends of intervals.

Looking at the networks more closely, we found that changes in feedback strength led to qualitatively different solutions (Figure 2A). When feedback was strong, it dominated the recurrent input to the pattern generator network close to interval borders. Because the feedback input was periodic, the network state, defined by the vector of all instantaneous unit activities, was reset to the same state at interval borders (Figure 2C), leading to periodic activity in the pattern generator network (see example unit dynamics in a network with $g_{FB} = 2$, Figure 2A). When feedback strength was low, the recurrent input dominated, causing non-periodic network activity (see example unit dynamics in a network with $g_{FB} = 1$, Figure 2A). Given these

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

qualitatively different dynamic regimes, we probed RNNs with varying degrees of feedback in terms of their capacity for robust and flexible time-keeping.

We first quantified the robustness of the network as a function of feedback by injecting 10ms perturbation pulses of different magnitudes into all units of the network similar to Laje and Buonomano (2013) (Figure 2D, Supplementary Figure 1B). We applied this perturbation during the second interval and found that the timing failure rate decreased with increased feedback, though not much beyond $g_{FB} = 1$. This implies that networks become more robust as the strength of feedback is increased.

We next examined how these networks cope with demands for timing flexibility by calculating their learning matrices (Figure 2E). We found that learning matrices are sensitive to the strength of the feedback, g_{FB}: at low feedback, off-diagonal elements are small, implying relatively low interference between intervals and high specificity. At high feedback, off-diagonal elements are almost as large as diagonal elements, implying high interference and low specificity. Interestingly, at high feedback, the first interval behaves differently than the others, and interferes much less. This is because the network state at the beginning of the first interval, set by an external start signal, is different than the network states at the beginning of other intervals (Figure 2C), which are determined largely by the feedback.

Timing flexibility was quantified by calculating the normalized learning matrix elements, $\left|\frac{M_{\beta\alpha}}{M_{\alpha\alpha}}\right| \times 100\%$, which reflect interference for all pairs of intervals as discussed above. Interference went from ~20% to almost 100% as feedback increased (Figure 2F). Taken together, our results (Figures 2C-F) show that feedback leads to a trade-off between robustness and timing flexibility, with strong feedback making the network more robust to perturbations, but less flexible in the timing domain.

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Thus far, we have inferred flexibility from gradient vectors in synaptic weight space and metrics derived from those. Biological implementations of reinforcement learning, however, must find those gradients. To understand how this affects timing flexibility, we simulated the process of reinforcement learning by implementing a biologically plausible rule that adaptively updates synaptic weights to maximize reward on average (Ila R. Fiete et al., 2007; Fiete and Seung, 2006) (Methods). Noise was added to the network units to ensure trial-by-trial variability in interval durations (Methods). Reinforcement was provided to the network in a manner analogous to our CAF experiments: if the target interval differed from the running average in the desired direction (Methods), reinforcement was delivered at the end of the interval (Methods). This learning rule successfully modified the duration of target intervals, with learning rates comparable to those seen during CAF (Figures 2G, H). When averaged across multiple runs of the simulated learning 'experiments', the interference patterns were largely predicted by and consistent with the learning matrix calculations (Figures 2I, J). Interference in single 'experiments', however, exceeded the predictions from gradient descent, i.e. the learning matrix, by a significant amount (Figure 2K), suggesting that experiment-to-experiment variability in changes to non-target intervals is large. The single experiment interference decreased with feedback and saturated around 35% (averaged across intervals, networks and simulations, Figure 2K, Methods).

To assess independence in the temporal domain, we targeted two intervals for modification as we had done experimentally in songbirds, with separate reinforcements given contingent on the duration of both intervals (Methods). As expected from the learning matrices, this resulted in destructive interference, meaning that learning rates for individual intervals decreased when they were modified in conjunction with another (Supplementary Figures 1C-E). This interference increased with feedback, leading to almost no learning for $g_{FB} > 2$ (Figure x).

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Time-keeping with dynamic attractor networks

Recently Laje and Buonomano (2013) proposed a new dynamic regime for RNNs, the dynamic attractor, which they showed can serve as a "population clock" suitable for reliably generating temporally precise output due to its robustness to perturbations. However, the extent to which the dynamic attractor allows flexible motor timing was not considered.

The dynamic attractor has a similar architecture to the fsRNN, but differs from it in that there is no feedback (Figure 3A). The RNN is instead trained using an "innate training" procedure (Laje and Buonomano, 2013), which applies the FORCE algorithm to neurons in the pattern generator network to stabilize a chaotic trajectory already produced by the network. This turns the trajectory into an attractor such that even if the network is perturbed, it returns to the original trajectory, yielding a reliable and precise output.

We trained instantiations of the dynamic attractor network, matched in size and other parameters to those of the fsRNNs (Methods), to produce output patterns from which interval durations could readily be extracted (Figure 3A, Methods). We found that a large fraction of the trained networks successfully produced all 10 intervals within 3 ms of their target duration (Figure 3B). In our analysis, we only considered networks that were successful >99% of the time (out of 400 simulations, 14 out of 20 trained networks). The robustness of these networks to perturbations was comparable to fsRNNs with low feedback ($g_{FB} \approx 1$), i.e. similar perturbation magnitudes led to similar failure rates (Figure 3C). These results ensured that our dynamic attractor networks were comparable to fsRNNs in reliability and robustness, allowing us to compare them in terms of their capacity to support flexible motor timing.

Learning matrices revealed less interference across intervals (~12%) than the low feedback RNNs (~23%) (Figures 3D,E). In contrast to the fsRNNs trained by the FORCE learning

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

algorithm, the interference between adjacent intervals was, on average, much larger (~35%) than the average interference (Figure 3E). This is consistent with the dynamic attractor trying to 'return' to its original pattern, compensating for a decrease/increase in the duration of an interval by increasing/decreasing the duration of adjacent intervals. In line with this, the learning matrix elements corresponding to adjacent intervals consistently had negative signs (Figure 3D).

When we modified the duration of targeted intervals with the same reinforcement learning algorithm that we used for the fsRNNs (Methods), learning rates were similar to the simulated experiments in fsRNNs and comparable to what we observe in songbirds (Figures 3F, G). Changes in the duration of target and non-target intervals were, over all, well predicted by the learning matrices (Figure 3H), though as with the fsRNNs, interference on average (~20%) exceeded what was expected by gradient descent (~12%) (Figure 3I).

To assess independence, we again targeted two non-adjacent intervals for simultaneous modification. Reductions in learning rates were well predicted by the associated learning matrices (Supplementary Figure 2). The changes in learning rates were less (~10%) compared to low feedback fsRNNs (~20%), consistent with dynamic attractors having less interference on average, with the interference affecting primarily intervals adjacent to the target.

These results suggest that dynamic attractor networks, as in Laje and Buonomano (2013), allow for more flexible adjustments to motor timing than fsRNNs. In contrast to the fsRNNs, the interference was largely concentrated on nearest neighbor intervals and changes to non-targeted intervals were opposite in sign to the target.

Flexible time-keeping in feedforward networks requires unique mapping between synapses and interval durations

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Feedforward networks have been proposed as alternatives to RNNs for generating temporal patterns (Buonomano and Laje, 2010; Fiete et al., 2004; Goldman, 2009) (Fig 5A). By organizing neurons into layers connected in a feedforward manner, these networks naturally map layer-specific neural activity to elapsed time (Buonomano and Laje, 2010). In this section, we focus on feedforward networks with a single neuron per layer. This simplification allows us to analytically study the constraints imposed on feedforward network dynamics by demands for flexible time keeping. Using the insight garnered from this, we discuss more realistic feedforward networks in the next section.

Applying the orthogonality constraint (eq. 1) to such networks revealed that flexible time-keeping requires synaptic weights to be mapped to interval durations in a unique one-to-one manner (Supplementary Note II). Consider an interval I_α, the duration of which depends on activity in layers $\alpha - 1$ and α. In Supplementary Note II we prove that for flexible time-keeping, changes to the synaptic weights between layers $\alpha - 1$ and α should only affect the duration of I_α: If W_{ij} is not a synapse between neurons in layers $\alpha - 1$ and α, then

$$\frac{\partial I_\alpha}{\partial W_{ij}} = 0. \tag{5}$$

Note that this requirement is naturally satisfied by the synaptic connections between downstream layers, as they cannot affect activity in upstream layers. Still, equation (5) constrains the types of dynamics a feedforward network can exhibit. As we show (see Supplementary Note II), the one-to-one mapping constraint (eq. 5) is satisfied only if synaptic weight changes alters the *timing* of the post-synaptic neuron's activity, but not its shape or magnitude. This ensures that downstream neurons encoding non-target intervals shift their activity in time by the same amount, keeping non-target interval durations unchanged, thus satisfying equation (5).

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

To further illustrate the constraints equation (5) imposes on dynamics, we present two simple and analytically tractable examples of feedforward networks, and show how they fail or succeed in flexible time-keeping. The first example is a chain of simple integrate-and-fire neurons. The initial synaptic weights are chosen such that when the first neuron in the chain produces a single spike, downstream neurons in the chain propagate activity by producing a single spike each. Spike timings mark interval boundaries. This network satisfies the one-to-one mapping constraint (eq. 5) for a finite volume of synaptic weights in weight space:

$$\frac{\partial I_k}{\partial w_l} = \begin{cases} -\dfrac{\tau}{w_l\left(\dfrac{w_l E(t_k)}{V_{th}} - 1\right)} & l = k \\ 0 & l \neq k \end{cases}, \qquad (6)$$

where $E(t)$ is the shape of the excitatory post-synaptic potential, w_l the weight of the synapse between neuron $l-1$ and l, τ the membrane time constant, and V_{th} the threshold potential for spiking. If w_l is increased, it will advance the firing of the l^{th} neuron, since it will reach spiking threshold sooner. However, as long as the neuron produces only one spike, the time it takes for the next neuron to spike will not change, keeping all subsequent intervals unchanged. Similarly, decreasing w_l causes the the duration of interval I_l to increase, without effecting the others. Note that this is exactly the mechanism described in the previous paragraph, as synaptic weight changes affect the timing of the post-synaptic response, but not its magnitude or shape. Therefore, constraint (eq. 5) is satisfied and the interference matrix is diagonal.

There are limitations to how much flexibility a chain of integrate-and-fire neurons can exhibit. Flexibility breaks down if the synaptic weights decrease below a point where the excitation in the post-synaptic neuron no longer drives it above spiking threshold. In this case chain propagation stops. On the other end, if the synaptic weight increases to a point where the

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

post-synaptic neuron produces multiple spikes, then the next interval gets shorter as the boost in excitation from the extra spike(s) will cause the next neuron in the chain to reach threshold sooner. Therefore, equation (6) holds for a finite range of synaptic weights (see Supplementary Note III for the exact expressions). Importantly, this range can be increased by other mechanisms that make the shape of the post-synaptic response less sensitive to changes in input, such as refractoriness or active burst-generating mechanisms (Long et al., 2010).

The second example is a chain of rate neurons with linear activation functions. If the first neuron in the chain is excited with a unit pulse, the activity of neurons in the network is given by $r_i(t) = \left(\prod_{j=1}^{i} w_j\right) \frac{t^n}{n!\tau^n} e^{-t/\tau}$, where τ is the neural time constant. The upward threshold crossings of the instantaneous firing rates mark interval boundaries. Let t_i be the time at which the instantaneous firing rate of the i^{th} neuron crosses its threshold, which can be set arbitrarily as long as neurons cross the threshold in the order of chain progression. Then,

$$\frac{\partial I_k}{\partial w_l} = \begin{cases} 0 & l > k \\ -\frac{1}{w_l}\frac{t_k}{k\tau - t_k} & l = k \\ -\frac{1}{w_l}\left(\frac{t_k}{k\tau - t_k} - \frac{t_{k-1}}{(k-1)\tau - t_{k-1}}\right) & l < k \end{cases}. \quad (7)$$

Equation (7) leads to an interesting conclusion: the orthogonality constraint (eq. 5) is satisfied only when intervals have equal duration, i.e. when $t_k/k\tau$ is constant. Any change from the equally spaced interval boundaries would lead to a violation of the constraint in equation (5), meaning that such networks cannot implement flexible time-keeping. The reason linear rate networks fail is that weight changes do not only shift the postsynaptic response in time, but also in magnitude. This leads to interference effects by altering the duration of subsequent intervals.

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

In realistic feedforward networks with multiple spiking neurons per layer, layer-to-layer propagation of excitation can occur in asynchronous or synchronous modes (Abeles, 1991). The rate-neuron chain operates effectively models the asynchronous mode, where each rate-neuron's activity is the population average of the spiking-neurons instantaneous firing rates within a layer. In such a network, aggregate activity in each layer depends on the magnitude of presynaptic input. Hence synaptic changes in one part of the chain will affect the activity of downstream neurons. The synchronous state, as we show next, is closer to the spiking-neuron chain example and conforms to the constraints on flexible time-keeper networks.

Synfire chains are flexible time-keepers

We simulated a chain of integrate-and-burst spiking neurons, modeled on the calcium-spike generated bursts observed in the putative time-keeper circuit of the songbird (Long et al., 2010). The network had 140 layers with 15 neurons per layer, for a total of 2100 neurons (Figure 4A, Methods), and was configured to propagate activity in a synchronous mode, hence a 'synfire chain' (Figure 4B). Output neurons received input from all neurons in every 14^{th} layer. The first spike of each output neuron marked the time of an interval boundary (Figures 4A,B), yielding 10 intervals of roughly 50 ms average duration. As expected from the analogy with the single spiking-neuron chain, the learning matrix for the synfire chain was diagonal, i.e. non-diagonal elements were close to zero, with the average non-target interference being $0.3 \pm 0.7\%$ (Figure 4C).

Modifying the duration of specific intervals using the same reinforcement learning algorithm as for the recurrent networks (Methods), resulted in learning rates comparable to what is seen in songbirds (Figure 4D). Changes in the duration of targeted and non-targeted intervals were, on average, well predicted by the learning matrices (Figure 4E). The interference

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

was an order of magnitude smaller than in the recurrent networks (~1-2%, Figure 4F), and consistent with what we observed in songbirds.

To assess independence, we targeted two non-adjacent intervals for simultaneous modification. The average reduction in learning rates was small (~3-5%) (Supplementary Figure 3), again consistent with our observations in songbirds (Figure 1G). These results suggest that synfire chains allow for flexible adjustments to motor timing.

Variability in temporal structure

To further constrain the likely network architecture underlying song timing in zebra finches (Long et al., 2010; Yamashita et al., 2008), we analyzed temporal variability of the song segments targeted in our CAF paradigm. We found that when segments were shortened, variability in their duration decreased, while when the segments were lengthened their duration variability increased (Figure 5A). This linear relationship held across multiple days of learning for the same target (Pearson's correlation coefficient, $r = 0.98$, $p = 6.1 \times 10^{-6}$, $n = 9$ days; Figure 5B). When compared across different targets (CV - normalized variability) in our population of birds, the correlation remained high ($r = 0.7$, $p = 6.5 \times 10^{-4}$; n=20 birds; Figure 5C).

We proceeded to analyze the extent to which the network models we have considered reproduce the linear relationship between mean change in interval duration and its variability. For both types of recurrent pattern generator networks, we found the relationship between target interval duration and variability to be inconsistent. Different networks produced different relations: variability could increase, decrease or have no correlation with interval duration. It could even have a convex shape (Figures 5D-E). The synfire chain networks however, showed a linear increase in variability as a function of target interval duration (Figure 5F), conforming to the experimental observations in songbirds.

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Discussion

Modifying the temporal structure of stereotyped motor sequences is an important means of adapting motor skills to new task demands (Ali et al., 2013; Fortune et al., 2011; Gentner, 1987; Sternberg et al., 1990). Here we combined learning experiments in songbirds with theory and network simulations to explore the flexibility with which pattern generator circuits, real and simulated, can modify the temporal structure of their outputs. We found that the timing of zebra finch song can be altered with a high degree of specificity (Figure 1), and derived formal prescriptions for the neural networks that implement such flexible motor timing.

We show that gradient vectors in synaptic weight space associated with different intervals in a sequence must be orthogonal for the intervals to be independently modifiable. Certain feedforward networks satisfy this criterion making them ideally suited as flexible time-keepers. The general topology of RNNs, however, does not guarantee such flexibility. Whereas synapses in feedforward networks naturally map onto specific time points in a sequence, synapses in RNNs that contribute to changing one interval are often not unique from those that drive changes in other intervals, leading to interference in the time domain.

We found that the degree of timing flexibility further depends on the specifics of the networks and the properties of its neurons. In RNNs, strong feedback from its output neurons back to the RNN, which can increase the network's robustness to perturbations (Figure 2D), makes the network less flexible in the temporal domain (Figure 2F). The dynamic attractor network of Laje and Buonomano (2013), which does not require any feedback, shows overall less interference (Figure 3). While feedforward networks are structurally better suited, we found that flexible time-keeping is compromised in these networks if synaptic changes between neurons in successive layers of the netork alter the post-synaptic spiking responses beyond just

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

shifting them in time. In this regime, modifying the strength of a synapse will not only affect the interval associated with that synapse, but it will also propagate the effect to downstream neurons causing interference in subsequent intervals (Supplementary Notes II and III).

The range over which feedforward networks operate as flexible time-keepers can be extended by making its neurons' spiking responses less sensitive to the magnitude of presynaptic input. Interestingly, projection neurons in HVC, the putative time-keeper circuit in songbirds (Hahnloser et al., 2002; Long et al., 2010; Long and Fee, 2008), generate calcium bursts that accomplish just that. Modeling a synfire chain network with integrate-and-burst neurons, akin to those described in HVC (Long et al., 2010), recapitulated the flexible timing observed in songbirds (Figure 4).

Additional support for the idea that song-timing is governed by a synfire chain-like network comes from the relationship between variability in interval duration and changes to its mean (Figures 5A-C). Whereas RNNs show no consistent relationship (Figures 5D,E), synfire chain predict the linear relationship between variability and mean duration we observe (Figure 5F). This relationship is a consequence of how interval durations are altered in feedforward networks: synaptic strengthening between neurons in subsequent layers makes the signal propagate faster, shortening the associated interval. But synaptic strengthening also makes the synapse less prone to noise and hence less variable (Garst-Orozco et al., 2015), while weakening the synapse has the opposite effects. Whether this characteristic linear relationship between interval duration and variability, suggestive of an underlying feedforward network architecture, is seen also in other flexible behaviors, such as speech, remains to be investigated (Cai et al., 2011).

While both recurrent and feedforward networks have been proposed to underlie the temporal structure of birdsong (Hahnloser et al., 2002; Yamashita et al., 2008), our findings

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

together with other recent experimental results (Long et al., 2010; Long and Fee, 2008), suggest that the network that controls song timing indeed functions as a synfire chain.

We also found that trial-and-error learning in RNN networks resulted in more temporal interference than predicted by gradient descent (Figures 2K and 3I). This is because individual learning experiments do not find the optimal synaptic updates, i.e. those that align with the gradient. This solution only emerges after averaging over many experiments. For an intuition as to why trial-and-error learning leads to more interference, consider two intervals whose gradients are orthogonal, one of which is targeted for modification by trial-and-error learning. Because the synaptic updates at the end of an experiment will not exactly point along the target interval gradient, the updates will not be orthogonal to the gradient of the non-targeted interval and will hence cause a change in its duration too, i.e. interference.

While we have limited our study to structurally feedforward networks, our results generalize also to functionally feedforward networks (Goldman, 2009). In these networks, specific modes of network activity (Schur modes), rather than individual neurons, are organized into layers and the connectivity matrix is feedforward in the basis of these modes. Our results suggests that timing flexibility in such networks would be compromised if synaptic changes alter the Schur modes beyond just shifting them in time. Functionally feedforward networks made of rate neurons will suffer from this, as do the feedforward networks of rate neurons we discuss (equation 7). What is needed for flexible time-keeping is a 'generalization' of the synfire chain to a functionally feedforward architecture. How such generalizations can be made remains to be understood. Asymmetric Hopfield networks (Kleinfeld and Sompolinsky, 1988; Sompolinsky and Kanter, 1986) might provide insight into how to build such networks. Even though these networks are built with binary neurons, they describe dynamics that evolve from one attractor to another in manner that approximates Markovian behavior, and their connectivity has a

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

feedforward structure in the basis of these attractors. We speculate that the attractor states of the asymmetric Hopfield networks might be analogous to synchronous spiking of neurons in a synfire chain.

One could also consider hybrid architectures, where separate RNNs encode the individual segments of a longer sequence (e.g. song) and are then coupled in a feedforward manner. If learning-induced modifications to the feedforward connections only shift the time at which the next recurrent network starts its activity, this architecture would also function as a temporally flexible pattern generator.

Modifications to timing could also occur by changing the spatiotemporal profile of the inputs to a pattern generator circuit, as opposed to synaptic plasticity within the pattern generator as considered here. This amounts to pushing the problem we have discussed upstream of the pattern generator, since it is now the network providing the input that must ensure that timing is modified flexibly. Interestingly, synfire chains provide a simple way for upstream networks to control timing in a flexible manner. Suppose all neurons in a layer gets a common tonic input from an upstream area, and these inputs are independent from layer to layer. Simply by increasing or decreasing the tonic input a layer gets, without changing its temporal profile, the upstream area can make the neurons in the layer fire earlier or faster, without interfering with other interval durations. Such control of timing by changes in the spatial, but not temporal, profiles of inputs is made possible by the one-to-one mapping of time to space a synfire chain provides.

In this study, we only discussed flexibility in motor timing. However, depending on the structure of the network controlling the behavior, temporal changes could also interfere with spatial aspects of motor output (Buonomano and Laje, 2010). In the songbird system, for example, the output neurons of the time-keeper network, HVC, synapse onto motor cortex

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

analogue neurons that, indirectly, drive vocal muscles (Fee and Scharff, 2010). If modifications to the song's temporal structure change the magnitude of the spiking response in the output neurons, this could change not only the timing of muscle activations but also their magnitude. Thus separating temporal and spatial aspects of movement control (Ali et al., 2013), requires separating the timing of an time-keeper neuron's firing from its magnitude. In general, neither RNNs nor feedforward networks ensure such a separation. However, synfire chains do, suggesting another reason why synfire chains might have the ideal topology and dynamics for implementing the timing of flexible motor behaviors.

Maximal flexibility in modifying motor skills would analogously require the capacity to change specific movement features independently, i.e. without having them interfere with other aspects of motor output. This is not generally ensured. For example, in a reinforcement learning model of a reaching task (Krakauer et al., 2000), interference between different movements were seen (Darshan et al., 2014), and this accounted for a significant slowing down of the learning process. The methods and formalisms we present for understanding the network constrains associated with flexible motor timing can be extended to study networks underlying flexibility in the spatial domain.

Methods

Birdsong Experiments

Animals

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The care and experimental manipulation of all animals were reviewed and approved by the Harvard Institutional Animal Care and Use Committee. Experimental subjects were adult male zebra finches (> 90 days post-hatch, n = 24).

Behavioral Experiments

Experiments to induce adaptive changes to the duration of targeted song segments were done as previously described (Ali et al., 2013). Briefly, we computed the duration of the target segment in real-time using a static threshold crossing of the smoothed amplitude envelope (5 ms boxcar filter with 1 ms advancement). If it did not meet the threshold, a burst of white noise (80-90dB) was played through a loudspeaker for 50-100 ms with a short latency (~1-3 ms). Syllable onsets are associated with rapid increases in sound amplitude, which makes the estimates of their timing more robust to noise. Thus, we mostly targeted 'syllable + gap' segments and estimated the target segments from the onset of the target syllable to the onset of the following syllable. In a typical single target experiment, birds underwent 3-5 days of CAF to lengthen the target followed immediately by 3-5 days of CAF to shorten the target (for a total of 6-10 days for each experimental block) and no CAF for at least 4 days before subsequent experiments.

When comparing learning rates across experiments in which one versus two targets were targeted for modification, we chose targets that were separated by at least 100 ms. We first ran the CAF protocol (n = 3 birds) for target 1, then for target 2, before targeting both intervals in the same experiment. In another group of birds (n = 3), we counterbalanced the order, running the two-target experiments first, followed by single target experiments. Two targets in the same song were modified in opposite directions.

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Birdsong analysis

All analyses of learning rates and variability of segment durations were done offline on "catch trials" during which the white noise feedback was turned off for up to 100 song renditions in the morning (AM) and again approximately 8 hours later in the evening (PM). The method for obtaining estimates of song segment durations has previously been described (Ali et al., 2013). To calculate learning rates, we computed the change in target duration from the start to the end of a CAF run (up or down) and divided it by the number of intervening days. We compared both morning and afternoon catch trials, averaging across them to obtain a more robust estimate of learning rate. Variability was measured as standard deviation (SD) of the distribution of interval durations while coefficient of variation (CV), as is standard, was defined as the SD divided by the mean duration of the same distribution. We compared variability for the same time period in the day in order to rule out potential circadian effects.

To compare changes in segment durations during CAF to baseline drift, we also analyzed the duration of song segments during days (6-10 days) when no CAF was imposed. We subtracted the absolute value of this baseline drift from the CAF-induced changes to obtain a better estimate of learning. We did the same for non-target interval, discarding the signs for each non-target interval change when averaging within a bird to rule out different signs negating each other that can potentially mask major non-target effects.

Statistical tests.

All statistics presented in the main text refer to mean ± SD, while error bars in the figures all represent SEM. For the birdsong analysis, we used parametric statistical tests as noted in the main text.

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Computational Methods

Feedback-stabilized RNNs

Network setup and training: We simulated recurrent neural networks with $N = 500$ neurons. The neural dynamics is described by,

$$\tau \frac{dx_i}{dt} = -x_i + \sum_{j=1}^{N} W_{ij}\, r_j(t) + \sum_{j=1}^{2} W_{ij}^{\mathrm{In}} y_j(t) + g_{FB} W_i^{\mathrm{FB}} z(t) + \sqrt{\tau_\eta}\sigma\eta_i(t),$$

$$r_i(t) = \tanh\big(x_i(t)\big),$$

$$z(t) = \sum_{j=1}^{N} W_{ij}^{\mathrm{Out}} r_j(t),$$

where $i = 1, \ldots, N$, $r_i(t)$ are the firing rates of the neurons, $y_j(t)$ are the inputs and $z(t)$ is the output. The neural time constant is $\tau = 10$ ms. Recurrent synaptic connectivity matrix, $\boldsymbol{W}$, is sparse: W_{ij} is 0 with probability $1 - p$, with $p = 0.1$. Non-zero elements of $\boldsymbol{W}$ were drawn independently from a Gaussian distribution with 0 mean and variance $1.5/(pN)$, ensuring that the network is in the chaotic regime without the feedback from the output (Sompolinsky et al., 1988). Elements of $\boldsymbol{W}^{\mathbf{In}}$ and $\boldsymbol{W}^{\mathbf{FB}}$ are independently drawn from a uniform distribution between -1 and 1. The strength of feedback g_{FB} was varied, as discussed in the main text. Unless stated otherwise, the networks were simulated with independently injected noise to neurons, including during training: $\eta_i(t)$ is a zero-mean Gaussian white noise with covariance $< \eta_i(t)\eta_i(t') >= \delta(t - t')$, $\tau_\eta = 10$ ms and $\sigma = 0.01$.

There are two types of inputs: 1) $y_1(t)$ is a unit pulse of duration 50 ms and amplitude 5. This input was delivered for all simulations. The end of the pulse marked the beginning of the first interval. 2) $y_2(t)$ is a unit pulse of duration 10 ms and amplitude P that was varied. It

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

arrived 120 ms after the start of the first interval. This input was used only when assessing the robustness of the network to perturbations (Figure 2D).

The desired output was a 530 ms waveform (Supplementary Figure 1A) constructed by:

(1) adding 10 Gaussian waveforms centered 50 ms apart with widths (standard deviation) of 10 ms.

(2) normalizing the waveform such that its maximum is 1 and minimum is 0.1, and

(3) shifting it in time such that the first threshold crossing (which is chosen to be 0.68) occurred 50 ms after the first interval.

Hence the desired output marked 10 equally spaced 50 ms intervals. $\boldsymbol{W}^{\mathbf{In}}$ was trained using the FORCE algorithm (Sussillo and Abbott, 2009) for 30 training trials. After training, 10 test runs were performed and the error between the network output and the desired output ('test error') calculated for each test run. The test error was defined by $\frac{\sqrt{\int_0^{530\text{ ms}}\left(z_{des}(t)-z(t)\right)^2 dt}}{\sqrt{\int_0^{530\text{ ms}} z_{des}(t)^2 dt}}$, where $z_{des}(t)$ is the desired output.

The dynamical equations governing the network dynamics were integrated with a first-order Euler method. The integration step size was $dt = 0.1$ ms, unless stated otherwise.

For Figure 2C, the initial state distance between two intervals was calculated by first making two *N=500* dimensional vector of the instantaneous firing rates of the network neurons at the beginning of the two intervals and then calculating the L_2-norm of the difference of these vectors.

Learning matrix: To calculate the learning matrices of trained fsRNNs, we calculated the gradients of interval durations numerically: one-by-one, each non-zero synaptic weight was increased by $dW = 0.05$ and the change in interval durations measured. In these simulations,

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

$\sigma = 0$ and $dt = 0.01$ ms. Our results were verified by running the same calculation with varying dt and dW parameters for some of the networks.

Reinforcement learning: We implemented a modified version of the synaptic plasticity rule of (I. R. Fiete et al., 2007; Fiete and Seung, 2006) to change non-zero elements of $\boldsymbol{W}$. Specifically, changes to synaptic weight W_{ij} is given by

$$\Delta W_{ij} = \gamma \int_0^{t_K} R(t) e_{ij}(t)\, dt,$$

where t_K is the end of the last interval, $\gamma = 0.004$, $e_{ij}(t)$ is the eligibility trace and $R(t)$ is the reward. γ is chosen to match learning rates observed in experiments with songbirds. The eligibility trace is given by

$$e_{ij}(t) = \int_0^t \frac{dt'}{\tau_e}\, e^{-(t-t')/\tau_e} \eta_i(t') r_j(t'),$$

where $\tau_e = 35$ ms. The reward signal is given by $R(t) = R^{tar_1}\, \delta\left(t - t_{tar_1}\right)$ in a single-target interval 'experiment' and by $R(t) = R^{tar_1}\, \delta\left(t - t_{tar_1}\right) + R^{tar_2} \delta\left(t - t_{tar_2}\right)$ in a two-target intervals 'experiment', where t_{tar_1} and t_{tar_2} are the timings of ends of 1st and 2nd target intervals respectively, and R^{tar_1} and R^{tar_2} are 0-1 rewards contingent on 1st and 2nd target intervals respectively. If the interval is not targeted, the reward is always 0. If the interval is targeted for modification, the rewards are calculated by comparing the interval duration in the current trial, I^{tar}, to the running average of the target interval duration, $\bar{I}^{tar}$, which is updated in each trial as follows: $\bar{I}^{tar} \leftarrow 0.995\, \bar{I}^{tar} + 0.005 I^{tar}$. If the interval is targeted for stretching, the reward is 1 if $I^{tar} > \bar{I}^{tar}$, and 0 otherwise. If the interval is targeted for shrinking, the reward is 1 if $I^{tar} < \bar{I}^{tar}$, and 0 otherwise.

Interval duration changes during a reinforcement learning experiment are calculated by a running (200 point window) across trials and subtracting from them their baseline values. The

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

baseline values are averages of the interval durations across 400 trials where no reinforcement is delivered.

Dynamic Attractor

Network setup and training: The architecture of this network, the dynamics of its neurons and the numerical integration of the dynamical equations, its connectivity and noise parameters, and its input and output waveforms are the same as the fsRNNs, except the feedback from the output was set to zero, i.e. $g_{FB} = 0$.

The networks are trained using the 'innate training' procedure of (Laje and Buonomano, 2013), which consists of two stages:

(1) 70% of non-zero elements of $\boldsymbol{W}$ are trained using the FORCE algorithm (30 training trials) to stabilize an innately produced trajectory of duration 530 ms. See (Laje and Buonomano, 2013) for details.

(2) $\boldsymbol{W}^{\mathbf{In}}$ is trained using the FORCE algorithm to produce the desired output waveform, exactly as in fsRNNs.

Learning matrix: The procedures for calculating the learning matrices were the same as for the fsRNNs.

Reinforcement learning: The procedures for reinforcement learning was the same as the fsRNNs, except that $\gamma = 0.004/3$.

Synfire Chain

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Network setup and training: Our synfire chain model consisted of 2,100 integrate-and-burst neurons organized into 1400 layers, with 15 neurons in each layer (Figure 4A). Neurons in a layer projected to all neurons in the next layer, forming a chain topology. The subthreshold membrane potential of the i^{th} neuron, V_i, obeys:

$$C\frac{dV_i}{dt} = -g_L(V_i - V_L) + I_{syn,i}(t) + I_{ext,i}(t) + \sqrt{\tau_\eta}\sigma\eta_i(t),$$

where $C = 1\ \mu\mathrm{F/cm^2}$ is the membrane time constant, $g_L = 0.4\ \mathrm{mS/cm^2}$ is the leak conductance, $V_L = -60\ \mathrm{mV}$ is the leak potential, $\eta_i(t)$ is a zero-mean Gaussian white noise with covariance $< \eta_i(t)\eta_i(t') >= \delta(t - t')$, $\tau_\eta = 10\ \mathrm{ms}$ and $\sigma = 260\ \mathrm{nA/cm^2}$. The synaptic input is given by $I_{syn,i}(t) = \sum_j W_{ij} \sum_k \varepsilon\left(t - t_j^k\right)$, where t_j^k denotes the k^{th} spike of j^{th} neuron, $\varepsilon(t) = \Theta(t)e^{-t/\tau_s}$ with $\Theta(t)$ being the step function and $\tau_s = 5\mathrm{ms}$, and W_{ij} is $271.8\ \mathrm{nA/cm^2}$ for synapses from a neuron to the neurons in the next layer and 0 otherwise. When the membrane potential of the integrate-and-burst neuron reaches threshold, $V_{th} = -50\ \mathrm{mV}$, the neuron emits 4 spikes with 2 ms intervals and the membrane potential is reset to , $V_R = -55\ \mathrm{mV}$ after a refractory period of 4 ms. Chain propagation was started by $I_{ext,i}(t)$, a 5 ms pulse input with magnitude $6.7\ \mu\mathrm{A/cm^2}$ applied only to the neurons in the first node.

To map network dynamics to timing, an integrate-and-fire readout neuron was connected to all the neurons in every 14th layer, making a total of 10 readout neurons. Their dynamics were governed by the same equations as the chain neurons, except they did not receive any external input or noise. They synaptic weights were $266.7\ \mathrm{nA/cm^2}$. The first spikes of the output neurons were taken to mark interval boundaries. The first interval is taken to start with the start of the external pulse.

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The dynamical equations governing the network dynamics were integrated with a first-order Euler method. The integration step size was $dt = 0.1$ ms, unless stated otherwise.

Learning matrix: To calculate the learning matrices of the synfire chain, we calculated the gradients of interval durations numerically: one-by-one, each non-zero synaptic weight was increased by $dW = 27.18\ \mu\mathrm{A/cm^2}$ and running the change in interval durations were measured. In these simulations, $\sigma = 0$ and $dt = 0.01$ ms. Our results were verified by running the same calculation with varying dt and dW parameters for some of the networks.

Reinforcement learning: The procedures for reinforcement learning was the same as the fsRNNs, except that $\gamma = 956.7\ \mathrm{nA/cm^2}$ and the eligibility trace is given by

$$e_{ij}(t) = \int_0^t \frac{dt'}{\tau_e}\, e^{-(t-t')/\tau_e} \eta_i(t') s_{ij}(t'),$$

where $\tau_e = 35$ ms, and $s_{ij}(t) = \frac{1}{\tau_s}\sum_k \varepsilon\left(t - t_j^k\right)$. The synaptic weights were not allowed to increase above $341.8\ \mathrm{nA/cm^2}$ nor fall below $216.8\ \mathrm{nA/cm^2}$ for stability of chain propagation: weights above the upper bound led to explosion, and weights below the lower bound led to activity propagation terminating before reaching the end of the chain.

Statistics

In figures, all reported error bars are standard deviations over different trained networks unless stated otherwise.

Figure legends

Figure 1: Flexible modification of motor timing in songbirds

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(A) Schematic showing different ways in which the temporal structure of a motor sequence (made from motor segments m1-m6) can change. (Left). Global change – duration of all segments change together. (Middle) Interference – Modifications to the duration of one segment (m3) interferes with (i.e. changes) the duration of others. (Right) Specificity – each segment can change independently of others. **(B)** (Top) Spectrogram of a zebra finch song, divided into four segments. One of these ('target') was targeted for lengthening/shortening by delivering aversive noise bursts after the end of the target segment when its duration was below/above a threshold (Bottom, see Methods). **(C)** Duration distributions for the segments in (B) when the target segment was lengthened by our CAF paradigm. **(D)** Summary statistics (mean ± SEM) of the daily changes to the duration of target and non-target segments (N=18 targeted segments and n=120 non-targeted segments in 18 birds). Changes in target intervals differed significantly from zero ($p = 1.6 \times 10^{-6}$, one-sample t-test), whereas non-targets did not (p=0.96, one-sample t-test). **(E)** Spectrogram of a song for which two segments were targeted for modification, either alone or in conjunction with each other. **(F)** Learning trajectories (mean ± SEM) during CAF experiments designed to change the duration of the targets in (E) either alone or together. Arrows denote the direction in which the duration of the targets were induced to change by the CAF paradigm. **(G)** Summary statistics (mean ± SEM) of the daily changes to the duration of targeted segments, as a function of whether they were targeted alone or together with another segment (n=12 segments in 6 birds, p = 0.54, paired-samples t-test).

Figure 2: Feedback-stabilized RNNs exhibit a robustness-flexibility trade-off

(A) (Left) Schematic showing the network architecture. Trained connections are marked in red. (Middle, right) Firing rates of example neurons pattern generator networks with different

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

feedback strengths, and the output of those networks. **(B)** Training performance (test error, Methods) and timing failure rate, i.e. the frequency with which the network fails to produce all 10 intervals within 6% of their target durations (calculated for 400 trials), shown as a function of feedback strength. **(C)** The distance between two firing rate vectors, each constructed from the instantaneous firing rates of all neurons in the network at the beginning of an interval (Methods), averaged over all interval pairs for intervals 3 to 10, 20 trained fsRNNs and 100 runs over each fsRNN. A small distance means that the firing rate vectors are more similar to each other at the beginning of intervals, signaling periodic activity. **(D)** Timing failure rate in response to a perturbation pulse (P) of varying magnitudes delivered to the network, plotted as a function of the feedback strength. Error bars show standard error across 20 trained fsRNNs. **(E)** Example learning matrices for fsRNNS with two different feedback strengths, one for a value where the network activity is periodic (g_{FB}=2) and one not periodic (g_{FB}= 1). **(F)** Relative change in the duration of non-targeted intervals as a function of the change in targeted intervals ('Average Interference'), calculated from the off-diagonal elements of the learning matrix. The plot shows averages over all pairings of intervals 2 to 10 for 20 trained fsRNNs at each feedback strength. **(G)** Examples of a reinforcement learning 'experiment' that targeted interval 3 for lengthening. Shown is the effect on all intervals after for 3000 trials for fsRNN with different feedback strengths (Methods). **(H)** Learning rate, defined as the absolute value of the change in the target interval duration after 1000 trials, averaged over 20 learning 'experiments' in 20 trained networks. **(I)** Interval duration changes (Methods) after 1000 trials for the 'experiments' in G, averaged across 20 different simulations, regressed to the elements of the 3rd row (i.e. those associated with the target interval) of the networks' learning matrices. Error bars show standard error across 20 simulations. **(J)** The analysis of **H** applied to 100 trained fsRNNs (20 for each of g_{FB}=0.5,1,1.5,2,3), 2 times each, one for stretching the 3rd interval and one for shrinking the 3rd

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

interval. Plotted is the median correlation coefficients over the 200 regressions performed. Error bars denote 25th and 75th percentiles respectively. **(K)** Change in non-targeted intervals in the first 1000 trials of reinforcement learning simulations, relative to the change in the target interval (interference). Average over all non-target intervals and 20 simulations across 20 trained fsRNNs as a function of feedback strength. The blue portion of the bars show the average interference calculated from the learning matrix (as in H).

Figure 3: Flexible timing with dynamic attractor networks

(A) Schematic showing the network architecture. Recurrent connections as well as the connections to the output neuron are trained. On the right are the activities of example neurons from the pattern generator network and the output neuron. **(B)** Box plots of training (test error, Methods) and timing (timing failure rate out of 400 trials) performances of 20 trained networks. Whiskers show maxima and minima. 14 out of 20 networks were successfully trained and used for the subsequent analysis in this figure. **(C)** Timing failure rates (out of 400 trials) of the network when perturbed by a pulse input of varying magnitude. **(D)** Example learning matrix. **(E)** Relative change in the duration of non-targeted intervals as a function of the change in targeted intervals ('Average Interference'), calculated from the off-diagonal elements of the learning matrix, as in Figure 2F except that the 1st interval is also taken into account. The average interference is compared to the interference in intervals adjacent to the target, and for fsRNNs with g_{FB}=1. **(F)** Example of a reinforcement learning 'experiment' that targeted interval 3 for lengthening. Shown is the effect on all intervals after for 3000 trials. **(G)** Learning rate averaged across networks as in Figure 2I. **(H)** Median correlation coefficient between the interval changes calculated from the learning matrix and from the reinforcement learning simulations (as in Figure 2I). Error bars denote 25th and 75th percentiles respectively. **(I)** Average interference in

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

reinforcement learning simulations as in Figure 2K. The blue portion of the bars show the average interference calculated from the learning matrix

Figure 4: Flexible timing with synfire chains

(A) Schematic showing the network architecture and interval (I) time readout. **(B)** Example run of the synfire chain. Only one neuron per layer is shown. On bottom is the activity of the read-out neurons, defining the interval boundaries. **(C)** The learning matrix of the synfire chain. **(D)** Examples of a reinforcement learning 'experiment' that targeted interval 5 for lengthening. Shown is the effect on all intervals after for 5000 trials. **(E)** Correlation coefficient between the interval changes calculated from the learning matrix and from the reinforcement learning simulations (as in Figure 2I). **(F)** Average interference per interval in reinforcement learning simulations, averaged across 20 learning simulations. Error bar shows standard deviation across simulations.

Figure 5: Changes in temporal variability with learning

(A) (Top) Spectrogram of a zebra finch song showing the target segment. (Middle) Histograms of mean durations of a target segment after it had been shortened and lengthened respectively using our CAF paradigm. Mean at baseline denoted by the dotted line. **(B)** Scatterplot of variability (standard deviation) as a function of mean segment duration for the same example as (A) across 9 days of CAF. Black line is the linear fit to the data. **(C)** Scatterplot of percent change in normalized variability (CV) as a function of percent change in mean segment duration (both relative to baseline before CAF). Black line is the linear fit to the data. **(D)** Standard deviation of target duration as a function of the target's mean duration for fsRNNs with g_{FB}=1. Target intervals were lengthened and shortened from baseline for 3000 trials each. Data from 200

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

'catch' (i.e. non-reinforced) trials interspersed between every 1000 trials, averaged across 20 simulations. Example from four fsRNNs are shown. Error bars are standard deviations across simulations. (Bottom) Box plot of all correlation coefficients between interval variability and interval duration, across 20 networks, 2 target intervals each (3^{rd} and 8^{th}). **(E)** Same as D for dynamic attractor networks. Box plot of all correlation coefficients across 14 networks, 1 target interval each (3^{rd}). **(F)** Same as D, but for synfire chains. Data from 200 'catch' trials interspersed between every 1000 trials, averaged across 20 simulations. Target interval was the 3^{rd}.

Supplementary Figure 1: Flexibility and robustness of fsRNNs

(A) The signal that the fsRNN and dynamic attractor networks were trained to produce. **(B)** An example of what happens when the network is perturbed, and fails to produce the right timing interval. Black shows a successful output when no perturbation is delivered. Dashed blue line shows the output when perturbation was delivered to the network during the time denoted by the brown bar. Even though the output crossed the threshold sufficiently many times, the interval durations were not within the desired 6% of their targets. **(C)** Example reinforcement learning simulations for a two-target 'experiments' run for 3000 trials in a fsRNN, where the 3^{rd} and 8^{th} interval was targeted for lengthening and shortening respectively. Shown for different feedback strengths. **(D)** Change in average learning rates of the 3^{rd} and 8^{th} intervals (across 20 simulations) when they are targeted together relative to when they are targeted alone. For each network, average learning rates (across 20 simulations) were calculated (as for Figure 2J) for both intervals in single-target and two-target 'experiments'. Error bars are standard deviations across 20 networks. **(E)** Reduction in learning rate is well predicted by the learning matrices. Each dot represents a network, pooled over 100 trained fsRNNs (20 for each of

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

g_{FB}=0.5,1,1.5,2,3), 2 intervals each (3rd and 8th). Reduction in average learning rate (across 20 simulations) plotted against $\frac{M_{\mu\nu}}{M_{\mu\mu}} \times 100\%$ (see equation 4 of main text) for that network.

Supplementary Figure 2: Two-target reinforcement learning in a dynamic attractor

(A) Example reinforcement learning simulations for a two-target 'experiments' run for 3000 trials in a dynamic attractor network, where the 3rd and 8th interval was targeted for lengthening and shortening respectively. **(B)** Change in average learning rates of the 3rd and 8th intervals (across 20 simulations) when they are targeted together relative to when they are targeted alone. Error bars are standard deviations across 14 networks. **(C)** Reduction in learning rate is well predicted by the learning matrices. Each dot represents a network, pooled over 100 trained networks, 2 intervals each (3rd and 8th). Reduction in average learning rate (across 20 simulations) plotted against $\frac{M_{\mu\nu}}{M_{\mu\mu}} \times 100\%$ (see equation 4 of main text) for that network.

Supplementary Figure 3: Two-target reinforcement learning in a synfire chain

(A) Example reinforcement learning 'experiment' for two-target 'experiments' run for 5000 trials in a synfire chain. 5th interval is lengthened and the 8th interval is shortened. **(B)** Change in average learning rates of the 5th and 8th intervals (across 20 simulations) when they are targeted together relative to when they are targeted alone. Error bars are standard deviations across 20 simulations. For comparison, learning rate changes for the fsRNNs and the dynamic attractor are also provided.

References

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abeles, M., 1991. Corticonics: Neural Circuits of the Cerebral Cortex, First. ed. Cambridge University Press.

Ali, F., Otchy, T.M., Pehlevan, C., Fantana, A.L., Burak, Y., Olveczky, B.P., 2013. The Basal Ganglia is necessary for learning spectral, but not temporal, features of birdsong. Neuron 80, 494–506. doi:10.1016/j.neuron.2013.07.049

Boström, K.J., Wagner, H., Prieske, M., de Lussanet, M., 2013. Model for a flexible motor memory based on a self-active recurrent neural network. Human Movement Science 32, 880–898. doi:10.1016/j.humov.2013.07.003

Brainard, M.S., Doupe, A.J., 2013. Translating birdsong: songbirds as a model for basic and applied medical research. Annu. Rev. Neurosci. 36, 489–517. doi:10.1146/annurev-neuro-060909-152826

Buonomano, D.V., Laje, R., 2010. Population clocks: motor timing with neural dynamics. Trends in cognitive sciences 14, 520–527.

Cai, S., Ghosh, S.S., Guenther, F.H., Perkell, J.S., 2011. Focal Manipulations of Formant Trajectories Reveal a Role of Auditory Feedback in the Online Control of Both Within-Syllable and Between-Syllable Speech Timing. J. Neurosci. 31, 16483–16490. doi:10.1523/JNEUROSCI.3653-11.2011

Carter, M.C., Shapiro, D.C., 1984. Control of sequential movements: evidence for generalized motor programs. Journal of Neurophysiology 52, 787–796.

Darshan, R., Leblois, A., Hansel, D., 2014. Interference and Shaping in Sensorimotor Adaptations with Rewards. PLoS Comput Biol 10, e1003377. doi:10.1371/journal.pcbi.1003377

Doupe, A.J., Kuhl, P.K., 1999. Birdsong and human speech: common themes and mechanisms. Annual Review of Neuroscience 22, 567–631.

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Doya, K., Sejnowski, T., 1995. A novel reinforcement model of birdsong vocalization learning. Advances in Neural Information Processing Systems 7, 101–108.

Fee, M.S., Scharff, C., 2010. The songbird as a model for the generation and learning of complex sequential behaviors. ILAR J 51, 362–377.

Fiete, I.R., Fee, M.S., Seung, H.S., 2007. Model of Birdsong Learning Based on Gradient Estimation by Dynamic Perturbation of Neural Conductances. Journal of Neurophysiology 98, 2038–2057. doi:10.1152/jn.01311.2006

Fiete, I.R., Fee, M.S., Seung, H.S., 2007. Model of Birdsong Learning Based on Gradient Estimation by Dynamic Perturbation of Neural Conductances. Journal of Neurophysiology 98, 2038–2057. doi:10.1152/jn.01311.2006

Fiete, I.R., Hahnloser, R.H.R., Fee, M.S., Seung, H.S., 2004. Temporal Sparseness of the Premotor Drive Is Important for Rapid Learning in a Neural Network Model of Birdsong. Journal of Neurophysiology 92, 2274–2282. doi:10.1152/jn.01133.2003

Fiete, I., Seung, H., 2006. Gradient Learning in Spiking Neural Networks by Dynamic Perturbation of Conductances. Phys. Rev. Lett. 97. doi:10.1103/PhysRevLett.97.048104

Fortune, E.S., Rodríguez, C., Li, D., Ball, G.F., Coleman, M.J., 2011. Neural mechanisms for the coordination of duet singing in wrens. Science 334, 666–670. doi:10.1126/science.1209867

Garst-Orozco, J., Babadi, B., Ölveczky, B.P., 2015. A neural circuit mechanism for regulating vocal variability during song learning in zebra finches. eLife Sciences 3, e03697. doi:10.7554/eLife.03697

Gentner, D.R., 1987. Timing of skilled motor performance: Tests of the proportional duration model. Psychological Review 94, 255–276. doi:10.1037/0033-295X.94.2.255

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Goel, A., Buonomano, D.V., 2014. Timing as an intrinsic property of neural networks: evidence from in vivo and in vitro experiments. Phil. Trans. R. Soc. B 369, 20120460. doi:10.1098/rstb.2012.0460

Goldman, M.S., 2009. Memory without Feedback in a Neural Network. Neuron 61, 621–634. doi:10.1016/j.neuron.2008.12.012

Hahnloser, R.H.., Kozhevnikov, A.A., Fee, M.S., 2002. An ultra-sparse code underliesthe generation of neural sequences in a songbird. Nature 419, 65–70.

Hoerzer, G.M., Legenstein, R., Maass, W., 2014. Emergence of Complex Computational Structures From Chaotic Neural Networks Through Reward-Modulated Hebbian Learning. Cereb. Cortex 24, 677–690. doi:10.1093/cercor/bhs348

Immelmann, K., 1969. Song development in the zebra finch and other estrildid finches, in: Bird Vocalizations (ed. Hinde, R.A.). Cambridge University Press, pp. 61–74.

Jaeger, H., Haas, H., 2004. Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication. Science 304, 78–80. doi:10.1126/science.1091277

Kleinfeld, D., Sompolinsky, H., 1988. Associative neural network model for the generation of temporal patterns. Theory and application to central pattern generators. Biophysical Journal 54, 1039.

Krakauer, J.W., Pine, Z.M., Ghilardi, M.-F., Ghez, C., 2000. Learning of Visuomotor Transformations for Vectorial Planning of Reaching Trajectories. J. Neurosci. 20, 8916–8924.

Laje, R., Buonomano, D.V., 2013. Robust timing and motor patterns by taming chaos in recurrent neural networks. Nat. Neurosci. 16, 925–933. doi:10.1038/nn.3405

Long, M.A., Fee, M.S., 2008. Using temperature to analyse temporal dynamics in the songbird motor pathway. Nature 456, 189–194. doi:10.1038/nature07448

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Long, M.A., Jin, D.Z., Fee, M.S., 2010. Support for a synaptic chain model of neuronal sequence generation. Nature 468, 394–399. doi:10.1038/nature09514

Rajan, K., Abbott, L.F., Sompolinsky, H., 2010. Stimulus-dependent suppression of chaos in recurrent neural networks. Phys. Rev. E 82, 011903. doi:10.1103/PhysRevE.82.011903

Sompolinsky, H., Crisanti, A., Sommers, H.J., 1988. Chaos in Random Neural Networks. Phys. Rev. Lett. 61, 259–262. doi:10.1103/PhysRevLett.61.259

Sompolinsky, H., Kanter, I., 1986. Temporal Association in Asymmetric Neural Networks. Phys. Rev. Lett. 57, 2861–2864. doi:10.1103/PhysRevLett.57.2861

Sternberg, S., Knoll, R.L., Monsell, S., Wright, C.E., 1988. Motor Programs and Hierarchical Organization in the Control of Rapid Speech. Phonetica 45, 175–197. doi:10.1159/000261825

Sternberg, S., Knoll, R.L., Turock, D.L., 1990. Hierarchical control in the execution of action sequences: Tests of two invariance properties, in: In M. Jeannerod (Ed.), Attention and Performance XIII. Erlbaum, pp. 3–55.

Sussillo, D., Abbott, L.F., 2012. Transferring Learning from External to Internal Weights in Echo-State Networks with Sparse Connectivity. PLoS ONE 7, e37372. doi:10.1371/journal.pone.0037372

Sussillo, D., Abbott, L.F., 2009. Generating Coherent Patterns of Activity from Chaotic Neural Networks. Neuron 63, 544–557. doi:10.1016/j.neuron.2009.07.018

Trofimovich, P., Baker, W., 2007. Learning prosody and fluency characteristics of second language speech: The effect of experience on child learners' acquisition of five suprasegmentals. Applied Psycholinguistics 28, 251–276. doi:10.1017/S0142716407070130

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Tumer, E.C., Brainard, M.S., 2007. Performance variability enables adaptive plasticity of "crystallized" adult birdsong. Nature 450, 1240–1244. doi:10.1038/nature06390

Viviani, P., Laissard, G., 1991. Timing control in motor sequences, in: Jacqueline Fagard and Peter H. Wolff (Ed.), Advances in Psychology, The Development of Timing Control and Temporal Organization in Coordinated Action Invariant Relative Timing, Rhythms and Coordination. North-Holland, pp. 1–36.

Williams, R.J., 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. Mach Learn 8, 229–256. doi:10.1007/BF00992696

Woolley, S.C., Doupe, A.J., 2008. Social context-induced song variation affects female behavior and gene expression. PLoS Biol 6, e62.

Yamashita, Y., Takahasi, M., Okumura, T., Ikebuchi, M., Yamada, H., Suzuki, M., Okanoya, K., Tani, J., 2008. Developmental learning of complex syntactical song in the Bengalese finch: A neural network model. Neural Networks 21, 1224–1231. doi:10.1016/j.neunet.2008.03.003

ACKNOWLEDGEMENTS

This work was supported by a grant from NINDS (R01 NS066408), a McKnight Scholar Award and Klingenstein Fellowship to BPÖ, and partly by a Swartz Foundation post-doctoral fellowship to CP.

Author contributions

BPÖ, CP, and FA designed the study. CP performed the theoretical analysis and simulations with input from the other authors. FA performed the birdsong experiments and analyzed the data

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

with input from the other authors. BPÖ supervised all aspects of the study. BPÖ, CP, and FA wrote the paper.

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

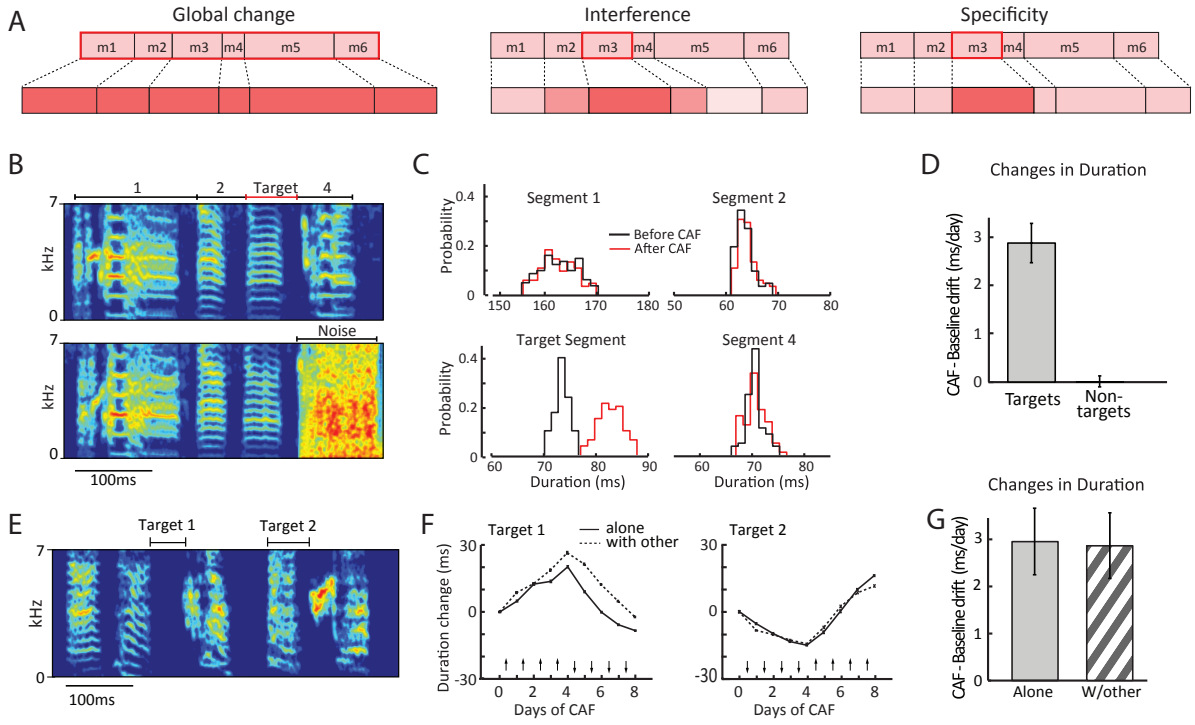


Figure 1
Pehlevan et. al 2015

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

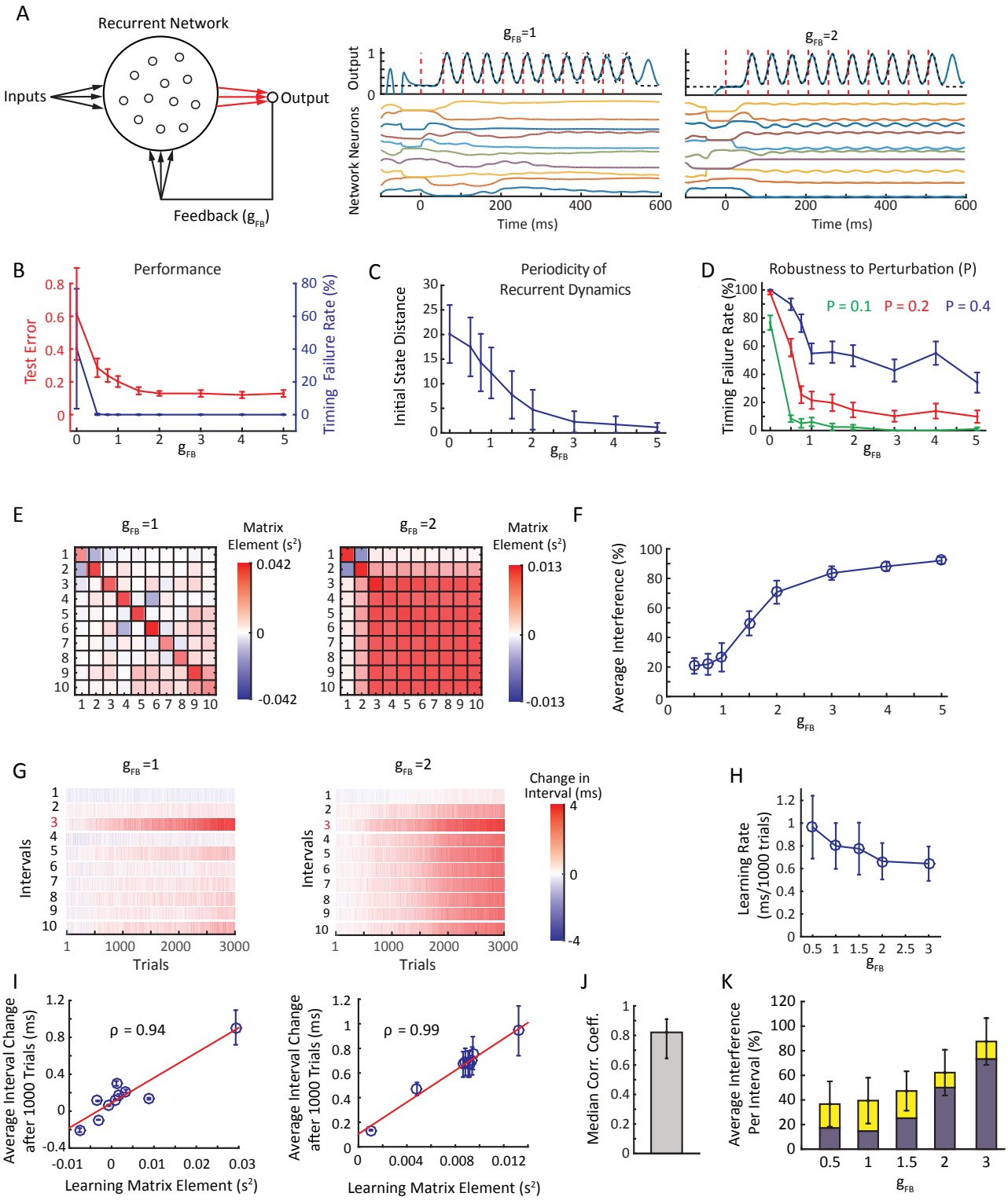


Figure 2
Pehlevan et. al 2015

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

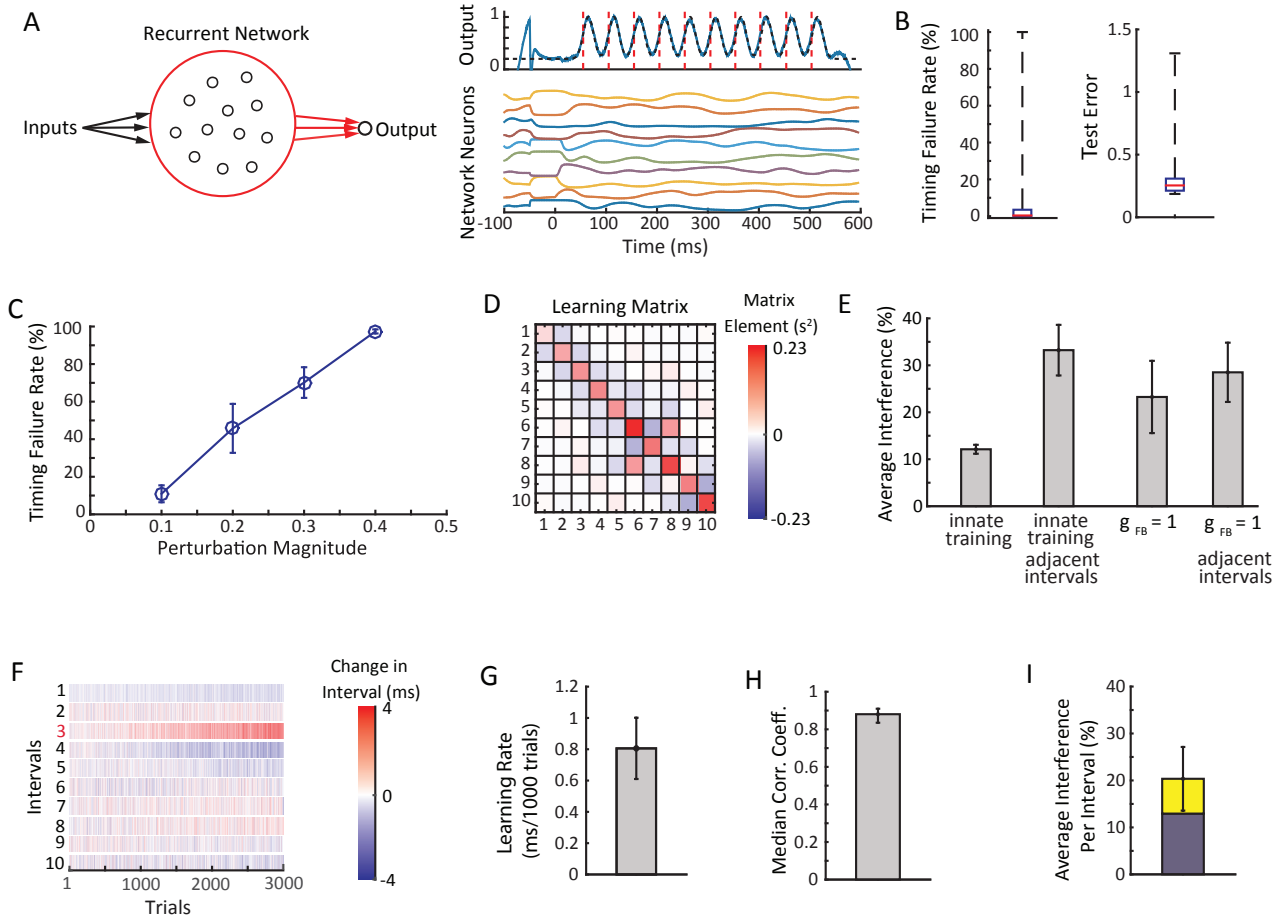


Figure 3
Pehlevan et. al 2015

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

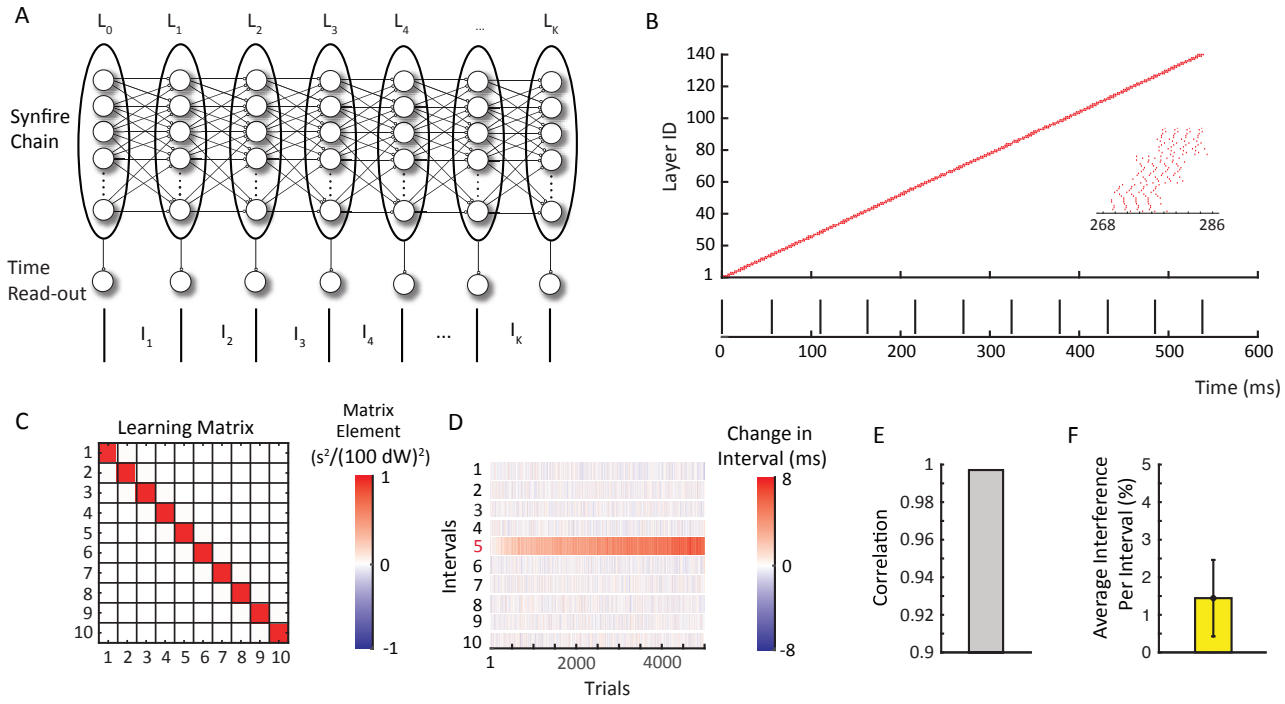


Figure 4
Pehlevan et. al 2015

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

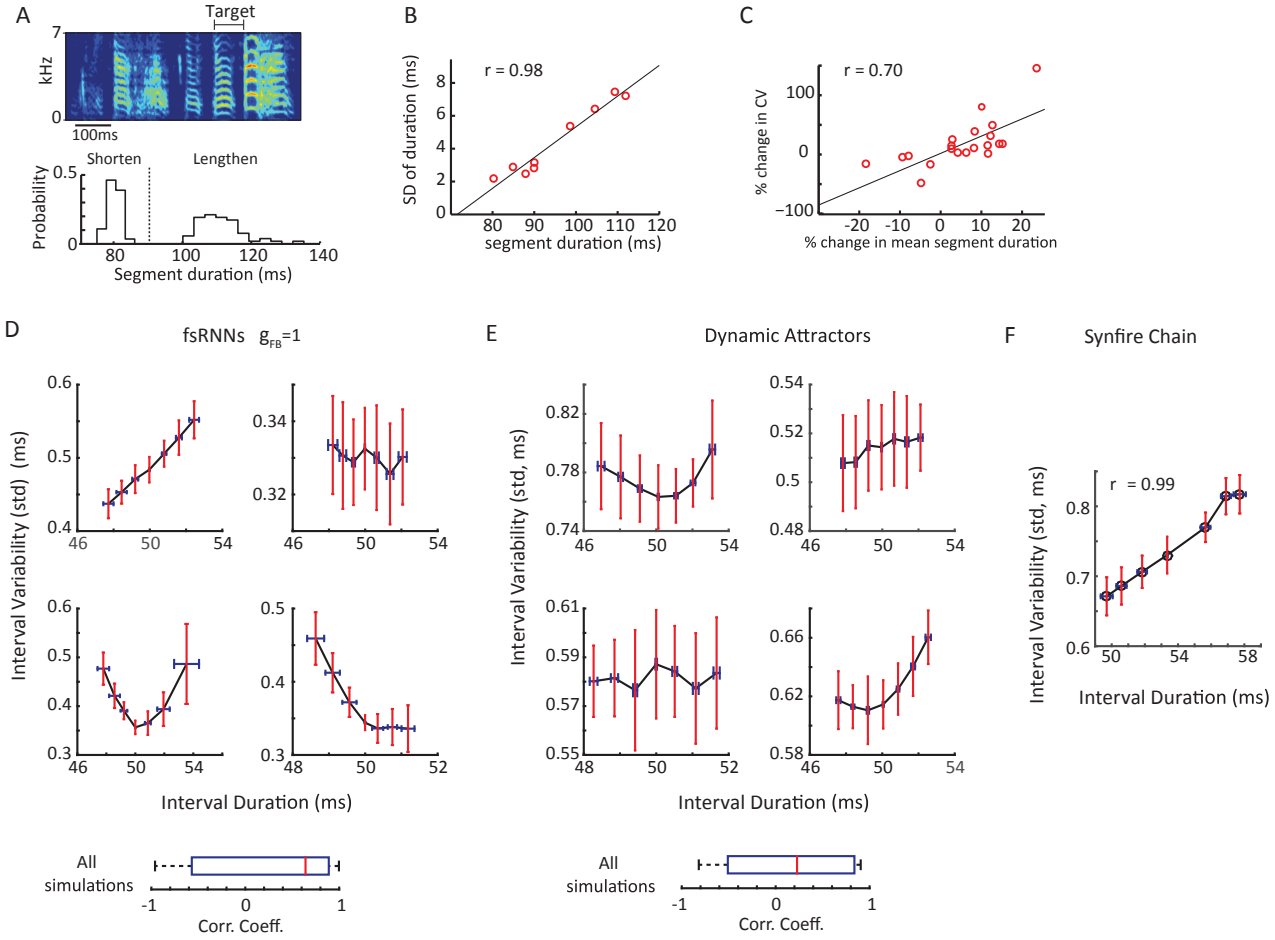
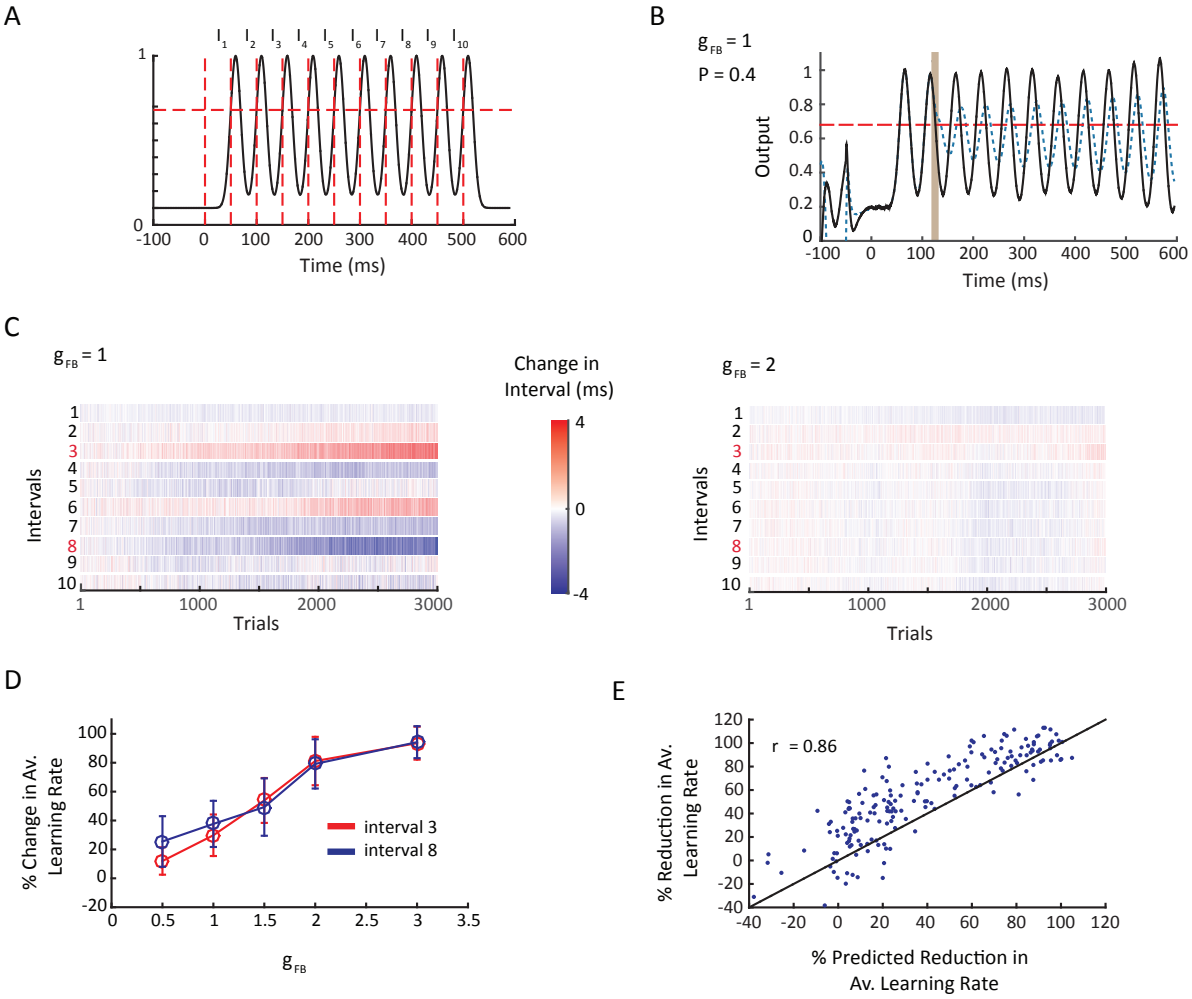


Figure 5
Pehlevan et. al 2015

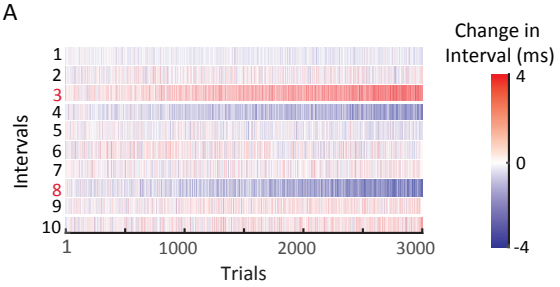
bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



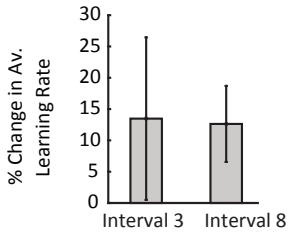
Supplementary Figure 1
Pehlevan et. al 2015

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

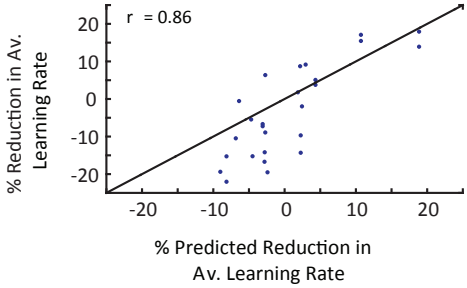
A



B

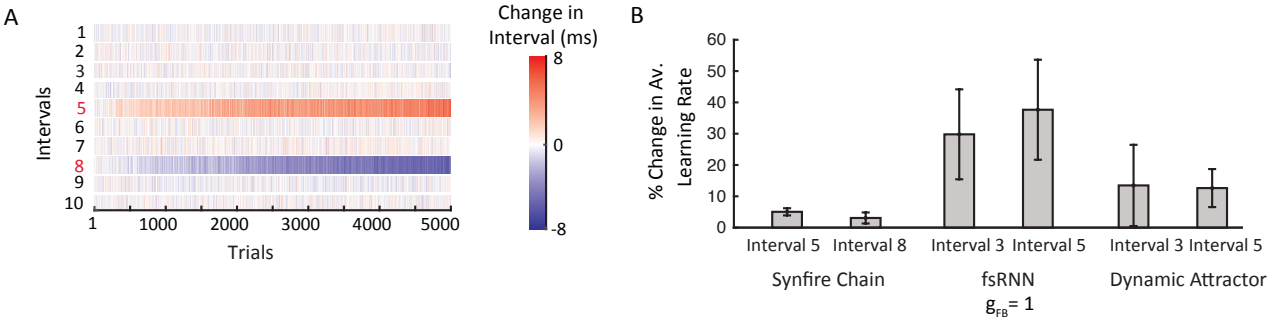


C



Supplementary Figure 2
Pehlevan et. al 2015

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



Supplementary Figure 3
Pehlevan et. al 2015

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Information

Cengiz Pehlevan, Farhan Ali, Bence Ölveczky

December 1, 2015

Contents

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

1 Supplementary Note I: Flexible modifications to temporal patterns by changes in synaptic weights

Here we lay out the general conditions under which a pattern generator network can perform specific and independent modifications to the temporal structure of its output by changes to its synaptic weights.

We consider a pattern generator network with N neurons. Let W_{ij} be the strength of synaptic connections from j^{th} to i^{th} neuron. The network produces K intervals, I_α, with $\alpha = 1, \ldots, K$. These intervals are functions of the synaptic connectivity matrix, a fact which we denote by the notation $I_\alpha(\mathbf{W})$.

1.1 Specific trajectories in synaptic weight space

Specificity of temporal modifications requires the existence of trajectories in synaptic weight space along which only a single interval monotonically changes. Suppose an interval β is to be modified specifically. Is there a trajectory, $\mathbf{W}^\beta(l)$, in synaptic weight space along which only interval β changes monotonically? We choose our parametrization such that $l = 0$ is the original network's synaptic weight matrix and $l > 0$ parametrizes a slowing down of the interval, without loss of generality. We assume that along these trajectories no new synapses are formed, i.e. synapses with zero weight remain with weight zero. Then, for specificity, along $\mathbf{W}^\beta(l)$, the following conditions must be satisfied

$$\begin{aligned} &I_\alpha(\mathbf{W}^\beta(l)) = I_\alpha(\mathbf{W}^\beta(0)) \qquad \text{if } \alpha \neq \beta, \\ &I_\beta(\mathbf{W}^\beta(l)) \text{ is an arbitrary monotonically increasing function of } l. \end{aligned} \tag{S.1}$$

The temporal range, that is the maximum and minimum values of $I_\alpha(\mathbf{W}^\beta(l))$, for which (S.1) must hold depends on the needs of the organism and the behavior that the pattern generator is governing. Complemented with such a range, Eq. (S.1) is the most general statement of specificity.

What can be learned from such a statement? Taking a derivative of both sides and applying the chain rule, one gets:

$$\left.\frac{\partial I_\alpha(\mathbf{W}_+)}{\partial \mathbf{W}_+}\right|_{\mathbf{W}_+ = \mathbf{W}_+^\beta(l)} \cdot \frac{d\mathbf{W}_+^\beta(l)}{dl} = \delta_{\alpha\beta} \frac{dI_\beta^\beta(l)}{dl}, \tag{S.2}$$

where $\mathbf{W}_+$ refers to a vector composed from the non-zero elements of $\mathbf{W}$, and $\delta_{\alpha\beta}$ is the Kronecker delta. Eq. (S.2) implies that along a specific trajectory, gradients of non-target interval durations with respect to synpatic weights must be orthogonal to the tangent vector of the trajectory.

Eq. (S.2) defines a non-linear system of K ordinary differential equations (one for each α), with initial conditions $\mathbf{W}^\beta(0) = \mathbf{W}$ for S unknowns, S being the number of non-zero elements of $\mathbf{W}$. The right hand sides of these equations are all zero except when $\alpha = \beta$, for which the equation is:

$$\left.\frac{\partial I_\alpha(\mathbf{W}_+)}{\partial \mathbf{W}_+}\right|_{\mathbf{W}_+ = \mathbf{W}_+^\beta(l)} \cdot \frac{d\mathbf{W}_+^\beta(l)}{dl} = \frac{dI_\beta^\beta(l)}{dl}. \tag{S.3}$$

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Let us remember that $I_\beta(\mathbf{W}^\beta(l))$ is an arbitrary monotonically increasing function of l. We can deal with this degree of freedom by reparametrizing l. We define

$$\bar{l}(l) := I_\beta(\mathbf{W}^\beta(l)). \tag{S.4}$$

In terms of this new parameter, (S.2) becomes,

$$\left.\frac{\partial I_\alpha(\mathbf{W}_+)}{\partial \mathbf{W}_+}\right|_{\mathbf{W}_+=\mathbf{W}_+^\beta(\bar{l})} \cdot \frac{d\mathbf{W}_+^\beta(\bar{l})}{d\bar{l}} = \delta_{\alpha\beta}. \tag{S.5}$$

Hence, we are left with K equation for S unknown functions with no extra degree of freedom. In ordinary situations, $K << S$ and hence these equations are underdetermined and solutions to them can be found, provided that the dependence of interval durations on synaptic weights are "well-behaved" (in a sense that will be made precise in the next paragraph) for a range of interval durations that are required by the behavioral task at hand. This argument can be separately repeated for all the intervals that the pattern generator produces. Then, requiring the existence of specific trajectories starting from the original network state does not restrict the network architecture.

In our discussion, we implicitly assumed that the pattern generator circuit is well-behaved, meaning that the following two cases are excluded:

1. $\frac{\partial I_\alpha(\mathbf{W}_+)}{\partial \mathbf{W}_+} = 0$. We exclude this case because the pattern generator doesn't allow local changes to α^{th} interval and hence is not flexible.

2. $I_\alpha(\mathbf{W}_+)$ is not differentiable. This situation could occur, for example, if the network is poised at a bifurcation point. We exclude this case, because such networks will not exhibit the robustness required of a pattern generator circuit.

Until now, we have only considered specific changes to intervals starting from the original network connectivity. What if a specific change to another interval is required after a specific change to the network has already been made? Such a scenario can be generalized to task-specific combinations of intervals and interval changes, each of which would add constraints of the form (S.2), starting from different network connectivities. In this paper, we only focus on changes from the original network connectivity.

1.2 Independent trajectories in synaptic weight space

Next, we discuss independent trajectories. These are trajectories along which multiple intervals can change simultaneously without interference, by which we mean that their rates of change do not suffer compared to how they would change on their own. Suppose $\beta_1, \ldots, \beta_r$ are the intervals targeted for change. Is there an independent trajectory, $\mathbf{W}^{\beta_1,\ldots,\beta_r}(l)$, in synaptic weight space for changing the duration of these intervals?

To quantify interference, we need to compare the rate of interval changes along $\mathbf{W}^{\beta_1,\ldots,\beta_r}(l)$ to the rate interval changes along specific trajectories. Therefore, to have a well-defined problem we need to assume the existence of specific trajectories (see above). Then, we can state the condition for independence as:

$$I_\alpha(\mathbf{W}^{\beta_1,\ldots,\beta_r}(l)) = I_\alpha(\mathbf{W}^{\beta_1}(l)) + \ldots + I_\alpha(\mathbf{W}^{\beta_r}(l)), \tag{S.6}$$

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

where $I_\alpha(\mathbf{W}^{\beta_i}(l))$ are specific trajectories. The parametrization of specific trajectories can be chosen freely for the current discussion, but in a biological setting that will be discussed in the next section, l relates to the learning rate during single- and multiple-target learning. Taking derivatives with respect to l and applying the chain rule on both sides gives:

$$\left.\frac{\partial I_\alpha(\mathbf{W})}{\partial \mathbf{W}_+}\right|_{\mathbf{W}=\mathbf{W}^{\beta_1,\ldots,\beta_r}(l)} \cdot \frac{d\mathbf{W}_+^{\beta_1,\ldots,\beta_r}(l)}{dl} = \sum_i \left.\frac{\partial I_\alpha(\mathbf{W})}{\partial \mathbf{W}_+}\right|_{\mathbf{W}=\mathbf{W}^{\beta_i}(l)} \cdot \frac{d\mathbf{W}_+^{\beta_i}(l)}{dl}. \tag{S.7}$$

Note that the terms in the summation of the right hand side are zero except when $\alpha = \beta_i$, due to specificity. Hence, independent paths must be orthogonal to gradients of non-target intervals. For each $\beta_1, \ldots, \beta_r$ combination, this amounts to K equations, one for each interval α, for S unknowns. In ordinary situations, $K << S$ and hence these equations are underdetermined and solutions to them can be found, provided that the dependence of interval durations on synaptic weights are "well-behaved" for the required range of interval durations. This argument can be separately repeated for all possible independent interval combinations. Then, the requirement for independence of trajectories starting from the original network state would not restrict network architecture.

Once a set of specific paths are found, locally (meaning around $l = 0$), a linear combination of specific paths gives an independent path. To be more precise,

$$\mathbf{W}^{\beta_1,\ldots,\beta_r}(\epsilon) = \mathbf{W} + \epsilon \sum_i \left.\frac{d\mathbf{W}^{\beta_i}(l)}{dl}\right|_{l=0} + \mathcal{O}(\epsilon^2), \tag{S.8}$$

where ϵ is a small parameter, solves (S.7) to order ϵ^2. There could, however, exist other independent paths.

The existence of specific and independent trajectories in synaptic weight space is a minimum and necessary requirement for flexible time-keeping, but such paths must be found using biologically plausible learning rules, and we discuss them next.

1.3 Biologically plausible reinforcement learning of specific and independent temporal changes

Until now we discussed the existence of trajectories in weight space along which modification to individual intervals is specific and independent. In reality, such trajectories have to be traced by synaptic plasticity during a trial-and-error learning. The information available at each synapse is very restricted: a signal about the produced target interval as well as knowledge about the synaptic weight of the synapse and the states of the pre- and postsynaptic neurons. This poses a serious problem for biologically plausible specific and independent modifications to timing. What neurons as a population can do, instead of solving (S.2) and (S.7), is to learn to increase reward. In fact, there are many biologically plausible reinforcement learning models in the literature, e.g. [1, 2, 3], which suggest plasticity rules that find the direction of maximum reward increase in synaptic weight space when averaged across many trials, i.e. the 'gradient'.

Leaving the trial-and-error aspect of learning aside (it is addressed in the main paper), let us for a moment assume that the network has calculated a trajectory that is a gradient

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

ascent on reward, $R^\beta = R^\beta(I_\beta)$ (β is the target interval):

$$\frac{d\mathbf{W}_+^\beta}{dl} = \eta \frac{\partial R^\beta}{\partial \mathbf{W}_+} = \eta \frac{dR^\beta}{dI_\beta} \frac{\partial I_\beta}{\partial \mathbf{W}_+}, \tag{S.9}$$

where η is a positive parameter that governs the rate of change. Then, using (S.2) and (S.9), we deduce that biologically plausible specific modification of interval duration requires orthogonality:

$$\frac{dR^\beta}{dP_\beta} \frac{\partial I_\beta}{\partial \mathbf{W}_+} \cdot \frac{\partial I_\alpha}{\partial \mathbf{W}_+} = \frac{dR^\beta}{dP_\beta} M_{\beta\alpha} = D_{\beta\alpha}(l), \tag{S.10}$$

where $D_{\beta\alpha}(l)$ is diagonal with positive diagonals, $\mathbf{M}$ is the learning matrix defined in the main text, and $\mathbf{W}^\beta(l)$ is now the trajectory obtained by the reinforcement learning algorithm. Note that this relation has to hold at each point along the trajectory. Equation (S.10) is a design constraint on a pattern generator network and in the main text we check whether the right hand side is indeed a diagonal function for various possible pattern generator networks (Figs. 2D, 3D and 4C).

Simultaneous and independent modifications to multiple intervals require multiple reinforcement signals delivered to the network. Here we assume that the combined effect of such reinforcement is given by an addition of each reinforcement. Borrowing notation from the previous section:

$$R^{\beta_1,\ldots,\beta_r} = R^{\beta_1} + \ldots + R^{\beta_r}. \tag{S.11}$$

With this assumption, specificity in a reinforcement learning experiment would imply independence, as, locally, independent paths will be linear combinations of specific paths, as in (S.8).

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

2 Supplementary Note II: Flexible time-keeping in feedforward networks require a one-to-one mapping between synapses and the interval durations they affect

Here we discuss flexible time-keeping in a pattern generator network with feedforward architecture, making the following general assumptions:

1. The mapping between network activity and time is layer specific: the k^{th} time point is a function of only the activity of neurons in the k^{th} layer.

2. Time increases with progression through the feedforward network. Activity in each layer codes for the start of an interval and end of the previous one.

3. Only the initial layer, which we call 0^{th} layer, receives external input.

For example, mean spike times of neurons in the k^{th} layer could code for the beginning of the k^{th} interval and the end of the $(k-1)^{\text{th}}$ interval. These very generic assumptions are sufficient to prove that in a feedforward network, specificity and independence of gradient ascent reinforcement learning is possible only if the duration of the k^{th} interval depends only on the synaptic weights between k^{th} and $(k-1)^{\text{th}}$ layers.

2.1 The case of a single neuron per layer

For simplicity, let's first assume a single neuron per layer. We will discuss the generalization of our results to multiple neurons per layer in the next section. We denote the activity of the i^{th} neuron by $r_i(t)$, which can be spikes or firing rate. The feedforward architecture makes $r_i(t)$ a function of the activity of the presynaptic neuron $r_{i-1}(t)$ and the connection between them w_i:

$$r_i(t) = r_i\left(w_i, r_{i-1}(t)\right). \tag{S.12}$$

Interval boundaries, on the other hand, depend on the activity of neurons

$$t_i = t_i\left(r_i(t)\right) = t_i\left(w_i, r_{i-1}(t)\right), \tag{S.13}$$

where t_i denotes the end of i^{th} interval.

Claim 1: In a feedforward network with a single neuron per layer, under the assumptions listed above, flexible time-keeping requires that the k^{th} interval, $I_k = t_k - t_{k-1}$ depends only on w_k, i.e.

$$\begin{aligned} \frac{\partial I_k}{\partial w_l} &= 0 \qquad \text{if} \quad \mathrm{k} \neq \mathrm{l}, \\ \frac{\partial I_k}{\partial w_l} &\neq 0 \qquad \text{if} \quad \mathrm{k} = \mathrm{l}. \end{aligned} \tag{S.14}$$

Hence, synaptic weights affect only specific interval durations.

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Proof. We will prove this claim using strong induction, but, first note that the statement always holds for $l > k$ as activity of postsynaptic neurons do not affect activity of presynaptic neurons in a chain,

$$\frac{\partial t_k}{\partial w_l} = 0, \qquad w > l. \tag{S.15}$$

We prove the claim for the first interval as the base case. Note that,

$$\begin{aligned} \frac{\partial I_1}{\partial w_1} &= \frac{\partial t_1}{\partial w_1} - \frac{\partial t_0}{\partial w_1}, \\ \frac{\partial I_1}{\partial w_i} &= 0, \qquad i > 1. \end{aligned} \tag{S.16}$$

Flexible learning requires the learning matrix to be diagonal with positive diagonal elements (see (S.10)). The first row and column of the matrix are given by:

$$M_{1,i} = M_{i,1} = \sum_j \frac{\partial I_1}{\partial w_j}\frac{\partial I_i}{\partial w_j} = \frac{\partial I_1}{\partial w_1}\frac{\partial I_i}{\partial w_1} \propto \delta_{i1}. \tag{S.17}$$

Hence, flexible learning requires

$$\frac{\partial I_i}{\partial w_1} \propto \delta_{i1}. \tag{S.18}$$

Now the (strong) induction step: Assume that the claim holds for $k = 1$ to $k = p - 1$. We prove that this implies that the claim holds for $k = p$. Note that,

$$\frac{\partial I_p}{\partial w_l} = \begin{cases} \frac{\partial t_p}{\partial w_l} - \frac{\partial t_{p-1}}{\partial w_l} & l < p \\ \frac{\partial t_p}{\partial w_p} & l = p \\ 0 & l > p \end{cases} \tag{S.19}$$

Flexible learning requires learning matrix to be diagonal with positive diagonal elements. The p^{th} row and column of the matrix are given by:

$$M_{p,i} = M_{i,p} = \sum_j \frac{\partial I_p}{\partial w_j}\frac{\partial I_i}{\partial w_j} \propto \delta_{ip}. \tag{S.20}$$

Looking at the right hand side, we note that the terms in the summation for which $j > \min(i, p)$ are 0 by (S.15) and for which $j < i$ are 0 by our induction assumption. Then, the only possibly non-zero term in the summation is:

$$M_{p,i} = M_{i,p} = \frac{\partial I_p}{\partial w_i}\frac{\partial I_i}{\partial w_i} \propto \delta_{ip}. \tag{S.21}$$

Since $\frac{\partial I_i}{\partial w_i} \neq 0$, flexible learning requires

$$\frac{\partial I_p}{\partial w_i} \propto \delta_{ip}. \tag{S.22}$$

□

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Next, we present a sufficiency condition for a feedforward network to exhibit flexible timing.

Claim 2: Suppose all neurons' activities are time-invariant with respect to their pre-synaptic partners, i.e. if the pre-synaptic neuron activity is shifted in time by t', the result is a temporal shift in the post-synaptic neuron's activity by t'. Then, a feedforward network exhibits flexible time-keeping if changing a weight w_k leads to a time-shift only of the k^{th} neuron activity.

Proof. If w_k perturbation leads to a time-shift of k^{th} neuron activity, all later neuron activities are time shifted by the same amount by time-invariance property, and therefore no other interval than I_k is affected. □

2.2 The case of multiple neurons per layer

The argument presented above applied to having a single neuron per layer. The argument can be extended to multiple neurons per layer with extra assumptions. We chose a notation that makes the layered structure of the feedforward network explicit. Let's denote by $r_{k,i}(t)$ to be the activity of the i^{th} neuron in the k^{th} layer and $w_{k,ij}$ to be the synaptic weight from the j^{th} neuron in the $(k-1)^{\text{th}}$ layer to the i^{th} neuron in the k^{th} layer. We assume that there are n_k neurons in each layer. The feedforward architecture makes $r_{k,i}(t)$ to be a function of the activity of the presynaptic neurons $r_{k-1,j}(t)$ and the synapses between them $w_{k,ij}$:

$$r_{k,i}(t) = r_{k,i}\left(w_{k,i1}, \ldots, w_{k,in_{k-1}}, r_{k-1,1}(t), \ldots, r_{k-1,n_{k-1}}(t)\right). \tag{S.23}$$

Interval boundaries, on the other hand, depend on the activity of neurons

$$\begin{aligned} t_k &= t_k\left(r_{k,i}(t), \ldots, r_{k,n_k}(t)\right) \\ &= t_i\left(w_{k,i1}, \ldots, w_{k,in_{k-1}}, r_{k-1,1}(t), \ldots, r_{k-1,n_{k-1}}(t)\right). \end{aligned} \tag{S.24}$$

Claim 3: In a feedforward network with a multiple neurons per layer, under the assumptions listed above, it is sufficient for flexible time-keeping that k^{th} interval, $I_k = t_k - t_{k-1}$ depends only on $w_{k,ij}$, i.e.

$$\begin{aligned} &\frac{\partial I_k}{\partial w_{l,ij}} = 0 \qquad \text{if} \quad \mathrm{k} \neq \mathrm{l}, \\ &\frac{\partial I_k}{\partial w_{l,ij}} \neq 0 \qquad \text{if} \qquad \mathrm{k} = \mathrm{l} \qquad \text{for at least one } \{i,j\} \text{ pair.} \end{aligned} \tag{S.25}$$

If we further have the extra condition that intervals cannot increase with weight increases, i.e. $\frac{\partial I_k}{\partial w_{l,ij}} \leq 0$, than (S.25) is both necessary and sufficient for flexible time-keeping.

Proof. We will prove this claim using strong induction. We note that the claim already holds for $l > k$ as activity of postsynaptic neurons do not affect activity of presynaptic neurons in a chain,

$$\frac{\partial t_k}{\partial w_{l,ij}} = 0, \qquad w > l. \tag{S.26}$$

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

We prove the claim for the first interval as the base case. Note that,

$$\frac{\partial I_1}{\partial w_{1,ij}} = \frac{\partial t_1}{\partial w_{1,ij}} - \frac{\partial t_0}{\partial w_{1,ij}} = \frac{\partial t_1}{\partial w_{1,ij}},$$
$$\frac{\partial I_1}{\partial w_{l,ij}} = 0, \qquad l > 1. \tag{S.27}$$

Flexible learning requires learning matrix to be diagonal with positive diagonal elements (see (S.10)). The first row and column of the matrix are given by:

$$M_{1,l} = M_{l,1} = \sum_m \sum_i \sum_j \frac{\partial I_1}{\partial w_{m,ij}} \frac{\partial I_l}{\partial w_{m,ij}} = \sum_i \sum_j \frac{\partial I_1}{\partial w_{1,ij}} \frac{\partial I_l}{\partial w_{1,ij}}. \tag{S.28}$$

If (S.25) holds, then

$$M_{1,l} = M_{l,1} \propto \delta_{l1}, \tag{S.29}$$

ensuring flexible learning. Further, if intervals cannot increase with weight increases, then (S.25) for $k = 1$ is necessary for flexible learning.

Now the (strong) induction step: Assume that the claim 3 holds for $k = 1$ to $k = p - 1$. We prove that this implies that the claim holds for $k = p$. Note that,

$$\frac{\partial I_p}{\partial w_{l,ij}} = \begin{cases} \frac{\partial t_p}{\partial w_{l,ij}} - \frac{\partial t_{p-1}}{\partial w_{l,ij}} & l < p \\ \frac{\partial t_p}{\partial w_{p,ij}} & l = p \\ 0 & l > p \end{cases} \tag{S.30}$$

Flexible learning requires learning matrix to be diagonal with positive diagonal elements. The p^{th} row and column of the matrix are given by:

$$M_{p,l} = M_{l,p} = \sum_m \sum_i \sum_j \frac{\partial I_p}{\partial w_{m,ij}} \frac{\partial I_l}{\partial w_{m,ij}} \tag{S.31}$$

Looking at the right hand side, we note that the terms in the summation for which $m > \min(l, p)$ are 0 by (S.26) and for which $m < l$ are 0 by our induction assumption. Then, the only possibly non-zero terms in the summation are:

$$M_{p,l} = M_{l,p} = \sum_i \sum_j \frac{\partial I_p}{\partial w_{l,ij}} \frac{\partial I_l}{\partial w_{l,ij}}. \tag{S.32}$$

If (S.25) holds, then

$$M_{l,p} = M_{p,l} \propto \delta_{pl}, \tag{S.33}$$

ensuring flexible learning. Further, if intervals cannot increase with weight increases, then (S.25) for $k = p$ is necessary for flexible learning.

□

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

3 Supplementary Note III: Flexibility range of a chain of integrate-and-fire neurons

Here we present a calculation of the range of synaptic weight strengths within which a chain of integrate-and-fire neurons exhibit flexible timing. Integrate-and-fire neurons are arranged in a chain with a single neuron per layer, and the first spikes of neurons mark interval boundaries. The sub-threshold dynamics of neuron k is given by:

$$\tau \frac{dV_k}{dt} = -V_k + w_k \sum_j E(t - t_{k-1}^j). \tag{S.34}$$

$E(t)$ is the excitatory post-synaptic current (EPSC) which is 0 for $t < 0$ and normalized to unit area, $\int_0^\infty dt\, E(t) = 1$, t_{k-1}^j is the j^{th} spike of the $(k-1)^{\text{th}}$ neuron. When the neuron reaches threshold, which we choose to be 1, the neuron produces a spike and the membrane potential is reset to 0. We assume that the neuron is at rest potential, 0, when the first pre-synaptic spike arrives. The membrane potential of the neuron after the first pre-synaptic spike and before its own spike is given by:

$$V_k(t + t_{k-1}) = w_k \int_0^t \frac{dt'}{\tau} e^{-(t-t')/\tau} E(t'). \tag{S.35}$$

We will work in the synaptic weight regime where the activity propagates in the chain with each neuron producing a single spike, therefore we will omit the j index in the rest of this section.

The minimum synaptic weight strength for producing a spike happens when the maximum value of the membrane potential hits the spiking threshold. Taking a derivative of the membrane potential and setting it to zero, we get an implicit equation for the time of the maximum potential, $t_{\max}$:

$$E(t_{\max}) = \int_0^{t_{\max}} \frac{dt'}{\tau} e^{-(t_{\max}-t')/\tau} E(t'). \tag{S.36}$$

Then, the minimum synaptic weight at which a spike is produced is given by:

$$w_{k,\min} = \frac{1}{\int_0^{t_{\max}} \frac{dt'}{\tau} e^{-(t_{\max}-t')/\tau} E(t')}. \tag{S.37}$$

As an example, assume that $E(t) = \Theta(t)\, e^{-t/\tau_s}/\tau_s$, where $\Theta(t)$ is the step function and τ_s is the synaptic time constant. When $\tau_s \neq \tau$, $t_{\max} = \frac{\tau\tau_s \ln(\tau/\tau_s)}{\tau - \tau_s}$ and $w_{k,\min} = \tau_s \left(\frac{\tau_s}{\tau}\right)^{\frac{\tau}{\tau_s - \tau}}$. When $\tau_s = \tau$, $t_{\max} = \tau$ and $w_{k,\min} = \tau e$. $w_{k,\min}$ monotonically increases with τ_s, starting with $w_{k,\min} = 0$ for $\tau_s = 0$, which is the delta-function EPSC limit. The intuition behind this behavior is that as τ_s get bigger, less of the EPSC falls in the time window of integration ($\sim \tau$) and the rise in membrane potential becomes smaller. To compensate for this effect, the synaptic weight has to get stronger.

The maximum synaptic weight strength for flexible timing depends on the configuration of intervals in the network. One can increase the synaptic weight until a second postsynaptic

bioRxiv preprint doi: https://doi.org/10.1101/033472; this version posted December 2, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

spike is generated without any effect on the durations of downstream intervals. The second spike, however, increases the excitation in the $(k+1)^{\text{th}}$ neuron, and may lead to earlier spiking of the $(k+1)^{\text{th}}$. If the $(k+1)^{\text{th}}$ neuron spiked before the second spike of the k^{th} neuron, there won't be a change in the duration of the $(k+1)^{\text{th}}$ interval, however, downstream intervals may become affected by the second volley of spikes that propagate down the chain. Therefore, a lower bound for the maximum synaptic weight strength for flexibility, is the strength at which a second spike is generated by the k^{th} neuron. The membrane potential after the first spike is

$$V_k(t+t_k) = w_k \int_{t_k(w_k)}^{t} \frac{dt'}{\tau} e^{-(t-t')/\tau} E(t'). \tag{S.38}$$

Note that t_k is implicitly a function of w_k as synaptic weight strength changes the timing of the first spike. The maximum of membrane potential after the first spike is given by the solution to

$$E(\bar{t}_{\max}) = \int_{t_k(w_k)}^{\bar{t}_{\max}} \frac{dt'}{\tau} e^{-(\bar{t}_{\max}-t')/\tau} E(t'). \tag{S.39}$$

The lower bound on maximum weight is when the membrane potential reaches spiking threshold at $t = \bar{t}_{\max}$:

$$w_{k,\max} = \frac{1}{\int_{t_k(w_k)}^{\bar{t}_{\max}} \frac{dt'}{\tau} e^{-(\bar{t}_{\max}-t')/\tau} E(t')}. \tag{S.40}$$

A refractory period, or a bursting mechanism such as generated by calcium spikes that extends through the timing of a possible second spike/burst will increase the lower bound on maximum synaptic weight.

References

[1] Ila R Fiete and H Sebastian Seung. Gradient learning in spiking neural networks by dynamic perturbation of conductances. *Physical review letters*, 97(4):048104, 2006.

[2] Ronald J Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning*, 8(3-4):229–256, 1992.

[3] Kenji Doya and Terrence J Sejnowski. A novel reinforcement model of birdsong vocalization learning. *Advances in neural information processing systems*, pages 101–108, 1995.