Dual pathway architecture underlying vocal learning in songbirds

Remya Sankar

THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEUR
DE L'UNIVERSITÉ DE BORDEAUX

ECOLE DOCTORALE: Mathématiques et Informatique

SPÉCIALITÉ: Informatique

Par Remya SANKAR

Dual pathway architecture underlying vocal learning in songbirds

Sous la direction de : Nicolas P. ROUGIER
Co-directeur : Arthur LEBLOIS

Soutenue le 12 décembre 2022

Membres du jury :

Mme Adrienne FAIRHALL	Professeur	University of Washington	Rapporteur
M. Mehdi KHAMASSI	Directeur de Recherche	ISIR	Rapporteur
Mme Catherine LE MOINE	Directrice de Recherche	INCIA	Examinatrice
M. Richard HAHNLOSER	Professeur	ETH Zürich	Examinateur
M. Nicolas P. ROUGIER	Directeur de Recherche	INRIA	Directeur
M. Arthur LEBLOIS	Directeur de Recherche	CNRS	Directeur

Résumé vulgarisé

L'apprentissage sensorimoteur est l'acquisition d'aptitudes motrices par un individu, guidée par la perception sensorielle, comme l'acquisition de la parole chez l'homme et les oiseaux chanteurs. Nous utilisons les oiseaux chanteurs pour étudier les circuits neuronaux qui soutiennent l'apprentissage sensorimoteur. Nous créons des modèles informatiques pour explorer l'utilité des voies corticales et sous-corticales parallèles qui régissent l'acquisition du chant, et nous étudions l'interaction entre la plasticité fonctionnelle et structurelle dans ces circuits. Inspirés par le comportement d'apprentissage vocal, nous étudions les mécanismes biologiquement plausibles pour optimiser l'apprentissage sensorimoteur. Nous étudions les corrélats neuraux de l'apprentissage par renforcement contribuant à l'apprentissage vocal en observant l'activité neuronale dans le système de chant. En utilisant les approches théoriques et expérimentales, nous espérons mieux comprendre l'apprentissage sensorimoteur.

Abstract

Sensorimotor learning refers to the acquisition of motor skills by an individual, guided by sensory perception, for instance, learning to speak. Human speech acquisition is rather similar to vocal learning by songbirds. Hence, we use songbirds to study the neural circuitry underlying sensorimotor learning. We build computational models to explore the utility of parallel cortical and subcortical pathways that govern song acquisition, and investigate the interplay between functional and structural plasticity within this circuitry. Inspired from the vocal learning behaviour, we look at biologically plausible mechanisms to optimise sensorimotor learning. We investigate the neural correlates of reinforcement learning contributing to vocal learning by observing the neural activity in the song system of male zebra finches. By using both theoretical and experimental approaches of neuroscience, we hope to advance our understanding of vocal learning, and, in general, sensorimotor learning.

Architecture à double voie sous-jacent l'apprentissage vocal des oiseaux chanteurs

Résumé:
Les êtres humains peuvent maîtriser nombre d'aptitudes motrices. L'apprentissage sensorimoteur est l'acquisition de multiple aptitudes motrices par un individu, guidée par la perception sensorielle. Ici, nous examinons les circuits et mécanismes neuronaux qui peuvent soutenir l'apprentissage sensorimoteur. Les oiseaux chanteurs et leur comportement d'apprentissage vocal constituent une voie intéressante pour étudier les circuits neuronaux qui sous-tendent l'apprentissage sensorimoteur.

La plasticité structurelle est centrale dans le développement des circuits cérébraux liés au chant et nous étudions son rôle surement critique dans l'acquisition du chant chez les jeunes oiseaux. Nous expliquons la contribution de la plasticité structurelle dans un circuit pour aider l'apprentissage sensorimoteur. Dans le chapitre 2, nous examinons les différents types de plasticité et leurs interactions en contribuant à la fonction du circuit. Nous explorons les études de modélisation actuelles intégrant la plasticité structurelle.

Dans le chapitre 3, nous présentons le comportement d'apprentissage vocal chez les oiseaux chanteurs comme un bon exemple d'apprentissage sensorimoteur, ainsi que les circuits qui le sous-tendent. Nous établissons la pertinence de l'étude des oiseaux chanteurs en démontrant leurs similitudes frappantes avec le comportement humain ainsi que les parallèles avec les circuits neuronaux. Il a été démontré que le transfert des structures sous-corticales de type BG vers les voies corticales joue un rôle dans leur apprentissage vocal. Nous utilisons cet exemple spécifique d'apprentissage sensorimoteur pour étudier les contributions possibles de la plasticité structurelle et fonctionnelle à l'apprentissage sensorimoteur, dans un paradigme d'apprentissage en deux étapes.

Tout d'abord, nous examinons la possibilité d'un rôle fonctionnel joué par la plasticité structurelle dans l'apprentissage sensorimoteur, au chapitre 5. Il existe des preuves montrant le retard de maturation de la voie corticale, par rapport à la voie BG-thalamo-corticale. Nous enquêtions si ce retard dans le développement de la voie motrice corticale contribue directement à l'apprentissage vocal. Nous examinons cette question en créant une architecture simplifiée du système vocal et en simulant une tâche non linéaire analogue d'apprentissage sensorimoteur par un bras pivotant multi-segmenté.

Ensuite, nous examinons les stratégies biologiquement réalistes adoptées par l'oiseau chanteur pour résoudre le problème de la navigation dans un paysage sensorimoteur complexe et inégal. Il a été démontré qu'une détérioration post-sommeil de la qualité du chant chez les oiseaux juvéniles peut être associée à une éventuelle imitation supérieure, du chant du tuteur, à l'âge adulte. Nous nous inspirons de ces preuves empiriques concernant le comportement de l'apprentissage vocal et son anatomie sous-jacente. Nous établissons ensuite des parallèles avec les algorithmes utilisés en machine learning pour compléter l'apprentissage par renforcement basé sur la descente de gradient. Nous étudions d'abord ces questions à l'aide d'un modèle conceptuel, puis nous le renforçons en utilisant un réseau de populations de neurones à fréquence codée.

Finalement, nous examinons le transfert d'un tel apprentissage basé sur le BG dans le système de chant chez le diamant mandarin. Nous étudions les corrélats neuraux de l'apprentissage par renforcement contribuant à l'apprentissage vocal, dans le cadre de la double voie. Nous testons

cette hypothèse en utilisant un modèle computationnel simplifié du système de chant en simulant la plasticité adulte. Nous utilisons ensuite l'électrophysiologie pour enregistrer les signaux neuronaux du système de chant chez des diamant mandarin, pendant l'apprentissage vocal, ce qui nous permet d'étudier l'interaction entre les noyaux du système de chant. En utilisant les approches théoriques et expérimentales, nous espérons mieux comprendre l'apprentissage sensorimoteur.

Mots-clés : L'apprentissage sensorimoteur, neuroscience, apprentissage par renforcement

Dual pathway architecture underlying vocal learning in songbirds

Abstract:
From learning to walk to riding a bike to playing tennis, human beings display a remarkable ability to master a wide range of motor skills. Sensorimotor learning refers to the acquisition of all such motor skills by an individual, guided by sensory perception. In this manuscript, we look into the neural circuitry and mechanisms that can support trial-by-trial sensorimotor learning. The animal model of songbirds along with their vocal learning behaviour provide an interesting avenue to study the neural circuitry underlying sensorimotor learning.

Structural plasticity is central to the development of song-related brain circuits and we investigate its possibly critical role in song acquisition in juvenile songbirds. We account for the contribution of structural plasticity within a circuit towards facilitating sensorimotor learning. In chapter 2, we look into the different types of plasticity and their interactions with each other while contributing towards the function of the circuit. We explore current modelling studies incorporating structural plasticity and unexplored questions which are raised by such models.

In chapter 3, we introduce the vocal learning behaviour in songbirds as a tractable instance of sensorimotor learning, along with the circuitry underlying it. We establish the relevance of studying the animal model of songbirds by demonstrating their striking similarities with respect to human behaviour as well as parallels with neural circuitry. The transfer from subcortical BG-like structures to cortical pathways has been shown to play a role in vocal learning in songbirds. We use this specific instance of sensorimotor learning to study the possible contributions of structural as well as functional plasticity to sensorimotor learning, within a two-step learning paradigm.

First, we look into the possibility of a functional role played by structural plasticity towards sensorimotor learning, in chapter 5. There is evidence showing the delayed maturation of the cortical pathway, as compared to the BG-thalamo-cortical pathway. We investigate if this delay in the development of the cortical motor pathway contributes directly towards vocal learning. We look into this by building a simplified architecture of the song system, and simulating an analogous non-linear task of sensorimotor learning by a multi-segmented pivoted arm.

Second, we take a closer look into biologically realistic strategies adopted by the songbird to solve the problem of navigating an extremely complex and uneven sensorimotor landscape. It has been shown that a post-sleep deterioration in song quality in juvenile birds can be associated with an eventual superior imitation of tutor song during adulthood. We take inspiration from such empirical evidences regarding both the behaviour of vocal learning and the underlying anatomy supporting it. We, then, draw parallels with algorithms used in machine learning to supplement gradient-descent based reinforcement learning. We, first, investigate these questions using a conceptual model and then proceed to ratify it using a network of rate-coded neuron populations.

Third, moving from the circuit level to a more granular scale, we look closely into the transfer of such BG-driven learning within the song system in zebra finches. We posit a hypothesis to investigate the neural substrates of reinforcement learning contributing to vocal learning, within the dual pathway framework. We test this hypothesis using a simplified computational model of the song system. We, then, use electrophysiology to record neural signals from the song

system in freely-behaving zebra finches, when subjected to a protocol to induce vocal learning in adult songbirds, allowing us to study the interaction between nuclei in the song system during sensorimotor learning.

By using both theoretical and experimental approaches of neuroscience, we hope to advance our understanding of vocal learning, and, in general, sensorimotor learning.

Keywords: Sensorimotor learning, neuroscience, reinforcement learning

---

Unité de recherche
Inria Bordeaux Sud-Ouest, 200 Avenue de la Vielle Tour, Talence, 33000
Institut des Maladies Neurodégénératives – CNRS UMR 5293, 146 rue Léo Saignat 33076 Bordeaux cedex
LABRI, 351, cours de la Libération F-33405 Talence cedex

To *Thatha*.

Contents

1 Sensorimotor learning

From learning to walk to riding a bike to playing tennis, human beings display a remarkable ability to master a wide range of motor skills. Sensorimotor learning refers to the acquisition of all such motor skills by an individual, guided by sensory perception. Such learning progresses by accounting for the perceived effect of the said action (Wolpert, Diedrichsen, et al. 2011). It involves efficiently processing the sensory information regarding the outcome of a motor action and thus, over time forming an expectation of the outcome of one's actions. Sensorimotor learning involves not only skill acquisition, but also maintenance of skill, for instance, in response to injury or ageing. Moreover, sensorimotor learning pertains to more than just the execution of motor action. It also involves the swift selection of the optimal action in response to a given stimulus or scenario (Krakauer et al. 2019). Certain instances of sensorimotor learning, such as habit formation, are intertwined with procedural learning, where several trials of a particular skill or sequence of actions, leads to an implicit representation of the skill which is not necessarily accessible to awareness (Knowlton et al. 2008). Sensorimotor learning can occur over a wide range of timescales. On one end of the spectrum, certain skills can be learnt in one shot, often by imitation. A child learning to produce a new word or an individual learning to draw a character from a new script are both possible via one shot learning (Carey et al. 1978; Lake et al. 2011) . On the other end of the spectrum, several motor skills require repeated practice over years, as commonly observed within procedural learning (Knowlton et al. 2008).

Similarly, a variety of processes can govern sensorimotor learning. A statistical distribution of perceivable effects of motor action can be built, akin to the prior of a Bayesian inference. This information can then be used to estimate the effect of future motor actions and, thus, help select the appropriate action. Such an estimate can be updated continually, and thus, support maintenance of a skill within an evolving body, or adaptation to change in the environment (Körding et al. 2006). Alternatively, sensorimotor learning can proceed via error-based learning. Progress in motor skill acquisition is usually measured by the reduction in perceived error of motor action or reaction time (Wolpert, Diedrichsen, et al. 2011). In error-based learning, both the desired goal and metric of improvement towards the goal can be quite subjective and purely intrinsic to the individual. Learning, in itself, can be both intrinsically or externally motivated. Curiosity, for instance, is a form of intrinsic motivation, which is well-exhibited by infants and children (Gottlieb et al. 2016). Another example is the completely intrinsic desire of a songbird to sing in isolation, despite it not necessarily being beneficial to the bird (Kim et al. 2021). On the other hand,

external motivation can be presented in the form of goals or rewards in the environment, either in concrete form, such as food, or abstract, such as appreciation. Such factors can influence the objective or error perception of a task by an individual.

Alongside error-based learning, random exploration of motor actions, for e.g. infants at play, can facilitate the formation of an internal inverse model, associating a particular effect in the sensory phase with a particular motor action initiated by the agent. Such an inverse model can be harnessed in imitative learning, where a motor skill can be learnt quickly by recalling motor actions corresponding to specific target effects (Kuperstein 1988). Moreover, an agent can build a forward model to predict the sensory consequences of a particular motor action (Ito 2000; Jordan et al. 1992; Sutton, Barto, et al. 1998). For instance, the amount of force one applies to lift an unknown object is scaled as per their initial visual perception, and subsequent tactile perception. Such a forward model (possibly also combined with the inverse model) can be used in error-based learning, where the improvement of a motor skill proceeds through an extensive trial-and-error process (Wolpert and Flanagan 2001). However, random sampling in the motor space can be rather inefficient and does not necessarily induce uniform sampling in the sensory space (Benureau 2015). The outcome of such error-based learning can be refined further using reinforcement learning (Wolpert, Diedrichsen, et al. 2011).

It has been widely hypothesised that the BG implements reinforcement learning in order to convey tutor signals towards vocal learning (Doya and Sejnowski 1998; M. Fee et al. 2011; Fiete, M. S. Fee, et al. 2007; Wickens et al. 2003) (discussed further in chapter 5). Reinforcement learning involves an agent repeatedly interacting with the environment and adapting their interactions, depending on their perception of the environmental response (Sutton, Barto, et al. 1998). Under this paradigm, the agent requires the ability to, one, modify its actions, two, perceive, evaluate and compare the effect of their actions against an internal objective, and, three, choose actions which provide more favourable results. There are several theories regarding the internal mechanisms that can provide these required elements and support reinforcement learning within neural circuitry, especially within the avian song system (Doya and Sejnowski 1998; Farries and Fairhall 2007) (discussed further in chapter 3). One, variability induced by different neural regions within the central nervous system, as well as, motor effectors can enable an agent to modify its actions and explore the varying outcomes of different actions. Two, dopaminergic projections from the mid-brain have been implicated in several studies to transmit information conveying the error in outcome prediction (Schultz et al. 1997). Such feedback provides the system with the ability to compare different action outcomes and evaluate them with respect to the desired goal. Three, the system can, then,"reinforce" preferable motor actions using experience-dependent synaptic plasticity within the relevant neural circuitry (Ding et al. 2004a; Mehaffey et al. 2015). Thus, reinforcement learning seems to be a biologically plausible mechanism to implement sensorimotor learning within neural circuits.

Several cortical and subcortical structures are known to be involved in governing sensorimotor learning across species. Cerebellar-cortical as well as basal ganglia-thalamo-

cortical loops have been shown to play a role in sensorimotor learning (Caligiore, Arbib, et al. 2019). The cerebellum is essential for proper sensory and motor timing in the range of milliseconds up to a second (Ivry et al. 1988). The cerebellum timing mechanisms support reinforcement learning in striatal-thalamo-cortical circuits, by providing error correction and sub-second timing signals to the basal ganglia (H. Chen et al. 2013; Schwartze et al. 2016). In sub-second interval detection tasks, as learning proceeds, the need for motor and temporal adjustments decreases along with decreasing reliance on the cerebellum (Caligiore, Arbib, et al. 2019; Ito 2013). Similarly, it has been observed that the basal ganglia can influence slower learning mechanisms in the cortex (Murray et al. 2020). For instance, during reversal training of a visuo-motor association task in monkeys, direction selectivity increased sooner and more abruptly in the caudate nucleus (a part of the striatum that receives direct projections from, and indirectly projects to, the PFC), compared with the pre-frontal cortex (PFC) (Pasupathy et al. 2005b). However, the literature on this matter does not lead to a consensus across species. For instance, Kawai et al. (2015) demonstrate transfer of learning in the opposite direction in rodents, where the motor cortex is required for learning but not for executing a motor skill.

In the context of transfer of learning, a two-step learning paradigm has been widely discussed, wherein the initial learning stages are reliant on subcortical structures and the learning is later progressively imprinted onto neocortical areas (Boraud et al. 2018a). The basal ganglia, the hippocampus, the cerebellum are all involved with cortical loops that can support this kind of transfer. The hippocampus has shown to be involved in the rapid formation of novel memories, while these memories are gradually consolidated within the neocortex (Buzsáki 1989; Milner et al. 1998). This has been evidenced in mice, amongst other species, where increasing the retention interval in a spatial discrimination task resulted in decreased hippocampal metabolic activity, along with a recruitment of certain cortical areas (Bontempi et al. 1999; Eichenbaum 2014). However, while memory formation is a specialised form of learning, it is not specifically sensorimotor learning. As per the model proposed by Doyon et al. (2003) and Nicolson et al. (2007)'s hypothesis, two types of mechanisms can govern procedural learning. They posit that one type of learning proceeds through perceptual-motor adaptation, and is dependent on the cortico-cerebellar loop. Meanwhile, the second type of learning is concerned with memorising perceptual-motor sequences, which depends on the cortico-striatal loop.

In this study, we focus on the striatal-thalamo-cortical network and how the basal ganglia may tutor neocortical motor areas, leading to the long-term storage of motor programs. In addition to primates as discussed above, the transfer from subcortical BG-like structures to cortical pathways has been shown to play a role in vocal learning in songbirds (Andalman et al. 2009a). The song-related BG circuit guides changes in motor output by generating an error-reducing motor bias that is rapidly incorporated into the cortical premotor network. In birds, as in mammals, it is believed that the adaptive motor changes elicited by the BG are learned through a reinforcement learning mechanism (discussed

further in chapter 3).

In this manuscript, we aim to investigate the neural circuitry and mechanisms underlying sensorimotor learning. More precisely, we look into the role of the structure of the circuitry as well as its function towards learning. Neural plasticity is the ability of the brain to undergo persistent functional or structural change, often in response to internal or external stimuli. Plasticity within nervous systems, is the primary substrate, which allows animals to exhibit learning or adapt to a changing environment. In **chapter 2**, we look into the different types of plasticity and their interactions with each other while contributing towards the function of the circuit. For instance, functional plasticity, i.e. experience dependent modification of synaptic strength, is crucial to implement learning within neural circuitry (Citri et al. 2008). On the other hand, the contribution of structural plasticity, i.e., the physical changes in neural connectivity, is seldom discussed in literature. Structural plasticity within neural circuitry is often critical to both the development of the central nervous system and the acquisition of complex behaviours (Butz, Wörgötter, et al. 2009). In this manuscript, we account for the contribution of structural plasticity within a circuit towards facilitating sensorimotor learning. In **chapter 2**, we introduce the various types of plasticity and specifically, review current evidences for structural plasticity across species and development scales. We discuss their interplay with functional synaptic plasticity and their significance from a functional point of view. We explore current modelling studies incorporating structural plasticity and unexplored questions which are raised by such models.

The animal model of songbirds along with their vocal learning behaviour provide an interesting avenue to study the neural circuitry underlying sensorimotor learning. Structural plasticity is central to the development of song-related brain circuits and we investigate its possibly critical role in song acquisition in juvenile songbirds. Moreover, the transfer from subcortical BG-like structures to cortical pathways has been shown to play a role in vocal learning in songbirds (Andalman et al. 2009a). In **chapter 3**, we introduce the vocal learning behaviour in songbirds as a tractable instance of sensorimotor learning, along with the circuitry supporting it. We establish the relevance of studying the animal model of songbirds by demonstrating their striking similarities with respect to human behaviour as well as parallels with neural circuitry. We use this specific instance of sensorimotor learning in order to study the potential contributions of structural and functional plasticity to sensorimotor learning, within the aforementioned two-step learning paradigm. We detail several features as well as constraints imposed by the vocal learning behaviour and underlying circuitry in songbirds, which we later incorporate in our theoretical study into the functioning of the dual pathway architecture.

Inspired by the behaviour as well as the structural and functional plasticity demonstrated by songbirds, we investigate the neural mechanisms of sensorimotor learning. Based on the aforementioned evidences, in **chapter 4**, we build our hypothesis about

the workings of the neural circuits governing vocal learning in songbirds, and detail the objective of our investigation.

In **chapter 5**, we design computational models to look into the utility of the parallel cortical and subcortical pathways that govern song acquisition in songbirds. We study the interplay between functional plasticity and structural plasticity within this vocal learning circuitry. Inspired from the vocal learning circuitry in zebra finches, we build a dual pathway framework underlying sensorimotor learning, and implement a two-step learning paradigm. We incorporate empirical evidences about the vocal learning behaviour and explore potential strategies used by these birds to supplement reinforcement learning and evade its shortcomings in navigating a complex, uneven sensorimotor contour. Ultimately, we draw comparisons with optimisation techniques used in machine learning in similar scenarios, in terms of their advantages, computational cost and biological feasibility.

In **chapter 6**, we further investigate the plausibility of such BG-driven reinforcement learning within the song system in zebra finches, and its subsequent consolidation within cortical circuitry. We posit a hypothesis to verify the mechanisms of BG-led exploration contributing to vocal learning, within the dual pathway framework. In a first approach, to make qualitative predictions about the hypothesis, we design a simplified computational model of the song system wherein we simulate adult plasticity. We, then, proceed to design an experimental protocol to test the predictions that emerge from this model. Within this protocol, we use electrophysiology to record signals from the neural correlates in the song system in freely-moving zebra finches. These songbirds are subjected to a standardised protocol to induce vocal learning in adulthood, allowing us to study the interaction between nuclei in the cortical and subcortical parts of the song system, during sensorimotor learning. We present a pilot study outlining our experimental protocol, preliminary data and methods of analysis, in order to verify the plausibility of reinforcement learning within the song system. Finally, we conclude the manuscript by discussing the relevance of our work, considering its various limitations and formulating possible directions to extend our investigation.[1]

[1]Note: The review article "Computational benefits of structural plasticity, illustrated in songbirds" (Sankar, Rougier, et al. 2022) has been edited and split into chapter 2 and 3. The article "Dual pathway architecture underlying vocal learning in songbirds" (Sankar, Leblois, et al. 2022) about the dual pathway architecture, presented at ICDL 2022, has been incorporated in section 5.2.

2 Structural Plasticity

Plasticity is defined as *the capacity of the neural activity generated by an experience to modify neural circuit function and thereby modify subsequent thoughts, feelings, and behavior* according to (Citri et al. 2008). This definition of plasticity encompasses a broad range of phenomena that have been observed to induce an alteration of the function or the structure of a neural component. For instance, plasticity can be induced by an external stimuli, an internal state, a lesion, etc. Such alterations of neural components occur across a large range of spatial scales and take place through different mechanisms: i) intrinsic plasticity relates to the continual alteration of a neuron's inherent biophysical characteristics by neuronal activity (Debanne 2009) ii) synaptic plasticity relates to the activity-dependent modification of the strength or efficacy of synaptic transmission at preexisting synapses (Citri et al. 2008) iii) representational plasticity relates to the re-organisation of distributed responses as a result of persistent external changes (Buonomano et al. 1998). On the temporal scale, short-term plasticity occurs on the sub-second to minutes time scale in response to an external event such as, for example, an external stimulation and may be transient. Long-term plasticity occurs on the hours, days or years time scale and encompasses long-lasting changes such as the modification of behavior, the formation of new memory, development of new connections across regions, etc.

Across these different spatial and temporal scales, plasticity operates via different means and there are two main forms in which plasticity manifests itself: functional and structural. Functional plasticity involves *changes in synaptic strengths without any change in the anatomical connectivity between neurons* (Butz, Wörgötter, et al. 2009). This can be realized through the insertion or the removal of synaptic receptors, change in the presynaptic release of transmitters or change in the thickness of the synapse and has been studied in a large number of works (Citri et al. 2008; Lledo et al. 2006). The transmission strength of a synapse can, thus, be modified, via potentiation and depression, over milliseconds, to seconds, to hours, to days, and possibly, even longer (Abraham et al. 2003). The experiences of an individual can modify their behaviour through long-term synaptic plasticity, i.e. activity-dependent, long-lasting modifications of synaptic strength (Bliss et al. 1973). Such synaptic plasticity plays a role in consolidation of memories in the hippocampus, in the development of the sensory cortex, in the induction of Pavlovian fear conditioning by the amygdala and in mediating addiction to drugs via the mesolimbic dopaminergic system, amongst several other crucial functions (Everitt et al. 2002; Malenka et al. 2004; Martin et al. 2000; Sigurdsson et al. 2007).

On the other hand, structural plasticity refers to changes in the physical anatomical connections (excluding the simple insertion or removal of synaptic receptors). There are numerous evidences of structural plasticity: from the excessive axonal branching during ontogeny with ensuing pruning to the rapid increase in synaptic density during infancy with subsequent slower synapse elimination to adult neurogenesis in the mammalian hippocampus (Butz, Wörgötter, et al. 2009; Eriksson et al. 1998; Huttenlocher et al. 1979; 1982; Portera-Cailliau et al. 2005). Major structural reorganisation of the primary somatosensory cortex has also been observed following amputation or sensory deprivation (Buonomano et al. 1998) in primates while the trimming of whiskers in juvenile rats have been shown to induce a realignment of the dendrites of stellate neurons in the barrel cortex (Datwani et al. 2002). Such evidences suggests that a change in neural activity may induce structural plasticity. Functional and structural plasticity work hand in hand with other mechanisms in the brain, such as homeostatic plasticity, in order to help stabilise the network activity by coordinating changes in network structure and maintaining the balance between excitation and inhibition within the circuitry (Tien et al. 2018).

The mechanisms underlying plasticity and its relationship with behaviour, learning and memory have been studied for several decades (Bliss et al. 1973; Livingston 1966; Wurtz et al. 1967) with a major focus on functional plasticity more than structural plasticity (Kerr 1975; Ooyen et al. 2017; Rutledge 1978). For example, long term potentiation, long term depression and spike time dependent plasticity have been intensively explored experimentally and widely incorporated in a number of computational works (Bi et al. 1998; Gerstner et al. 1996; Hartley et al. 2006; Markram et al. 1997; Shouval et al. 2002). However, with the advent of new experimental techniques, such as diffusion tensor imaging, the number of studies in structural plasticity has increased (Innocenti et al. 2005) and several of these studies from the last decade have clearly demonstrated the importance of structural rearrangements (Bernardinelli et al. 2014). It has been widely discussed that structural plasticity confers several advantages, such as improving energy efficiency of network formation, increasing information capacity, amongst others (discussed further in section 2.3) (Chechik et al. 1998; Chklovskii et al. 2004; Knoblauch and Sommer 2016). Moreover, it holds the capacity to improve learning in a circuit, while conserving resources (Spiess et al. 2016). This leads us to explore the role played by such structural plasticity in learning and development. While there are several instances across species, from ontogeny to seasonal plasticity (discussed in chapter 2.1 and 3.3), where the role played by such plasticity can be deduced based on the hypotheses prevalent in literature, benefits specific to a particular circuit and function may emerge in certain ecological contexts. Using this particular case of sensorimotor learning in the vocal learning circuitry of birds, where the function fulfilled by structural plasticity is not immediately evident, we investigate if structural plasticity can additionally contribute directly to circuit function.

In this chapter, we review current evidences for structural plasticity and their significance from a computational point of view. To do so, we start by examining evidences for structural plasticity across species and categorizing them along the spatial axis as well as

along the time course of development. Finally, we discuss properties that can be exploited computationally, and unexplored questions which are raised by such models.

2.1 Structural plasticity across spatial and development scales

Structural plasticity pertains to the physical changes in neuronal connections. It comprises of *changes in synapse numbers, axonal fibre densities, axonal and dendritic branching patterns, synaptic connectivity patterns, and even neuronal cell numbers* (Butz, Wörgötter, et al. 2009). Structural plasticity can be categorised according to the spatial scale (e.g. synaptic, axonal, network), the temporal scale (e.g. sub-seconds, minutes, days) and the development period (e.g. sensitive period, childhood, adulthood). Depending on the development period, we can further categorise it based on the cause for structural plasticity: hormonal, learning, pathological, injury-induced.

2.1.1 Spatial scale

Structural plasticity occurs across the whole range of cerebral scales through morphological changes, such as enlargement, growth or apoptosis, that target spines to axon terminals to glia (Bernardinelli et al. 2014). Recent breakthroughs in recording techniques such as *in vivo* confocal microscopy have made it possible to gain deeper knowledge onto the underlying mechanisms.

Synapses go through an exuberant growth during early brain development and most of them will permanently disappear as a result of a competitive process involving neurons, synapses and neural growth factors (Le Bé et al. 2006). It has further been observed that lesions in adult brains can lead to an alteration in the synaptic connectivity patterns, due to structural plasticity (Raisman 1969). Besides sprouting and pruning, structural plasticity can also manifest through neuro-degeneration, neurogenesis (Lledo et al. 2006) and synaptic rewiring, i.e. modification of existing connections, e.g. dendritic spines or axonal branches, by dissociating a pre/post synaptic element, and later, linking it to a different target (De Roo et al. 2008; Trachtenberg et al. 2002).

Dendritic spines are both highly motile and transient structures. For example, in the barrel cortex, about two thirds of the spines remain for less than a month while some of them appear and disappear within a day (Trachtenberg et al. 2002). It has been observed in mice that spine turnover and stabilisation can also correlate with learning and memory consolidation, respectively (Xu et al. 2009). In addition to their motility, dendritic spines occur in a wide range of size, shape and organelles, which in turn potentially affects different functional properties such as, the synaptic strength, its stability, the postsynaptic receptors, etc (Bernardinelli et al. 2014). Morphology does indeed have a direct impact on

the functional role, with large spines helping to form stable connections, and the transience of thin spines being an aid in learning, for instance.

Axon arbor structure can be manipulated within a circuit-level mechanism of learning. In macaque monkeys, trained to perform a contour detection task, Van Kerkoerle et al. (2018) found that there is an extensive sprouting and pruning of axonal collaterals in cortical regions which correspond to the trained area of the visual field. In addition to neurogenesis, neurons have the capacity to degenerate part of their axon in order to refine connections through axon pruning or more drastically, cell apoptosis (Geden et al. 2019; Portera-Cailliau et al. 2005). Both, the mechanisms underlying such formation of short-range or long-range axonal connections as well as the selective degeneration of unoptimal axonal branches could potentially have long-term impacts on circuit function (Tessier-Lavigne et al. 1996).

At the **network** level, significant structural changes occur that affect large cerebral regions. For instance, in blind Braille readers, representation of the reading fingers in the somatosensory cortex is larger than nonreading fingers or any finger of non-Braille readers (Pascual-Leone et al. 1993). Similarly, it has been shown that a massive reorganization of the motor cortex occurs within a few hours following the transection of the facial nerves in the rat (Gilbert 1998). There is, thus, overwhelming evidence that structural plasticity serves a functional role by manipulating various elements of the central nervous system, from spines to synapses to axonal arbors to neurons. We won't detail adult neurogenesis since it is beyond the scope of this manuscript, but we redirect the reader to the studies by Gould (2007), Paredes et al. (2016) and Pytte (2016).

2.1.2 Time course across development

Structural plasticity can be initiated by several factors at several timepoints in one's lifespan. During development, dendritic spines exhibit structural plasticity, by variable sprouting and pruning, depending on experience. Synaptic turnover is extremely high during this period, and a majority of newly sprouted dendritic spines in the mice somatosensory cortex, are lost, within the span of a mere few days (Bufill et al. 2011). As development of an individual progresses, the course of puberty also marks some structural changes. The beginning of higher rates of pruning of dendritic spines in the human frontal cortex, and cortical thinning in humans have been shown to correspond with pubertal development (Boivin et al. 2018).

During development, structural plasticity is limited in time by what is known as 'sensitive periods'. These periods designate limited periods of time wherein several connections or skills are modified permanently and significantly, based on the information acquired through experience (Hensch 2005). Sensitive periods occur in the prenatal brain and continue throughout development but are very limited during adulthood. For instance, newly hatched chicks memorize the characteristics of the first moving object they encounter, and subsequently show a preference for it. This "imprinting" behavior can

only be acquired within the first few days following hatching, and involves age-dependent remodeling of neural networks in the visual and associative areas of the chicks' brains (Nakamori et al. 2013). In humans, language acquisition is, often, a highly cognizant and arduous task for adults, as compared to the relative passive learning by children. It can be argued that, before the end of puberty, certain neural circuits are more susceptible to assimilating such skills (Patton et al. 2019). A prominent study by Hubel et al. (1979) found that closing one eye of kittens, but not of adult cats, causes the permanent loss of visual responsiveness of neurons in the primary visual cortex to stimulations to the deprived eye (ocular dominance plasticity). Moreover, axonal sprouting and branching of thalamic neurons is curbed by visual deprivation during development in rodents, and retinal ganglion cells extend their receptive fields in turtles which have been raised in the dark (Berry et al. 2016; Butz, Wörgötter, et al. 2009). Hence, while the beginning and end of sensitive periods are triggered by molecular signals delivered timely during development, sensory experience is crucial and can modulate the opening and closing of sensitive periods, especially during development (Hensch 2005; Yazaki-Sugiyama 2019).

Beyond development as well, structural plasticity has an impactful presence throughout the lifespan of an animal, triggered by various reasons, ranging across hormonal, lesion-induced, pathological and training. Hormonal changes lead to structural changes in neural circuits. Structural changes in the circuitry is crucial for the seasonal control of reproduction. Morphological rearrangements cause seasonal inhibition of a certain hormonal secretion in adult ewes (Migaud et al. 2011). Similar hormonal changes can be internally triggered, as in the previous case, or caused due to external chemical influence. Administration of drugs, such as, amphetamine or methylphenidate alters the organization of dendrites in the prefrontal cortex, causing reduced play initiation, as well as impaired working memory functioning (Kolb et al. 2011). Furthermore, lesions or adverse conditions can induce changes in the network. Lesioning a part of the retina leads to an adjustment in the cortical topography, by causing the receptive fields of cells in the cortical scotoma to adapt to representing the retinal area surrounding the lesion (Gilbert 1998). Pathological conditions can cause major upheaval in the neuronal networks. Uncharacteristically higher spine densities has been observed in certain parts of the frontal, parietal and temporal lobes in patients with Autism Spectrum Disorder (Forrest et al. 2018). Continuous exposure or targeted training can also potentially cause structural changes. The difference in proficiency, between professional musicians, amateur musicians, and non-musicians, in discriminating tones has been traced to certain structural differences, including the higher gray matter volume in the left Heschl's gyrus in musicians (Steven et al. 2004). Structural plasticity clearly plays a sensitive and functional role in several processes during and after development.

2.2 Interaction between structural and functional plasticity

The previous section has demonstrated the ubiquitous presence of structural plasticity across species and across developmental and spatial scales. Structural plasticity encompasses a variety of mechanisms that depends on the level under consideration; the processes underlying spine turnover are very different from the ones that drive axon arborisation (Bosch et al. 2012; Lewis Jr et al. 2013). However, such processes of structural plasticity do not occur in isolation. We find abundant evidences wherein structural plasticity interacts with the other various forms of plasticity within neural circuits. As discussed earlier, plasticity in neural circuits ranges from the alteration of the intrinsic biophysical properties of a neuron, and the experience-dependent modification of the efficacy of synaptic transmission, to a massive overhaul of axonal connections within and across brain regions. Inducing synaptic plasticity not only affects the generation of action potentials by a neuron but also affects the intrinsic properties of a neuron, such as synaptic integration. For instance, induction of long-term potentiation in slices of CA1 pyramidal and Purkinje neurons from the rat hippocampus and cerebellum, respectively, has been shown to have strong effects on the intrinsic excitability of a neuron as well as dendritic integration (Belmeguenai et al. 2010; Campanac et al. 2008). Conversely, alteration of neuronal excitability, by blocking after-hyperpolarisation (AHP), enhances induction of long-term potentiation in the hippocampal pyramidal neurons (Sah et al. 1996). Reduced AHP of pyramidal neurons in the rat piriform cortex, following operant conditioning, also has been linked with improved learning capacity (Saar et al. 1998). Similarly, functional and structural plasticity are tightly inter-connected. Initiation of structural plasticity may be triggered due to an external lesion, or by an internal hormone release (Ooyen et al. 2017). In this wide landscape, there is an interesting case where structural plasticity is driven first and foremost by neural activity. De Roo et al. (2008) observed an increase in spine turnover following induction of long term potentiation in slice culture of neurons. Further, Oh et al. (2013) showed an activity-dependent shrinkage of dendritic spines. It has been also observed in hippocampal slice cultures that stimulation led to an increased stabilisation of stimulated synapses and faster pruning of non stimulated synapses (Bernardinelli et al. 2014; Leuner et al. 2010). On the other hand, such synaptic pruning can lead to an increased inhibitory nature of the circuitry, which can have varied impacts. For instance, the late development of an inter-neuron subset in the primary visual cortex causes the maturation of the inhibitory circuitry which plays a crucial role in the opening and closing of the sensitive period (Hensch 2005). Thus, we can see that structural and functional plasticity have a strong mutual interaction, and are both closely related to the neural activity of the related network. While, in the former examples, a functional change is able to trigger structural plasticity, in the latter case, structural plasticity shows a prominent functional impact. At a larger scale, the increase in inhibitory activity curtails the gener-

ation of new synapses and the elimination of existing ones, which ultimately leads to the stabilisation of the circuitry (Schaefer et al. 2017).

These observations suggest that certain instances of structural plasticity stem from functional changes, and in turn, have the potential to induce a prominent functional impact. It seems more than relevant to investigate the interactions between structural plasticity, neuronal activity and the function of the related networks. Given the complex interactions between network structure, activity and function, theoretical investigations in computational models will likely be required to shed light on the underlying mechanisms. In the next section, we discuss how structural plasticity has been sporadically incorporated into computational models.

2.3 Computational perspectives: From biology to machine learning

We have seen that structural plasticity is shared among different species, occurs (non-uniformly) across the whole lifespan and scales from the level of dendritic spines up to distributed cortical regions. It leads us to question why plasticity of this kind is as prevalent, and if it confers any benefits to the development of neural circuitry. Models of experience-dependent structural plasticity have shown that rules of dendritic and synaptic growth, derived from experimental data (Butz and Ooyen 2013), improve information transmission in networks with small-world topology (i.e. high clustering coefficients and short path lengths) (Butz, Steenbuck, et al. 2014). Knoblauch, Körner, et al. (2014) simulate structural plasticity by accommodating potential future synapses and find that this leads to a higher storage capacity per synapse than networks with only synaptic weight plasticity. Moreover, Chechik et al. (1998) show that energy efficiency of memory storage, indeed, is improved by using the strategy of excessive sprouting followed by synaptic pruning. Structural plasticity can also influence the development of synaptic connectivity within networks. Poirazi et al. (2001) show that memory formation or long-term information storage might also be governed by the co-activation of synapses on a shared dendritic branch rather than merely synaptic weights. Stepanyants et al. (2002) complement this study, using geometric analysis to demonstrate that dendritic spines improve information capacity by increasing the specificity of dendritic connections. Sailor et al. (2016) further demonstrate the functional role of dendritic arborisation of adult-born granule cell inter-neurons in the olfactory bulbs of adults. Simulations by Spiess et al. (2016) show that structural plasticity improves learning by decreasing response noise, and when combined with pruning, also reduces the training time. Further, models such as Lightheart et al. (2013) incorporate the effect of structural plasticity by using constructive algorithms to simulate network growth which also account for the creation of new neurons. Such comparatively slower structural plasticity could balance the trade-off between stability and plasticity in order to help evade catastrophic forgetting (Knoblauch, Körner, et al. 2014; Knoblauch and Som-

mer 2016). Knoblauch, Palm, et al. (2010) further demonstrate that this strategy also helps optimise information capacity in biological networks (Knoblauch 2017).

Despite these clear advantages of incorporating structural plasticity, it has seldom been used in machine learning and computational neuroscience literature. Most of the time, the architecture of a model is chosen prior to the learning phase and remains fixed until the end. There are of course notable exceptions. One of the earliest and best known example of a computational model taking full advantage of structural plasticity is the growing neural gas model (Fritzke 1995) that learns the topology of the input thanks to the dynamic topology of the network. Connections between units can be created or removed dynamically during the learning phase according to a precise criterion that depends on both the data and the history of the model. This dynamic reconfiguration allows the model to faithfully map the topology of the data as opposed to, for example, a regular self-organising map (Kohonen 1982) that uses a fixed topology.

A couple of years before this model was introduced, pruning algorithms (Reed 1993) were popularized by the Optimal Brain Damage algorithm (LeCun et al. 1990) that aim at finding the optimal size (in terms of the number of connections) of a feed-forward model in order to establish a trade-off between the complexity and the error magnitude of the resulting model. This algorithm starts from a fully connected feed-forward model (multi-layer perceptron) and removes, after training, the weights that have the lowest saliency. With the advent of deep-learning during the past decade and the inflating size of models, there has been a renewed interest in such pruning techniques (Blalock et al. 2020). However, pruning is the simpler side of structural plasticity since you start from a situation where you can objectify the influence of a neuron A onto a neuron B based on the existence of a connection between them. The careful analysis of this influence can then help to decide if the connection can be removed or not.

On the other hand, the case of sprouting is quite different because there is a need to establish a new connection between a neuron A and a neuron B that are, by definition, not connected. This implies that, from a structural point of view, this selection cannot be based on the correlation of activity, for instance, since no connection exists. Hence, one obvious way to perform the selection is to randomly choose the source and the target. This is the technique used in a number of generative models that search for the best architecture (Angeline et al. 1994). There exist however non random techniques, such as marker induction, as proposed by (Willshaw et al. 1979). In this model, markers are used to induce presynaptic fibers to connect to the post-synaptic sheet. This results in a neighbouring presynaptic region to connect to a neighbouring postsynaptic region in an ordered fashion and this model is used to explain the formation of ordered retinotectal projections in amphibians and fishes. There exist similar models for the development of the retinogeniculo cortical pathway (Elliott et al. 1999) and more generally, the formation of topographic maps (Bogdan et al. 2018; Sirosh et al. 1996). In addition to several widely applicable advantages of structural plasticity such as optimising energy consumption, as discussed above, structural plasticity, in the case of topographic maps (e.g. visual maps), can help

to *embed pattern of synaptic strengths in the network topology* (Bamford et al. 2010) and may explain the existence of a critical period as well. However, from a computational point of view, the advantage of such structural plasticity compared to, for instance, an initially fully connected model that is later pruned, is not clear. Departing from the pure formation of topographical maps, the model by Kappel et al. (2018) explains and exploits structural plasticity to give a functional account on the development of a dedicated architecture for solving reward based tasks. As explained by the authors, this model uses *reward-driven network plasticity as continuous sampling of network configurations* that results in policy sampling. This model is interesting because it provides a functional interpretation of structural plasticity and justifies its presence. Without such plasticity, it would be more difficult or longer for the model to solve the task.

However, we suggest that the functional role of the sprouting phase should be considered beyond the formation of topographic maps. We hypothesise that via the wide range of interactions discussed in this chapter, structural plasticity could potentially fulfil, not merely an auxiliary role in the functioning of a circuit (e.g. energy optimisation), but also a more direct role towards facilitating the intended function of a circuit. We will advocate in the next chapter that vocal learning in songbirds is a useful paradigm to understand the implications of structural plasticity on the organization of neural networks during development and beyond, on the network function and the underlying neural code, and ultimately on behavior.

3 Vocal learning by songbirds

The vocal learning behaviour of songbirds provide an avenue to study sensorimotor learning, along with the role of structural plasticity therein. Birds and mammals share striking similarities in both a behavioural and neuroanatomical sense. We outline the evidence accumulated in support of the hypothesis that the two taxa share behavioural traits as well as general principles in the organisation of brain networks, the constraints bearing on their construction during development and reorganisation following injury or natural re-modelling of brain circuits. Whether these common principles are conserved, or the result of convergent evolution, is open to discussion (see Güntürkün (2012)). Once the parallels have been established, we highlight that structural plasticity is ubiquitous in the bird brain, just as in mammals, as seen in chapter 2. Thus, we emphasize that songbirds form a suitable animal model to gain insight into sensorimotor learning and its neural mechanisms. We then focus our survey on a specific instance of sensorimotor learning, i.e. the vocal learning behaviour exhibited by songbirds. We look into the brain circuits involved in the acquisition and production of complex vocalisations in songbirds, where the relation between neural structure, activity and function has been best studied. To this end, we briefly describe the vocal learning circuitry, then delve into the details of structural plasticity in this circuitry and finally extract computationally useful properties herein.

3.1 Songbird behaviour

Akin to speech acquisition by infants, juvenile songbirds imitate the vocalisations of adult tutor songbirds. In this section, we advocate for the use of songbirds as a pertinent animal model to study sensorimotor learning in humans, owing to the parallels in behaviour between mammals and songbirds, and the tractability of the vocal learning behaviour exhibited by zebra finches.

3.1.1 General behavior in birds

The successful evolution of mammals in diverse ecological contexts is thought to rely, at least in part, to the evolution of their large brains, conferring them with behavioral flexibility and cognitive abilities and making them efficient predators and competitors. These cognitive abilities are diverse and include (but are not limited to) complex senso-

rimotor coordination and adaptation, memorization, planning and anticipation capacity, analytical and deductive reasoning, sophisticated social interactions, and introspective judgement. While many of these complex cognitive skills have first been thought to be the hallmark of humans, or exclusively present in mammals gifted with large brains, abundant experimental evidence now demonstrates that several bird species also display many, if not all, of the same cognitive functions as mammals (Güntürkün and Bugnyar 2016; Karten 2015). Sophisticated sensorimotor learning abilities are demonstrated in songbirds by song acquisition, with remarkable parallels to human speech learning, including its reliance on high auditory processing of auditory feedback, tight coordination of vocal muscles and social interactions (Doupe and Kuhl 1999; Mooney 2009). Various bird species display food caching behavior and rely on months-old spatial memory traces to retrieve their spread-out cached food (Sherry and Hoshooley 2010), illustrating their excellent episodic-like memorization and planning capacities. Tool use and other problem-solving tasks have been successfully solved by several bird species, many from the corvidae family (Striedter 2013). Many species of birds, including starlings, crows and parrots, have been found to maintain complex social organization among colonies of thousands (and more) of subjects (Boucherie et al. 2019; Downing et al. 2020), relying at least partly on vocal communication. Finally, parrots display cognitive abilities that we have until now considered unique to humans and other primates (conceptual representation, combinatory learning, counting (Pepperberg 2006)) and magpies can recognize themselves in a mirror (Prior et al. 2008), a skill thought to require a representation of oneself and only sparsely present in mammals.

3.1.2 Song acquisition and production

Songbirds use learned vocalizations to communicate during courtship or aggressive behaviors. Akin to speech learning in humans, vocal learning in young birds requires the coordination of vocal muscles to reproduce previously experienced adult vocalizations. Singing is a sensorimotor skill and song acquisition in juvenile birds is also highly dependent on hearing the adults they will imitate, as well as themselves as they practice, and this dependence wanes as the birds mature. Strikingly, the gene FoxP2, linked to speech learning in humans (MacDermot et al. 2005), is also implicated in avian song learning (Haesler et al. 2004). Similar neural mechanisms underlying vocal learning are most probably involved in humans and birds (Doupe and Kuhl 1999), and the study of these mechanisms in birds could shed light on the neurobiological bases of speech learning.

Interestingly, in birds, as in humans, imitative vocal learning is characterized by several phases (Doupe and Kuhl 1999): (i) the sensory phase enables human infants or juvenile birds to build a neural representation of adult vocalizations, which will guide later vocal production; (ii) during the sensorimotor phase the young subjects start to vocalize, initially producing babbling sounds and then adapting its vocal output to imitate previously heard vocalizations; (iii) finally the produced vocalization becomes more and

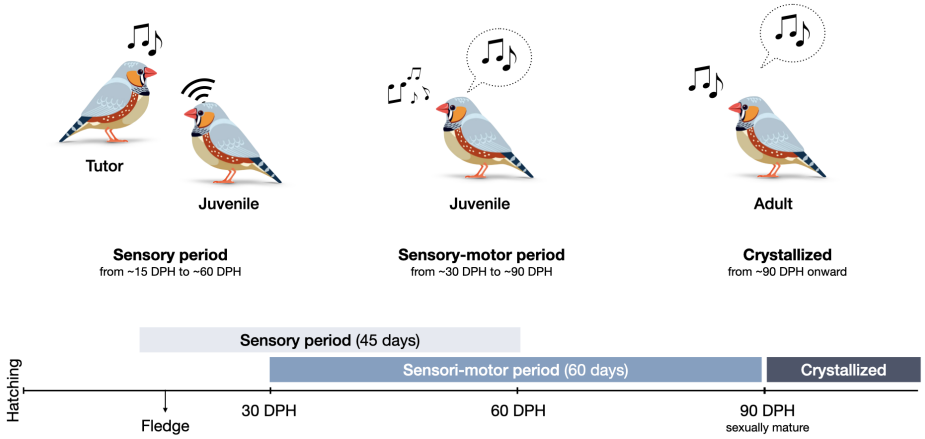


Figure 3.1: Phases of the vocal learning behaviour in zebra finches. Juvenile songbirds form a neural representation of tutor vocalisations by listening to the adult male tutor in the initial sensory phase of sensorimotor learning. In the sensorimotor period, the bird produces random vocalisations and gradually attempts to imitate the tutor song. Towards the crystallised phase, the produced vocalisations become more stereotyped with a reduction in variability.

more stereotyped and vocal plasticity significantly drops (Figure 3.1). This final phase is called crystallization in birds. Each of these phases is bounded by a sensitive period (Mooney 2009). In most bird species, if the young subjects have not experienced conspecific adult vocalizations before a species-specific age limit (e.g. 60 days post hatch (dph) in zebra finches, around 2 years in children), imitation will be virtually impossible because the sensory phase of vocal learning is closed past that age. Similarly, the sensorimotor phase closes with puberty in many species (around 90 dph in zebra finches, 12-14 years in humans) when vocal exploration drastically decreases (and vocalizations become more stereotyped), making vocal imitation more difficult and reducing greatly the vocal plasticity. Seasonal birds may re-open the sensorimotor phase yearly and therefore display cyclic sensitive periods locked to seasonal changes. These sensitive periods and the underlying phases of vocal learning are driven by physiological signals that constrain the development of neural circuits in the central nervous system, and thereby affect the behavioral plasticity. It is important to note that while avian song learning may share some similarities with human speech acquisition, bird song does not share the semantical complexities of human language.

3.1.3 Salient features of vocal learning

Having established a parallel between human speech acquisition and avian vocal learning, we now focus on features that are characteristic of vocal learning in songbirds and more specifically, zebra finches. Later, we attempt to incorporate some of these features within our theoretical investigation of the circuitry underlying vocal learning.

As noted in the previous subsection, vocal learning in zebra finches (*Taeniopygia guttata*) proceeds through three stages: the sensory phase, where the juveniles form a memory representation of the tutor song, the sensorimotor phase, where the juvenile engages in vocal exploration and attempts to imitate the tutor song, and finally, the crystallised phase, where each rendition of the bird's song is more stereotyped. The duration of each song learning stage differs across species. For zebra finches, the sensory stage lasts until around 65 dph, while the sensorimotor stage lasts for a duration of approximately 60 days, between 30-90 dph (see Figure 3.1) (Immelmann 1969) .

Birds raised in isolation, or deafened before tutor exposure, produce abnormal songs, termed 'isolate' songs. Aligned with the idea of sensitive periods, birds exposed to song post their sensory phase, are incapable of producing a recognisable imitation of tutor song (Marler 1970). Juveniles are not limited to imitating only the song of their male parent. They can learn to imitate the song of a non-parental conspecific or heterospecific adult tutor (Marler and Peters 2010).

In zebra finches, each song motif is composed of stable sequence of syllables and intervals of silence. During the sensorimotor and crystallised phase, birds can produce around 1000 to 3000 song motifs per day, with each motif being 300-1000ms long (Glaze et al. 2013). Birds deafened post tutor exposure are also unable to learn tutor song. This demonstrates the need for birds to listen to their own vocalisations during the sensorimotor stage, possibly to improve their imitation via self-evaluation (Konishi 1985).

Trial-to-trial variability in song renditions reduces gradually over these stages, with adult birds displaying a coefficient of variation of 2-3% in the fundamental frequency of their syllables (Arnold 1975; Tumer et al. 2007). While birds retain variability during adulthood as well, it is further reduced when the song performance is directed towards a female, as compared to undirected singing (Kao et al. 2005; Sakata et al. 2008). It is interesting to note that when female songbirds are presented to male juveniles around 60dph (well before crystallisation), the male songbirds are already capable of producing song which is more similar to their eventual adult song than their current undirected song (Kojima et al. 2011). This indicates the ability of a juvenile songbird to continue vocal exploration of its sensorimotor capabilities, despite having learnt more preferable imitations. Additionally, songbirds demonstrate a wider repertoire of vocalisations during their sensorimotor phase, than during adulthood (Nelson et al. 1994). Akin to human speech acquisition, social interactions play a key role in vocal learning. Quality of song imitation is superior when juvenile birds are trained using live models than by passive playback, similar to human infants (Catchpole et al. 2003; Kuhl 2010).

The role of sleep during song learning in birds is currently under investigation. Derégnaucourt et al. (2005) observe a pronounced deterioration in song quality post sleep, although a similar deterioration is not observed post long breaks in song production during awake periods. Most of the improvement in imitation quality, within a day, occurred in the post-sleep recovery period, and little improvement was observed in the later periods of the day. The daily deterioration was much less pronounced in the crystallised phase. Interestingly, the extent of such post-sleep deterioration in the sensorimotor period also seemed indicative of quality of tutor song imitation in the crystallised phase. This observation suggests the possibility of a non-monotonic progression in quality of imitation.

A similar non-monotonic trajectory in progression of sound acoustics has been observed by Tchernichovski et al. (2001). During song learning, juvenile birds often engage in pitch doubling. For instance, instead of gradually reducing the fundamental frequency of a particular vocalisation to match the target template, the bird increases the fundamental frequency and then doubles the period in order to produce an imitation of the target vocalisation. Thus, considered along with the daily post-sleep deterioration observed in song similarity during vocal learning, there are multiple indicators that songbirds may follow non-monotonic trajectories towards the target song. Additionally, using a novel technique of repertoire dating, Kollmorgen et al. (2020) observe that after 40dph, the within-day change in vocal behaviour can be partly misaligned with the slower change over the timescale of weeks.

Once the song has been learnt, birds experience deterioration in song quality, if deafened. This deterioration is more pronounced if the bird is deafened during adolescence than during adulthood. This may serve as an indicator to the need for active maintenance of song, as well as, the diminishing ability to alter vocal behaviour, with age (Nordeen et al. 1992).

The ability to alter vocal behaviour, or learn new songs, is dependent on the species of songbirds. Zebra finches, etc, belong to a set of 'closed-learners' who experience a single song learning period during their lifetime. On the other hand, canaries, etc, are 'open-learners' and can continue to learn new songs during adulthood, for instance, in each season (F. Nottebohm, M. E. Nottebohm, L. A. Crane, et al. 1987). Songbirds are capable of demonstrating adult plasticity. They retain a limited ability to introduce slight modifications to their song structure in adulthood, both in terms of maintenance of song as well as in response to external stimulus (Tumer et al. 2007; Warren et al. 2011). When subjected to a differential feedback protocol (detailed further in section 6.3.1), adult birds are able to selectively adapt the corresponding syllabic features.

What are the changes undergone during vocal learning by the related neural circuits? Can we understand the various behavioural phases in light of circuit changes? In other words, how does structural plasticity related to the development of brain circuits underlying vocal learning affect the imitation process? We will illustrate in the next section how behavioral, anatomical, neurochemical and electrophysiological data collected in song-

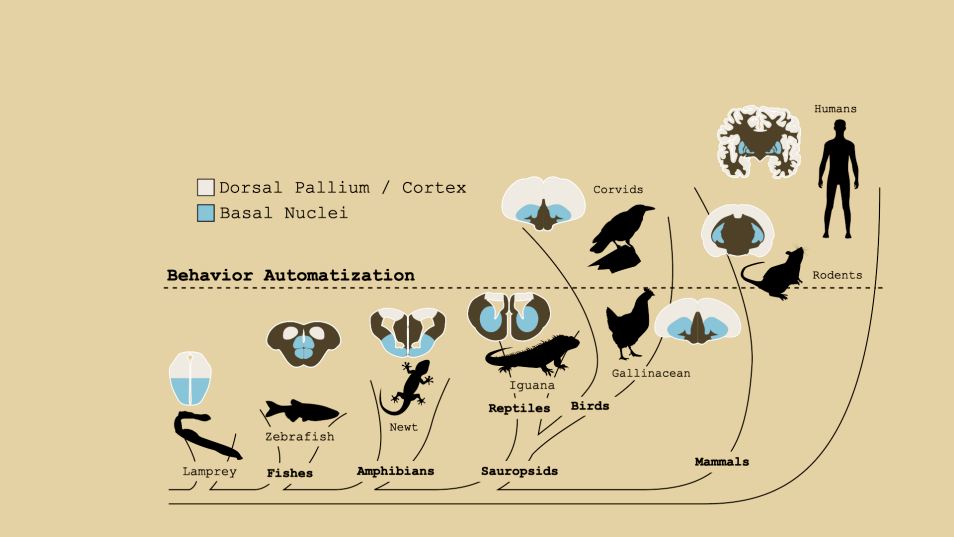


Figure 3.2: Comparative evolution of the striatum and pallium in vertebrates (Adapted from Boraud et al. (2018a)).

birds (mostly in zebra finches) may help bridge the gap between structural plasticity and function.

3.2 Songbird anatomy

As described in the previous section, the similar cognitive abilities of birds compared to mammals suggest that the organization of their nervous system face similar challenges to optimize their behavior. Studying the emergence of brain function during development or re-organization in adulthood in birds is, thus, just as promising as it is to rely on mammalian animal models. In this section, we look into the general configuration of brain structures in songbirds, and more specifically, at the neural circuitry underlying the vocal learning behaviour.

3.2.1 General brain anatomy in birds

The complex behaviors displayed by birds and briefly summarized above rely on brain circuits that display, for the most part, major biochemical, anatomical, and physiological differences as compared to their mammalian counterparts, but still remain surprisingly analogous in their functional organization (Güntürkün 2012; Karten 2015). Basic avian brain anatomy already differs from that of mammals, the most notable discrepancy being the absence of a laminated neocortex in the avian telencephalon (Güntürkün 2005; Herold et al. 2019). Rather, the avian pallium is inherited from the reptilian Dorsal Ventricular Ridge and is organized into a largely continuous field of nuclei. For this reason, the avian

telencephalon has long been considered to mainly consist of a hypertrophied basal ganglia structure (or paleostriatum, see (Reiner, Medina, et al. 1998; Reiner, Perkel, et al. 2004)). While this organization appears very different, the nuclei of the avian pallium show similar connectivity, neuronal types and functional properties to those of the mammalian cortex, amygdala, and claustrum. Within this area, birds clearly display a hippocampus, piriform cortex and olfactory bulb. The hyperpallium in birds (previously/also called the Wulst), also included in this area, displays a strong analogy with the primary visual area and a primary somatosensory area of the mammalian cortex, in terms of thalamic input, connectivity, and electrophysiological properties (Medina et al. 2000). The Nidopallium contains various auditory areas that display a similar functional organization as the mammalian auditory cortex, although the spatial arrangement of the various network components is very different in birds (Calabrese et al. 2015; Wang et al. 2010). At a finer scale, comparative developmental analysis and neurochemical data reveal a surprising extent of similarities in the neuronal subtypes and among birds and mammals (Butler et al. 2011; Montiel et al. 2016).

In the ventral part of the avian telencephalon, 'subcortical' structures are older from an evolutionary point of view as compared to the pallium, and therefore share even more similarities with mammalian subcortical structures. In particular, the avian brain contains homologues of the mammalian septum, basal ganglia (BG) and several other nuclei, as unveiled by developmental, topological, neurochemical, cellular, connectional and functional data (Doupe, Perkel, et al. 2005). Moreover, a modern revision of the avian anatomical nomenclature has now provided a common language for studying the function of the avian subcortical nuclei (Reiner, Perkel, et al. 2004).

Concerning the BG, there is an avian circuit that has been looked deeper into: the song-related BG-thalamo-cortical loop of songbirds. Indeed, songbirds have specialized a portion of their forebrain–BG circuitry expressly for the purpose of song learning. Recent advances in anatomical, physiological and histochemical characterization of avian BG neurons and circuitry have revealed a homology between this circuit and the mammalian motor loop of the BG-thalamo-cortical network (Bottjer and Johnson 1997a; Reiner, Medina, et al. 1998). In particular, the song-related BG nucleus Area X differs from mammalian BG in its gross anatomical structure, but it displays surprisingly similar circuitry at a finer scale. It is embedded within a region homologous to the mammalian striatum and receives a particularly strong dopaminergic projection (Lewis et al. 1997). It contains neuron classes corresponding electrophysiologically and morphologically to those in the mammalian striatum (Carrillo et al. 2004; Farries and Perkel 2002). In addition, Area X contains a class of neurons with pallidal properties, which directly project to the thalamic nucleus DLM (Leblois, Bodor, et al. 2009; Luo and Perkel 1999). We will focus below on the song-related circuits in songbirds, including this BG-thalamo-cortical loop.

Given the strong homology (for subcortical structures) or analogy in the organization of brain circuits driving complex behavior between birds and mammals, it is most likely that the functional organization of these circuits rely on the same basic principles. There-

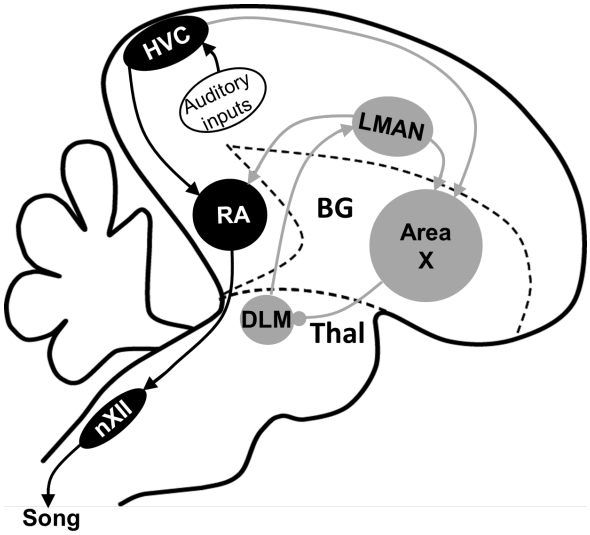


Figure 3.3: Schema of the neural substrates involved in vocal learning in zebra finches. The cortical pathway, shown in black, comprising the cortical nuclei HVC (premotor) and RA (motor), is primarily responsible for vocal production. The anterior forebrain pathway (AFP, show in grey), consists of the cortical nucleus LMAN, Area X (song-related BG nucleus) and thalamic DLM, and plays a crucial tutor role in vocal learning.

fore, studying the computational advantages of structural plasticity in birds will improve our understanding of this basic principle and provide generalizable theories for how development and remodeling acts to (re)shape brain function. Now, we will discuss the specific case of the vocal learning circuitry in songbirds.

3.2.2 The song system: anatomy and function

In birds, the sensorimotor skill of song production and its learning has a dedicated set of interconnected brain nuclei collectively known as the "song system" (see Figure 3.3), that ultimately coordinate the patterned breathing and vocal muscle activity necessary for vocalization. This dedicated circuit makes the songbird an outstanding model to study the neural mechanisms of vocal learning and more generally, the function of neural circuit in sensorimotor learning.

The song system includes a 'motor pathway' that is required throughout life for normal song production, and a BG-thalamo-cortical circuit necessary for song learning, plasticity and maintenance, called the 'anterior forebrain pathway' (AFP) (Brainard et al. 2002). The song system receives auditory information through the projections from the high-level auditory areas Field L, NCM (caudo-medial nidopallium) and CM (caudal mesopallium), and the pallial regions analogous to the auditory cortices of mammals (see section 3.2.1), to the nucleus HVC (used as a proper name) (Hahnloser and Kotowicz 2010).

The motor pathway includes the premotor cortical nucleus HVC and the robust nucleus of the arcopallium (RA), which is functionally equivalent to the laryngeal motor cortex. The HVC is involved in generating the timing and sequencing of song (Hahnloser, Kozhevnikov, et al. 2002; Long et al. 2008). RA receives inputs from the HVC and projects directly to respiratory centers and to the brainstem motor neurons controlling the vocal organ; the respiratory centers send recurrent information back to the HVC via the thalamus, reflecting the importance of bidirectional coordination between telencephalic and brainstem structures in vocal control (Brainard et al. 2013; Schmidt et al. 2012).

The BG-cortical loop indirectly connects the HVC and RA, and like the cortico-basal ganglia circuitry in mammals, plays a crucial role in motor learning. It consists of three nuclei connected in a loop: the BG nucleus Area X, the medial portion of the dorsolateral thalamic nucleus (DLM), and the lateral magnocellular nucleus of the anterior nidopallium (LMAN). This loop is closed in the macroscopic sense (i.e. the projections form a recursive loop). The LMAN - Area X, Area X-DLM and DLM - LMAN projections within the loop are topographic (Luo, Ding, et al. 2001; Vates et al. 1995), thus demonstrating that this BG-cortical loop is also microscopically closed. The song-related BG nucleus Area X receives strong dopaminergic innervation from the midbrain (Bottjer and Johnson 1997a). Because this specialized cortico–BG circuit is discrete and devoted to a specific well-defined and naturally learned sensorimotor task rather than a broad range of motor behaviors, it is particularly tractable for elucidating the interwoven sensory, motor and reward signals carried by BG, and the function of these signals in skill learning and execution (Doupe, Perkel, et al. 2005).

The cortical motor pathway from the HVC to the RA is responsible for song production, by projecting to downstream structures that control respiration and syringeal musculature. The HVC is composed of at least three main types of neurons (Mooney 2005). One set of neurons project to the downstream nucleus RA and is necessary for song production (J. Kirn et al. 1991). The second set of projection neurons signal to the avian BG homologue, area X, and is required for vocal plasticity (Prather et al. 2010). The third set of neurons function as inhibitory inter-neurons (Kubota et al. 1998). Hahnloser, Kozhevnikov, et al. (2002) show that individual RA-projecting HVC neurons produce sparse bursts of less than 10ms, at a single, precise time during the song motif. Thus, as a population, these HVC neurons may form a representation of time in the sequence. In turn, these HVC neurons project to the RA, the site of motor control. The RA consists of excitatory projection neurons and inhibitory interneurons, as well. Outside of singing bouts, the RA projection neurons display a tonic regular spiking activity of 10-20Hz. During singing, RA neurons display highly variable firing patterns in juvenile birds (Ölveczky, Otchy, et al. 2011). As development proceeds, the motor program underlying crystallized song in adult birds, develops into phasic activity of highly stereotyped sequences of bursts in RA (Yu et al. 1996). These sparse bursts are precisely timed-locked to the song (Ölveczky, Otchy, et al. 2011). Immediately post singing, all RA neurons experience strong inhibition, lasting 400 to 800 ms, after which the neurons return to their baseline tonic activity. The

RA neurons modulate song at a timescale of 10-20ms (FEE et al. 2004; Sober, Sponberg, et al. 2018; Vu et al. 1994). Notably, surgical transection of the HVC-RA fibres completely eliminates normal adult song (Aronov et al. 2008). Thus, the cortical motor pathway is indispensable for generating vocalisations.

On the other hand, during song acquisition, the integrity of the song-related BG-cortical loop is required for proper imitation of adult vocalizations (Bottjer, Miesner, et al. 1984; Scharff et al. 1991). Although its precise role in song learning has long remained mysterious, recent experimental studies have shed light on its function during learning. Multiple lines of experimental evidence support a role of the BG-cortical loop in driving and modulating acoustic variability in song. In particular, lesions or inactivations of the cortical output of the loop (LMAN) reduce the acoustic variability of plastic song in juvenile birds (Ölveczky, Andalman, et al. 2005; Warren et al. 2011) and changing dopaminergic input to the circuit triggers changes in song variability (Leblois, Wendel, et al. 2010). The song-related BG circuit also guides changes in motor output by generating an error-reducing motor bias that is rapidly incorporated into the cortical premotor network (Andalman et al. 2009a). In birds, as in mammals, it is believed that the adaptive motor changes elicited by the BG are learned through a reinforcement learning mechanism. Accordingly, dopamine delivery in the BG provides an online evaluation of song quality (Gadagkar and Goldberg 2016) analogous to reward prediction errors signaled by dopamine during conditioning in mammals (Schultz et al. 1997). This dopaminergic signal provides a reinforcement signal that drives error correction during learning (Hisey et al. 2018; Xiao et al. 2018). The dopaminergic input is known to modulate activity-dependent synaptic plasticity in area X in birds (Ding et al. 2004b), akin to mammals (Wickens et al. 2003). Thereby, the BG output implements behavioral adaptations guided by dopaminergic reinforcement signaling to optimize motor output. These behavioral adaptations are ultimately consolidated in cortical networks following extensive training (Andalman et al. 2009a; Hélie et al. 2015).

The whole network of brain nuclei involved in song production in songbirds - the song system - offers a great model for studying brain circuits and their role in behavior. The numerous genes implicated in human diseases that are enriched in the song system nuclei and exhibit differential expression, both at specific points in development and during behaviors, such as listening to song or singing, also opens up interesting paths for the study of pathophysiological processes (M. S. Fee and Scharff 2010; Wada et al. 2006). We argue that studying the benefits and consequences of structural plasticity in this system, through a functional standpoint, may shed light on the general principles that make structural plasticity advantageous for the central nervous system in the perspective of learning complex skills.

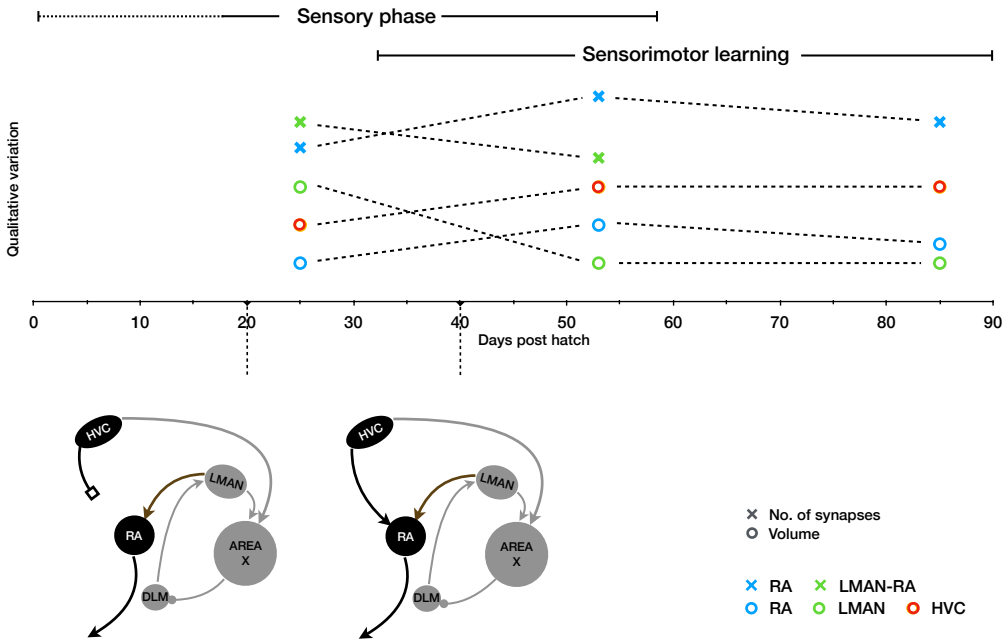


Figure 3.4: Timeline of development of the vocal learning circuitry in zebra finches (Bottjer, Miesner, et al. 1986; Herrmann and Arnold 1991; Herrmann and Bischof 1986; Johnson et al. 1992; Konishi 1985; Mooney and Rao 1994b; Nordeen et al. 1988; Sohrabji et al. 1993). The bottom panel shows the axons from HVC entering the RA, to form the cortical pathway responsible for vocal production, much after the anterior forebrain pathway (theorised to provide a tutor signal for vocal learning) is completed (bottom panel). This is accompanied by significant changes in the neural regions involved. During the initial sensorimotor period, the volume of the RA and HVC increase with a significant decrease in the LMAN volume, while the synapses within the RA increase, both from HVC axons and from RA interneurons. In the later stages of the sensorimotor learning, decreasing song variability is accompanied with a stabilisation of the volume of the LMAN and HVC and a slight decrease in RA volume.

3.3 Structural plasticity in songbirds

Similar to mammals, birds display plasticity from the scale of dendritic spines to synapses to networks throughout their lifespans, both during development and in adulthood. Internal factors (such as the concentration of circulating hormones) and external factors (such as daily light cycles, but also injury or lesion), can trigger plasticity in birds. For instance, the avian olfactory pathway which projects into paralimbic areas, before hatching, is displaced competitively by the development of visual, auditory and motor pathways after hatching, as a result of higher audio-visual exposure (Teuchert-Noodt et al. 1991). Neurogenesis and circuit reorganization is also common in adult birds. The hippocampus undergoes seasonal neurogenesis and reorganization in food-storing birds and brood parasites (Sherry and MacDougall-Shackleton 2015), and higher exposure leads migrating species (passerines) to undergo more neurogenesis in areas that process spatial information, than non migratory species (LaDage et al. 2011). Interestingly, similar mechanisms underlie adult neurogenesis in both avian and mammalian species (Brainard et al. 2013). The brain circuit where structural plasticity has been most studied in birds is arguably the song system in songbirds (Brenowitz 2004). In this section, we provide a short review on the current evidence for structural plasticity in the vocal learning circuitry in birds.

During development

Song acquisition happens in juvenile birds while their brain is still in development. As a consequence, the song-related neural circuits are still in the process of being built and there are many forms of structural plasticity in the network, during the song learning period.

Auditory experience may already have an influence on brain development before hatching (Mariette et al. 2016) and the auditory pathway from the ears to high auditory areas in the pallium is built early in development. The anatomical structure of the auditory network is thus largely mature in young birds when the sensory learning phase starts (dph 20), with all major anatomical connections being present. Exposure to adult vocalization, at that stage, shapes the auditory responses of neurons in the high-level auditory nuclei of the pallium where song-selective responses occur relatively early (before dph 35) (Amin et al. 2007). Receiving strong afferents from these auditory nuclei, neurons in HVC and its efferent structures show a progressive emergence of song selectivity during the sensory period of song learning (Doupe and Kuhl 1999). A. N. Chen et al. (2020) demonstrate that during development, there are significant changes in the phasic excitability in the caudal mesopallium CM, a cortical auditory region which exhibits selective responses to familiar conspecific songs. Exposure to tutor song has been shown to induce cell-type specific changes within the ion channel expression patterns of HVC neurons, depending on whether they project to the RA or the Area X. These changes happen at a magnitude that can potentially alter network function (Ross et al. 2017). Moreover, daytime expo-

sure to tutor song had profound effects on the distributions of inter-spike intervals of RA bursting activity during sleep, depending on the tutor song (Shank et al. 2009). Deafening adult birds, on the other hand, decreases dendritic spine size and stability of Area X projecting HVC neurons, along with an increase in mean spontaneous action potential firing rates and decrease in inter-spike intervals (Tschida et al. 2012). Daou et al. (2020) investigate the intrinsic properties of Area X projecting HVC neurons further, and discover that birds displaying good copies of their tutor's song exhibit similar spike morphology, spike trains, and ion current magnitudes, indicating that these intrinsic dynamics may be molded by auditory and vocal experience. Additionally, considering conditioning using distorted auditory feedback also elicits a prominent effect on the intrinsic properties of HVC neurons in adults, it has been suggested that such intrinsic plasticity could form an alternate/complementary mechanism for learning via their effect on synaptic efficacy and network interactions (Daou et al. 2020; 2021). Further, functional plasticity is likely involved in the shaping of auditory responses in this network (Fiete, Senn, et al. 2010). The synaptic rewiring in HVC is also evidenced by a large turnover of dendritic spines early in the sensory learning phase, followed by a rapid stabilisation of spine dynamics (Roberts et al. 2016). This correlation of circuit changes with tutoring may indicate a functional role of structural plasticity in sensorimotor learning. On the other hand, with development, spine turnover in the HVC decreases in untutored birds too. Hence, this stabilisation might potentially underlie the closing of a sensitive period, where they are more receptive to learning behavioural change. The underlying cellular and molecular mechanisms triggering the changes in plasticity are still under investigation. Interestingly, perineuronal nets surrounding groups of neurons and limiting axonal and dendritic processes could mediate changes in plasticity (Cornez et al. 2018). Further work is however required to confirm their causal role in the opening and closing of sensitive periods.

Contrary to the ascending auditory pathway, the song system is highly immature when birds start the song learning process, and structural plasticity is ongoing both during the sensory and the early sensorimotor period of song acquisition (Figure 3.4). Among the two pathways of the song system, the BG-cortical loop develops first and is fully formed by dph 20, before the sensorimotor phase begins (Mooney and Rao 1994b). Indeed, in juvenile male zebra finches, at the onset of song learning, axonal terminals from LMAN enter Area X (dph 20), Area X terminals enter DLM by dph 20 (Sohrabji et al. 1993), and DLM terminals enter LMAN before dph 15 (Johnson et al. 1992). This BG-thalamo-cortical circuit is topographic in nature and reaches maturity before dph 20. (Iyengar et al. 1999; Vates et al. 1995). Concerning the input pathway of the BG-thalamo-cortical loop, HVC neurons projecting to Area X innervate this target by dph 20 (Alvarez-Buylla, Theelen, et al. 1988; Nordeen et al. 1988). The output pathway of the loop is also formed early, with LMAN projections entering the RA by dph 15 (Herrmann and Arnold 1991). The number of synapses made by LMAN axons onto RA neurons decreases substantially (-70% between dph 20 and 60) over the course of vocal development (Herrmann and Arnold 1991). While anatomical connections between the nuclei are stable from dph 20 to adulthood,

the nuclei of the BG-cortical pathway nonetheless experience dramatic volume changes. Between 20 dph and 60 dph, LMAN volume and neuron number decline by 50%, while the volume of Area X increases by 50%, due to neurogenesis (Bottjer, Glaessner, et al. 1985; Nordeen et al. 1988) .

The motor pathway, on the contrary, develops very late in the song learning process. First, both HVC and RA grow slowly, reaching adult size only after dph 40 and dph 70, respectively (Herrmann and Bischof 1986). In HVC, the number and size of neurons both increase dramatically (Bottjer, Miesner, et al. 1986). Meanwhile, the synapses in the RA, from the HVC and LMAN are drastically rearranged at the onset of the sensorimotor phase. HVC axons reach the dorsal border of RA by dph 15, but, unlike the LMAN, they form synapses with the RA only between dph 30 and 40 (Herrmann and Arnold 1991; Mooney and Rao 1994b). In the sensorimotor period, between dph 20 and 60, the reduction of synapses between the LMAN and RA is accompanied by a tripling of the number of synapses within the RA and an increase of the RA volume (Bottjer, Miesner, et al. 1986; Herrmann and Arnold 1991; Herrmann and Bischof 1986).

The most striking consequence of the late development of the song system is that the motor pathway, which drives adult vocalizations, does not contribute to the babbling vocalization, called subsong, produced by juveniles in the early sensorimotor phase of song learning (Aronov et al. 2008). At that stage, the HVC-to-RA pathway is not yet functionally strong enough to drive singing and LMAN projections are driving subsong-related activity in RA. During the following days, the formation and strengthening of inputs from HVC to RA leads to the emergence of plastic song, driven both by LMAN and HVC (Garst-Orozco et al. 2014b; Herrmann and Arnold 1991). As learning progresses, the influence of the LMAN inputs to the RA reduces due to the strengthening of the synaptic connections from the HVC. Further, the variability of the song decreases with development and practice (dph 60–90), along with a significant pruning of the HVC inputs to the RA, albeit with a strengthening of the surviving connections (Garst-Orozco et al. 2014b).

During seasonal plasticity

Post-development, hormonal changes and external factors continue to induce structural changes in the brain in many vertebrates beyond the developmental period (Jacobs 1996). Most temperate songbird species breed seasonally and display a pronounced seasonal plasticity in their singing behavior (Ball et al. 2010). The song initially learned as a juvenile undergoes a pattern of yearly changes triggered by changes in the light cycle and mediated by sex steroid signaling in the brain (Alward et al. 2013; F. Nottebohm, M. E. Nottebohm, L. A. Crane, et al. 1987). Producing highly stereotyped songs during breeding, birds stop singing during the summer molt, resume singing in fall with short-duration songs of variable structure at lower rate and volume and gradually sing longer, louder and more stereotyped songs until the next breeding season (Brenowitz 1997; F. Nottebohm 1981). These

changes in song are accompanied by a large structural reorganization of the neural circuits in the song system (F. Nottebohm, M. E. Nottebohm, and L. Crane 1986), auditory areas (Caras et al. 2012) and beyond (De Groof et al. 2008). This naturally occurring plasticity in the songbird brain is perhaps the most pronounced observed in any adult vertebrate. The volume of many song-related brain regions, as well as their number of neurons and synapses, increase dramatically in anticipation of the breeding season (Brenowitz 2004; F. Nottebohm 1981), while the physiological and functional properties of their neurons are altered (Del Negro et al. 2002; Meitzen et al. 2007). The entire volumes of several song nuclei, including HVC, RA and Area X, are considerably larger during the spring breeding season than during autumn and winter in seasonal birds (Tramontin et al. 2000). The increase in volume may reflect an addition of new neurons through neurogenesis, notably in the HVC (Alvarez-Buylla and J. R. Kirn 1997; Goldman et al. 1983), or dendritic growth and synaptogenesis, with an increase in the number of dendritic spines, as seen in the RA or Area X (Hill et al. 1991). Meanwhile, neuron number in LMAN and their inputs to RA remain relatively constant (Thompson et al. 2007). Altogether, it is interesting to note that the strengthening of the HVC-RA pathway before the breeding season recapitulates at least in part the late development of the motor pathway during song acquisition in juveniles. Interestingly, seasonal plasticity is regulated by the same endocrine signals as juvenile song learning (Korsia et al. 1991; Whaling et al. 1995) and may therefore exploit similar mechanisms as those acting during early ontogeny (Cajal 1928).

3.3.1 Functional impact of structural plasticity

There are parallels in the manifestation of structural plasticity between mammals and birds. Insights drawn from the songbird literature can be used to understand neural circuitry and behaviour in mammals, as well. They present an avenue to explore the contributions of changes in neural circuitry beyond functional plasticity.

Certain fundamental elements of structural plasticity can potentially confer benefits when incorporated into computational models. During song learning, there is an early excessive sprouting of neurons in HVC and RA, followed by a significant pruning of the HVC-RA synapses, and strengthening of the surviving connections (Herrmann and Arnold 1991). This form of abundant sprouting followed by selective pruning is a commonly adopted technique to model changes in network connectivity. This phenomenon can also be employed to encode the trade off between exploration and exploitation, to represent decreasing tendencies to explore, as learning proceeds (Garst-Orozco et al. 2014b; Tumer et al. 2007). Also, spine turnover and terminal size can be useful parameters in modeling the consolidation of learned motor action or behaviour, in lieu of solely modifying synaptic weights (Herrmann and Arnold 1991; Roberts et al. 2016). Beyond structural plasticity, it could also be interesting to consider integrating features of intrinsic plasticity within computational models, for instance, by simulating the alteration of intrinsic

excitability of a neuron by introducing activity-dependent modifications within the activation functions in typical neural network models, as suggested by Daou et al. (2021).

Apart from the general usage of these basic elements to computationally exploit the properties of structural plasticity, there are specific insights that can be drawn from vocal learning in songbirds, where structural plasticity directly has a crucial functional impact on learning. As we have seen in the section 3.3, a common feature of the developmental and seasonal structural plasticity in the song system is the progressive strengthening of the motor pathway and its increased synaptic drive on RA neurons compared to the early shaped and steady input from LMAN. Interestingly, in both cases, induction of large variability in the song by LMAN occurs in the condition where HVC inputs to RA are not as prominent (Alliende et al. 2017; Aronov et al. 2008). The putative effect of this relative change in the balance of inputs from HVC and LMAN in RA is crucially influencing the activity of neurons in RA and how they mediate changes in song.

Both in the early phase of juvenile learning and during fall in seasonal songbirds, LMAN drives variable patterns of activity in RA, resulting in less structured and more variable songs (Alliende et al. 2017; Ölveczky, Andalman, et al. 2005). Accordingly, the singing related activity of RA neurons gradually changes from highly variable firing patterns to precise and sparse bursts of spikes locked to song motifs during juvenile learning (Ölveczky, Otchy, et al. 2011). LMAN input is necessary for the expression of RA firing variability, and the change from variable to stereotyped firing patterns in RA throughout development could be explained by the strengthening and pruning of HVC inputs to RA (that drive stereotyped patterns) while LMAN inputs remain unchanged (Garst-Orozco et al. 2014b). As HVC input to RA becomes larger, it drives stronger bursting in RA interleaved with hyperpolarized periods of silence due to recurrent inhibition (Ölveczky, Otchy, et al. 2011). LMAN inputs to RA are mediated mostly through voltage-dependent NMDA receptors (Mooney and Konishi 1991; Ölveczky, Andalman, et al. 2005), and its influence on RA firing is weak in this condition. The influence of LMAN on RA is thus diminished, resulting in a progressively more stereotyped song (Ölveczky, Andalman, et al. 2005). Interestingly, this scenario does not require any seasonal or developmental modification in the BG–forebrain circuit to see its influence on song strongly modulated. It ensures that LMAN-driven variability is expressed long before the HVC driven temporal structure of the song motif emerges, and likely results in the two-phase learning, including early babbling in subsong followed by temporally structured plastic song.

The strong influence of the LMAN on the RA early in development initially drives subsong and then drives the variability in plastic song; thereafter, HVC plays an increasingly strong role in driving the stereotyped firing of RA as the bird approaches crystallization. The same scenario seems to be recapitulated at least partially during seasonal plasticity due to the regrowth of HVC and the entrance of new axons in RA. Why is HVC input arriving so late in the RA during song acquisition or during seasonal relearning? And why is learning divided in these two phases - babbling and plastic song? In the next

chapter, we draw a parallel with optimisation techniques, commonly used in machine learning, to investigate the potential roles of these phenomena in sensorimotor learning.

3.4 Two-stage learning

Beyond the specific case of birds, for which non-seasonal structural plasticity has been hypothesised to be critical for both acquisition and automatization in juvenile songbirds, we may question to what extent such structural plasticity is similarly critical in vertebrates and more specifically, in mammals. The similarity in brain organization between birds and mammals is indeed striking, especially when considering the cortex-basal ganglia loop. As discussed in (Boraud et al. 2018a) how *the development of automatized skills relies on the BG teaching cortical circuits and is actually a late feature linked with the development of a specialized cortex or pallium that evolved in parallel in different taxa*. Alongside this hypothesis, there is a growing number of computational models of decision making that takes into account this dual pathway hypothesis, where the BG acts as a general training machine for cortico-cortical connections (Hélie et al. 2015). In this context, it is natural to wonder if the initial delay in connection might exist as well in other species and if this would provide a similar benefit in early learning. In humans, this would correspond to the phase preceding the babbling phase that is known to be characterized by rapid structural and functional changes. However, as explained by Vasung et al. (2019), the study of early human brain development remains a challenge.

Beyond the interactions between the basal ganglia and the cortex, other learning-related brain circuits function with a 'two-stage' learning process where the initial learning stages rely on subcortical structures while long-term memory is imprinted in neocortical areas through on-line and/or off-line training of cortical networks by the subcortical regions. Concerning motor learning, BG and cerebellar circuits may teach neocortical motor areas where the motor programs are engrained for long-term storage of motor tasks (Caligiore, Pezzulo, et al. 2016). Here, we have shown how BG motor output matures earlier than cortico-cortical connections in songbirds, and how this may be beneficial for song learning. Similarly, the cerebello-neocortical circuit driving motor cerebellar-dependent learning may mature before the cortico-cortical connection in motor areas of the motor cortex are mature enough to store the motor memories. Even though this is purely speculative, this hypothesis could be tested experimentally by comparing the contribution of cerebellar and neocortical circuits to motor learning across development. Whether delaying the maturation of cortico-cortical connections benefits learning could be tested further in theoretical models of the BG-cortical and cerebello-cortical networks (Doya 2000).

For declarative memory, the experimentally confirmed 'two-stage' theory (Bontempi et al. 1999; Marr 1971; Maviel 2004) posits that new memories are transiently encoded in the hippocampus before they are gradually transferred in the prefrontal cortex for long-

term retention. The hippocampus has appeared before the neocortex in evolution, and it is tempting to speculate that ontogeny partially recapitulates evolution with an early maturation of hippocampal circuits involved in declarative memory before the maturation of the prefrontal cortex, which only terminates in adolescence or even later (Teffer et al. 2012). Such two-stage motor learning and episodic memory formation continues in adulthood, albeit with diminished influence of the cerebullum and hippocampus, respectively. Whether such two-stage developmental process could be beneficial for the development of memory circuits, during development, remains to be investigated both experimentally and theoretically.

4 Objective

In chapter 2 and 3, we have established that songbirds present a tractable animal model to study sensorimotor learning. We discussed evidence of BG-led sensorimotor learning in songbirds, and its subsequent consolidation within cortical motor pathways, suggesting a two-step paradigm of learning. The avian BG-thalamo-cortical pathway is necessary for learning vocalisations and is uniquely situated to implement reinforcement learning, which enables it to act as a tutor. In parallel, the cortical motor pathway is critical for producing these learnt vocalisations. In this manuscript, we exploit the vocal learning behaviour in songbirds, specifically zebra finches, and the underlying song system, in order to understand how neural circuits govern sensorimotor learning, from the perspective of both structural and functional plasticity.

Does structural plasticity have a functional role?

As discussed in chapter 2, the interactions between functional and structural plasticity are still unclear and often overlooked in computational models. In chapter 3, we have highlighted the presence of structural plasticity within the dual pathway system underlying vocal learning in songbirds. Several studies have provided sufficient empirical evidence showing the delayed maturation of the cortical pathway, with respect to the subcortical circuit through the BG in songbirds. The BG-thalamo-cortical loop (AFP) is fully developed before the sensorimotor phase of vocal learning, while in the cortical pathway, the HVC axons do not innervate the RA until two weeks later (Mooney and Rao 1994a). In this study, we are interested in exploring the functional role of structural plasticity in sensorimotor learning. Could structural plasticity potentially facilitate the transfer of sensorimotor learning from subcortical to cortical structures when constraints imposed by behaviour and anatomical development are taken into account? More specifically, is this delayed maturation of the cortical pathway, with respect to the subcortical circuit, merely an artefact of development? Or does the late innervation of RA by the HVC axons have a functional role in song learning? Could this lag be advantageous for sensorimotor exploration led by the BG during the initial phase of vocal learning (before 35dph)?

In chapter 5, we investigate the potential functional role played by structural plasticity towards sensorimotor learning. We hypothesize that the delayed growth of the cortical pathway, within the song system, confers functional advantages to vocal learning. In order to investigate this hypothesis, we design a simplified model of the song system that conducts sensorimotor exploration using a non-linear multi-segmented pivoted arm, as

an analogy to vocal learning. Within this scenario, we simulate both the simultaneous development of the parallel cortical and BG-thalamo-cortical pathways, as well as the delayed maturation of the cortical motor pathway and use this simple model to demonstrate the functional role of structural plasticity.

How can sensorimotor learning be optimised under complex landscapes in a biologically plausible manner?

The highly non-linear nature of the relationship amongst neural activity, the musculature in the syrinx and the resulting vocalisations creates an extremely complex and uneven sensorimotor landscape. In chapter 5, we look into biologically realistic strategies adopted by the songbird to navigate such a landscape. In chapter 3, in addition to anatomical substrates, such as the parallel cortical and subcortical pathways governing vocal learning and production, we have elucidated various behavioural features or constraints of vocal learning, such as the duration of sensorimotor learning period, the number of trial motifs per day, etc. Moreover, a daily post-sleep deterioration has been observed in song quality during early sensorimotor learning, which has been associated with an eventual superior imitation of tutor song during adulthood (Derégnaucourt et al. 2005). A similar non-monotonic trajectory in progression of sound acoustics has been observed during song learning, where juvenile birds often engage in pitch doubling (Tchernichovski et al. 2001). Thus, there seem to be multiple indicators of an apparent non-monotonic trajectory of improvement in song quality during vocal learning.

We look into possibly strategies employed by the songbird to optimise vocal learning in complex, uneven landscapes, by taking into consideration empirical evidences regarding the behaviour of vocal learning and the underlying anatomy supporting it. Given the dual pathway architecture of the song system, do the parallel pathways aid sensorimotor learning by executing different roles? What is the effect of structural plasticity within such a dual pathway framework? Furthermore, is there a potential neural mechanism that could implement the aforementioned post-sleep deterioration? Could the mechanisms underlying such a saltatory trajectory facilitate navigation in a complex sensorimotor landscape? How do such learning mechanisms compare with standard machine learning algorithms that optimise gradient-descent based reinforcement learning?

In section 5.2 and 5.3, we investigate the interplay between the cortical and subcortical pathways during vocal learning. We look into the different roles played by the two pathways and the potential functional benefits conferred by the delayed maturation of the cortical pathway to vocal learning. Further, we interpret the post-sleep deterioration observed during early sensorimotor learning as a facilitatory factor towards vocal learning. In order to investigate this hypothesis, we design a conceptual model of the song system that conducts sensorimotor exploration on a performance landscape generated using a model of the avian syrinx, as an analogy to vocal learning. We ratify the conceptual model using a network model with sigmoidal rate-coded neurons. Finally, we draw

parallels between the functioning of the model and optimisation techniques of gradient descent approaches used in machine learning.

Neural correlates of reinforcement learning and transfer of sensorimotor learning from subcortical to cortical circuitry in songbirds

It has been widely hypothesised that reinforcement learning, driven by the BG-thalamo-cortical loop, is a primary substrate for sensorimotor learning (M. Fee et al. 2011). In chapter 3, we inspect the song system and study the necessary substrates for the neural implementation of reinforcement learning within the dual pathway architecture. There are indeed evidences demonstrating, one, the ability of the output nucleus of the AFP, the LMAN, to inject exploratory variability into the system, two, mid-brain dopaminergic projections providing feedback about performance evaluation to the area X, and, three, activity-dependent plasticity within both the BG and cortical pathways of the song system. Thus, the avian BG-thalamo-cortical pathway displays most of the ingredients necessary to implement reinforcement learning.

We scrutinise the plausibility of such BG-led exploration influencing vocal behaviour and being consolidated within the cortical motor pathway in songbirds. Now, if the BG does indeed implement reinforcement learning, the output nucleus of the AFP, i.e. the LMAN, needs to be able to drive variations in behaviour. It has been shown that variations in premotor RA activity are correlated with variations in acoustic features, such as, pitch. The LMAN could potentially use its direct excitatory projections to the RA to influence behaviour. In this case, during vocal learning, variations in LMAN activity would also show a correlation with variations in acoustic features. Moreover, for the tutor signals from the AFP to be consolidated within the motor pathway as well, the LMAN needs to be able to exert its influence on the cortical pathway. If the LMAN does indeed drive variability in the RA, in order to exact changes in behaviour, we hypothesise that during sensorimotor learning, there would be an increase in correlation between, one, LMAN activity and resulting behaviour, as well as, two, LMAN and RA activity, presumably within the premotor window.

In chapter 6, we look further into this hypothesis regarding the feasibility of reinforcement learning in the song system. First, we build a simplified model of the song system and induce adult plasticity, by emulating a widely used protocol of conditioned auditory feedback (CAF), in order to make qualitative predictions concerning our hypothesis. We then proceed to test the prediction which emerges from the model, by designing an electrophysiology study to observe the activity in the neural correlates of vocal behaviour, in freely moving zebra finches, exhibiting song learning. We induce adult plasticity in male zebra finches, using the aforementioned conditioned auditory feedback protocol. During this protocol, while the bird learns to alter its song, we perform electrophysiological recordings in the RA, the locus of motor control, and LMAN, the output nucleus of

the AFP, and analyse whether the activity of LMAN is correlated with behaviour, and whether LMAN drives RA activity during sensorimotor learning.

Thus, we investigate BG-driven sensorimotor learning within the avian song system, from both a theoretical and experimental perspective in order to assess our hypothesis concerning the contribution of structural plasticity, biologically plausible optimisations of learning mechanisms and transfer of reinforcement learning within a dual pathway architecture. More specifically, we begin by using theoretical models to understand the interaction between the parallel pathways in the song system and the functional role played by structural plasticity, in chapter 5. We posit that within the dual pathway architecture, the delayed maturation of the cortical pathway directly facilitates sensorimotor learning. We then proceed to incorporate behavioural features of vocal learning and look into biologically plausible ways of optimisation in pathological situations for gradient-descent based learning. Finally, in chapter 6, we look into the feasibility of reinforcement learning in a two-step paradigm, within the song system, using both theoretical and experimental methods. To investigate the predictions made by the computational model, we design an experimental protocol and collect preliminary electrophysiological data from the neural correlates of vocal learning in zebra finches.

5 Computational benefits of the dual pathway framework

The vocal learning behaviour exhibited by songbirds makes them an excellent model to study sensorimotor learning. As discussed in chapter 3, song acquisition and production in songbirds is governed by a dedicated neural circuitry that involves two parallel pathways: a cortical pathway required for production and a basal ganglia-thalamo-cortical (BG) pathway necessary for plasticity. The BG pathway induces variability in production during vocal exploration, receives a performance signal via midbrain dopaminergic projections and drives a motor bias that corrects vocal errors. This dopamine-modulated change in vocal output, induced by the BG is gradually consolidated within the cortical pathway.

To understand BG-driven learning, several models based on Hebbian learnng have been proposed (Hanuschkin et al. 2013; Pagliarini et al. 2018). These models hypothesize that early sensorimotor exploration could enable the bird to form an internal inverse representation of the possible vocalisations and corresponding neural motor codes. This inverse model can be later used to produce the appropriate motor command or vocalisation at a given point of time within an independently learnt sequence structure. Alternatively, Yildiz et al. (2011) generate vocalisations with a three-level heirarchical model, instead of an inverse model, where neural ensembles in the HVC are sequentially activated, and drive the activity in RA from one attractor ensemble to the next, which in turn generates vocalisations using a model of vocal tract dynamics (Laje et al. 2002). Troyer et al. (2000) implement Hebbian learning at HVC-RA synapses to build a forward model and map neural codes of motor action to its auditory consequences. They, then, use the AFP as a critic to evaluate this mapping. Such an inverse or forward model could provide the foundation for dopamine-dependent reinforcement learning (RL) within the BG circuitry to eventually learn an accurate imitation of the tutor song.

Reinforcement learning (RL) has been widely hypothesised to govern sensorimotor learning (Wickens et al. 2003). Schultz et al. (1997) show that dopaminergic signals from the ventral tegmental area and substantia nigra report ongoing prediction errors for reward, which is reminiscent of the error signals propagated over time in temporal difference learning algorithms, a class of RL algorithms. Likewise, in songbirds, dopaminergic neurons have been shown to encode performance error during vocal learning, with a suppression of dopaminergic activity following worse than expected performance and

vice versa (Gadagkar and Goldberg 2016). Unlike the reward prediction error seen in the former study, which predicts value based on external factors, performance errors are an intrinsic evaluation of perceived motor output with respect to an internal objective. The facilitators of reinforcement learning within the BG-thalamo-cortical circuitry in songbirds have been discussed further in section 3.2.2 and 5.2.1. In the context of BG-driven vocal learning in songbirds, Doya and Sejnowski (1998) propose that the LMAN acts as a variability generator, which perturbs the HVC-RA connections, leading to exploration of the bird's range of vocalisations. Fiete, M. S. Fee, et al. (2007) expand on this idea in a more biologically realistic manner, where instead of perturbing synaptic connections, the cortical pathway is influenced via perturbing RA neurons or 'nodes'. However, the purely gradient descent based approach adopted here is susceptible to getting trapped in local optima when traversing uneven sensorimotor landscapes (Bottou et al. 1991; Gori et al. 1992). An important proposed hypothesis for BG-led exploration is that the AFP is involved in gradually biasing the motor system towards the desired target (M. Fee et al. 2011). Finally, Farries and Fairhall (2007) take a step towards verifying the biological plausibility for the reinforcement learning theory by using spike-time dependent plasticity rules to train a neural network such that given a patterned synaptic input, neurons generate firing patterns that lead to neural ensembles developing the appropriate population response.

Thus, reinforcement learning can enable the BG pathway to guide the cortical pathway towards producing desired vocalisations. However, the non-linear and redundant nature of the relationship between neural control, syrinx musculature and resulting acoustics, makes song learning a complicated problem to solve (Srivastava et al. 2015). Such complex and uneven sensorimotor landscapes present a challenge to reinforcement learning implemented using direct gradient descent. Several pathological landscapes, such as the Rastrigin function, divulge the shortcomings of gradient descent when several local optima are present ("Global Optimization" 1999). Now, we turn to the vocal learning behaviour and underlying circuitry to explore the strategies adopted by songbirds to navigate such sensorimotor landscapes. As discussed in chapter 3, the song system, with its parallel subcortical and cortical pathways, gradually transfers BG-led reinforcement learning to cortical motor networks, within a two-step paradigm. It has been observed that within the cortical pathway, the HVC axons wait at the dorsal border of RA for two weeks before innervating the RA (Mooney and Rao 1994b). Is this delayed maturation of the cortical pathway merely an artefact of development? Or could it be advantageous for sensorimotor exploration led by the BG during the initial phase of vocal learning? Moreover, a daily post-sleep deterioration has been observed in song quality during early sensorimotor learning, which has been associated with higher quality of vocal imitations in the longer timescale (Derégnaucourt et al. 2005). Could such a saltatory trajectory aid reinforcement learning in navigating a complex sensorimotor landscape? What could be the underlying mechanism for the post-sleep deterioration during early sensorimotor learning?

In this chapter, we examine the aforementioned characteristics of the vocal learning behaviour and circuitry in songbirds, and theoretically investigate the potential benefits

conferred by them to vocal learning. We, first, look into a possible functional role of structural plasticity towards sensorimotor learning. We posit that structural plasticity may play a crucial role in vocal learning by regulating the transfer of information from sub cortical to cortical networks and moderating their relative influence towards motor output. Having established a possible role for structural plasticity towards learning, we proceed to look at the dual pathway framework with a more holistic approach. We begin with a conceptual model which investigates the transfer of learning within a dual pathway framework while incorporating empirical evidences from both behaviour and anatomy. Using this conceptual model, we look into 1. the interplay between the two parallel pathways involved in vocal learning using both functional and structural plasticity, 2. strategies to navigate a complex sensorimotor landscape, inspired from the non-monotonic trajectory towards the vocal imitation and 3. the analogies between functioning of the biologically feasible dual pathway framework and more computationally intensive machine learning algorithms. We proceed to validate this conceptual model further using a scaled biologically realistic model with rate coded neurons.

5.1 Functional advantage of structural plasticity

Based on the empirical observations we have reviewed in the chapter 3, we introduce in this section a simplified model of vocal learning in songbirds (see Fig 5.1). While its purpose is not to precisely model the circuitry underlying the vocal learning process, this model attempts to demonstrate the benefits of adopting features of structural plasticity within computational frameworks. Furthermore, this simplified model will help us to illustrate the putative role of structural plasticity in the early phase of learning. More specifically, we aim at showing that structural plasticity plays a functional role and the absence of such early plasticity is indeed detrimental to learning. This will constitute the ground for a discussion about the crucial role of the structural plasticity in sensorimotor learning as well as the rather scarce use of structural plasticity in computational models.

5.1.1 Hypothesis

Now, we hypothesise a role for structural plasticity in directly aiding the process of song learning within a dual pathway framework, by incorporating evidences from anatomical studies of songbirds, specifically zebra finches. As we have seen in earlier sections, in the case of songbirds, the vocal learning circuitry comprises of two parallel pathways. The primary cortical motor pathway (involving the HVC and RA) is believed to be responsible for controlling song generation, while a secondary BG-thalamo-cortical pathway (involving the LMAN, Area X and DLM) functions as a tutor and sends signals to influence the course of vocal production. Area X receives strong dopaminergic innerva-

tion from the midbrain, which provides an online evaluation of song quality (Lewis et al. 1997). Moreover, there is experimental evidence showing activity-dependent synaptic plasticity at HVC and LMAN projections to spiny neurons in area X (Ding et al. 2004a). Thus, along with LMAN-driven variability, Area X forms an ideal site to facilitate reinforcement learning. On the other hand, activity dependent synaptic potentiation and depression have also been described at the HVC-RA synapses(Mehaffey et al. 2015). This provides an avenue for synaptic plasticity based on Hebbian learning, or co-activation of pre-synaptic and post-synaptic terminals, within the cortical motor pathway. There is a range of empirical evidence to support the hypothesis that the tutor signals from the BG-thalamo-cortical pathway guide song learning and are eventually consolidated within the cortical pathway(Ölveczky, Andalman, et al. 2005; Ölveczky, Otchy, et al. 2011). Based on these evidences, we model the vocal learning circuitry using a dual pathway framework.

Additionally, morphological studies have shown that, during the development process, the structural formation of the BG-thalamo-cortical (or tutor pathway) is completed prior to the formation of the primary cortical motor pathway. More precisely, it has been shown that by day post-hatch (dph) 20, i.e. at the onset of singing, the axons from LMAN enter the RA, thus completing the tutor pathway. However, the axons from HVC (part of the primary pathway) reach the dorsal border of RA by dph 15, but do not innervate the RA before day 30 (Herrmann and Arnold 1991; Mooney and Rao 1994b). Thus, the HVC axons wait at the dorsal border of RA and enter the RA to form synapses only after the formation of the BG-thalamo-cortical loop.

This leads us to investigate if this delayed circuit completion, induced by structural plasticity, might, in fact, have a crucial functional role in the process of vocal learning using the dual pathway architecture. We hypothesise that this manifestation of structural plasticity, in the form of delayed axonal connections, has a functional purpose and confers a functional advantage in the song learning process.

5.1.2 Virtual arm exploration as an analogy to song learning

To verify our hypothesis, we investigate this phenomenon, using an analogous paradigm of virtual arm exploration, albeit in a simplified manner[1]. Given a multi-segmented arm, pivoted at a point, we model the process of the arm learning to reach a specified target. We assume that sensory learning has conferred a sensory representation of the tutor song to the juvenile and is modeled here directly as a position on the sensorimotor space. The RA plays a crucial role in controlling the bird's syrinx and respiratory muscles, which are essential to the production of song. Imitating tutor song requires learning this complex non-linear transformation of the activity in the RA to the functioning of the vocal

[1]The scripts are available on https://github.com/rsankar9/Review-Arm-Exploration-Model and archived in Zenodo (https://doi.org/10.5281/zenodo.4063714) (Sankar, Rougier, et al. 2020).

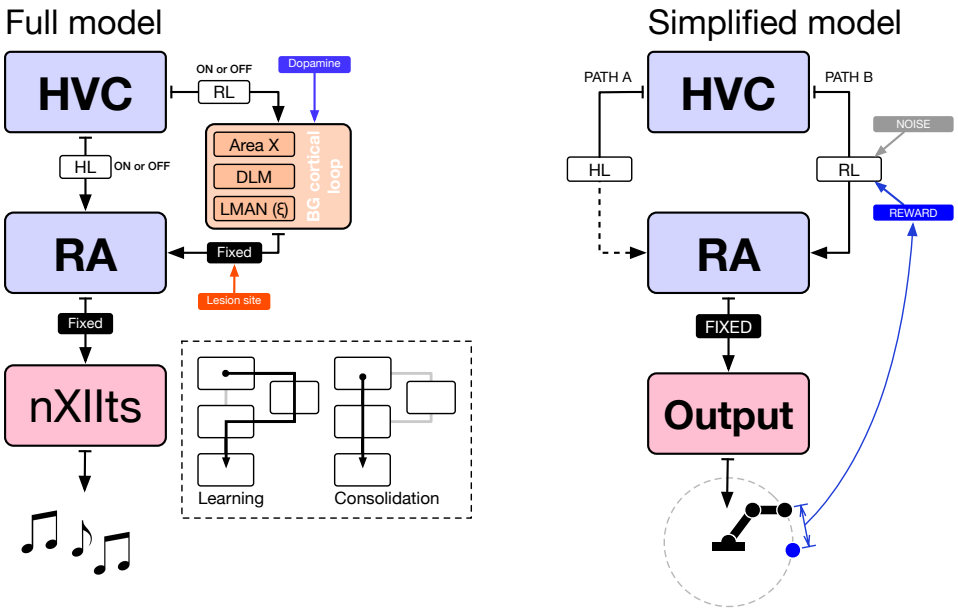


Figure 5.1: Full and simplified model architecture illustrating our hypothesis regarding learning and consolidation. The anterior forebrain pathway, via the avian BG, provides a tutor signal which is eventually consolidated within the cortical pathway to generate the desired behaviour. The simplified model illustrates a minimal functioning of the parallel pathways, built using Hebbian learning and reinforcement learning, without detailed description of the neural components.

musculature of the bird. Sensorimotor exploration via the manipulation of the angles between the arm segments serves as a minimal task to represent this process.

5.1.3 Methods

Architecture

The model's network comprises of three layers of rate-coded neurons: the HVC, the RA and the output layer, as shown in Fig 5.1. The HVC is used to encode the desired target. It contains multiple binary rate-coded neurons, and the configuration of these neurons can be used to denote different targets. The RA represents the locus of motor control, and consists of several highly sigmoidal neurons (Eq 5.2). The output layer comprises as many neurons as the number of arm segments(n_{seg}) (Eq 5.3). The angles between the arm segments is governed by the output of the model.

The model is mainly governed by two pathways connecting the HVC and the RA. Path A, analogous to the cortical motor pathway, connects the HVC and RA. In parallel, Path B minimally represents the BG-thalamo-cortical tutor pathway, without explicitly detailing the BG-thalamic components. Path A (W_A) and path B (W_B) are fully connected via two pathways (Eq 5.1). In path A, the connections are initialised to zero, and in path B, the connections are initialised to random values, between $W_{B_{min}}$ and $W_{B_{max}}$. The input to RA is a summation of the inputs from both Path A and Path B, normalised by the size of the HVC. The RA and the output layer are fully connected with fixed synaptic weights (W_C). The activation function of RA is a sigmoidal function with a steep slope (ss). The output of the model is thus a non-linear function of the total output of two parallel paths. The motor output provides the angle (in radians, θ) between each arm segment w.r.t. the horizontal plane, which is thereafter converted into Cartesian coordinates (Eq 5.9).

$$J_{RA} = sig(\frac{J_{HVC} \cdot W_A + J_{HVC} \cdot W_B}{n_{HVC}}) \tag{5.1}$$

$$sig(x) = \frac{1}{1 + e^{-1*(x-sm)*ss}} \tag{5.2}$$

$$output = \frac{J_{RA} \cdot W_C}{n_{RA}} - \pi \tag{5.3}$$

where n_{HVC} is the number of neurons in HVC, n_{RA} is the number of neurons in RA, J_X is the activity of layer X, ss is the slope of the sigmoid, and sm is the mid-point of the sigmoid.

Learning

Path A, analogous to the cortical motor pathway, is built on Hebbian learning between the HVC and RA. Path A updates the weights of its synaptic connections via potentiation when both presynaptic and postsynaptic neurons are active, and depotentiation when either of them are inactive. Path B minimally represents the BG-thalamo cortical tutor pathway by using reinforcement learning between the HVC and RA, without explicitly detailing the BG-thalamic components. Path B induces randomised noise in the network in order to explore and evaluates its sensorimotor exploration using feedback from the motor output. This scalar feedback is indirectly proportional to the distance of the position of the arm's end-effector from the target position.

Path B: Reinforcement Learning

The weights of Path B are updated using reinforcement learning, specifically using the covariance rule. Path B conducts exploration for the desired goal, by introducing randomised noise directly into the weights of the Path B. The noise (ξ) injected into Path B perturbates the activity of RA, which in turn affects the output of the model. The error is calculated as the normalised Euclidean distance between the position of the end effector of the arm and the desired target (Eq 5.4). A scalar reward signal, indirectly proportional to the error, is then sent as feedback to path B (Eq 5.5). Path B uses this information to reinforce its weights if the injected noise improves the reward received and vice versa (Eq 5.7). The weights are soft-bounded to $W_{B_{min}}$ and $W_{B_{max}}$.

$$E = \frac{\sqrt{(x - x^*)^2 + (y - y^*)^2}}{2 * \Sigma_{k=1}^{n_{seg}} l_k} \tag{5.4}$$

$$R = e^{-E^2/\sigma_R^2} \tag{5.5}$$

$$R_{rec} = \frac{\Sigma_{t=nt-25}^{nt} R[t]}{25} \tag{5.6}$$

$$\Delta W_B = \eta_B * \xi * (R - R_{rec}) * J_{HVC} * J_{RA} \tag{5.7}$$

where η_B is the learning rate of path B, x, y denote the output coordinates, l is the length of each segment and nt is the trial number.

Path A: Hebbian Learning

The weights of Path B are updated as per Hebbian learning. If the presynaptic neuron in the Hvc and the postsynaptic neuron in the RA are both active, the corresponding synaptic weight is potentiated, and if either one is inactive, the corresponding synaptic weight is depotentiated (Eq 5.8). The weights are soft-bounded to $W_{A_{min}}$ and $W_{A_{max}}$.

Table 5.1: Parameter values used for simulations in section 5.1.

Parameter	Symbol	Value
No. of arm segments	n_{seg}	3
Path B learning rate	η_B	0.15
Path A potentiation rates	η_p	0.0015
Path A depotentiation rate	η_d	0.0015
Reward sensitivity	σ_R	0.35
Noise limit	ξ_{max}	0.5
No. of training trials	NT	6000
Path A, B weight bounds	$W_{A_{min}}, W_{A_{max}}$	-1, 1
Path B weight bounds	$W_{B_{min}}, W_{B_{max}}$	-1, 1
Path C weight bounds	$W_{C_{min}}, W_{C_{max}}$	0, 25

$$\begin{aligned}\Delta W_A = &\eta_p * J_{HVC} * J_{RA} \\ &- \eta_d * J_{HVC} * (1 - J_{RA}) \\ &- \eta_d * (1 - J_{HVC}) * J_{RA}\end{aligned} \tag{5.8}$$

where η_p and η_d are the rates of potentiation and depotentiation.

Task

The objective of the model is to learn the right combination of angles between the segments of a pivoted arm in order to reach a specified target. The model receives, as input, one, the length of the arm segments (l), two, the configuration of the network, and three, the target position (x^*, y^*), i.e. the Cartesian coordinates of the desired target (Eq 5.9). The length of the arm segments are selected to be equal and to sum up to 1. The configuration of network includes the information about the fixed weights between the RA and the output layer, as well as the randomly initialised weights between the HVC and RA via path A and path B. Further, it is verified that the network configuration has the capacity to produce outputs (representing the angles between the arm segments) in a range larger than 0 to 2π. The target position is chosen such that it lies within the range of the pivoted arm, i.e. a circle of diameter twice the arm length.

$$x, y = \Sigma_{k=1}^{n_{seg}} l_k.cos\theta_k, \Sigma_{k=1}^{n_{seg}} l_k.sin\theta_k \tag{5.9}$$

where n_{seg} is the no. of arm segments.

Simulation

Each simulation consists of a training phase and a testing phase. The training phase (nt trials) is ten times longer than the testing phase. During the training phase, path B is active, while during the testing phase, it is deactivated. This represents lesioning of the LMAN outputs, which has been shown to reduce variability in song production, as well as arrest song learning in juveniles. Thus, testing phase consists of $nt/10$ trials where the model is controlled only by path A. For condition 2, the training consists of nt trials where both Path A and Path B are active. For condition 1, Path B is active throughout the training phase while path A is deactivated for the initial third of the training phase, i.e. the training consists of $nt/3$ trials in the beginning where only Path B is active. This particular delay was chosen as, in 90% of simulations, path B takes longer than one-third of the training phase to converge (Path B convergence: M=3840, SD=1136), when path A is inactive throughout. The convergence point of a simulation, as shown in Fig 5.2, was determined as the trial at which 1. the current average error is within 1% of the minimum average error of the simulation, and 2. the standard deviation of the average error of the recent 500 trials is less than 0.03%. Here, the average error is the running average over 100 trials.

Hypothesis test

We test the learning process of this multi-segmented pivoted arm, under two conditions:

1. This condition mimics the delayed innervation of RA by HVC axons. Path B is present and active from the beginning of the training phase, while path A is not. Path A is formed and activated after a stipulated delay, in this case, the initial third of the training phase. Path A starts to learn after a delay with respect to when path B has started its exploratory learning.

2. This condition denotes simultaneous progression of both pathways. Paths A and B are present and active from the beginning of the simulation, and learn simultaneously. Path A is active throughout the training phase in condition 2.

Path B is active throughout the training phase in both conditions.

Analysis

The average convergence of path B, without any influence of path A, is calculated in order to set the delay. For analysis, each condition is simulated with two hundred different random seeds using the parameter set, shown in Table 5.1. Each seed corresponds to a specific network configuration, initial position of the arm, and the target destination. The mean normalised error during the testing phase of each simulation is used as a metric for further statistical comparison of the distributions. The proportion of seeds for which

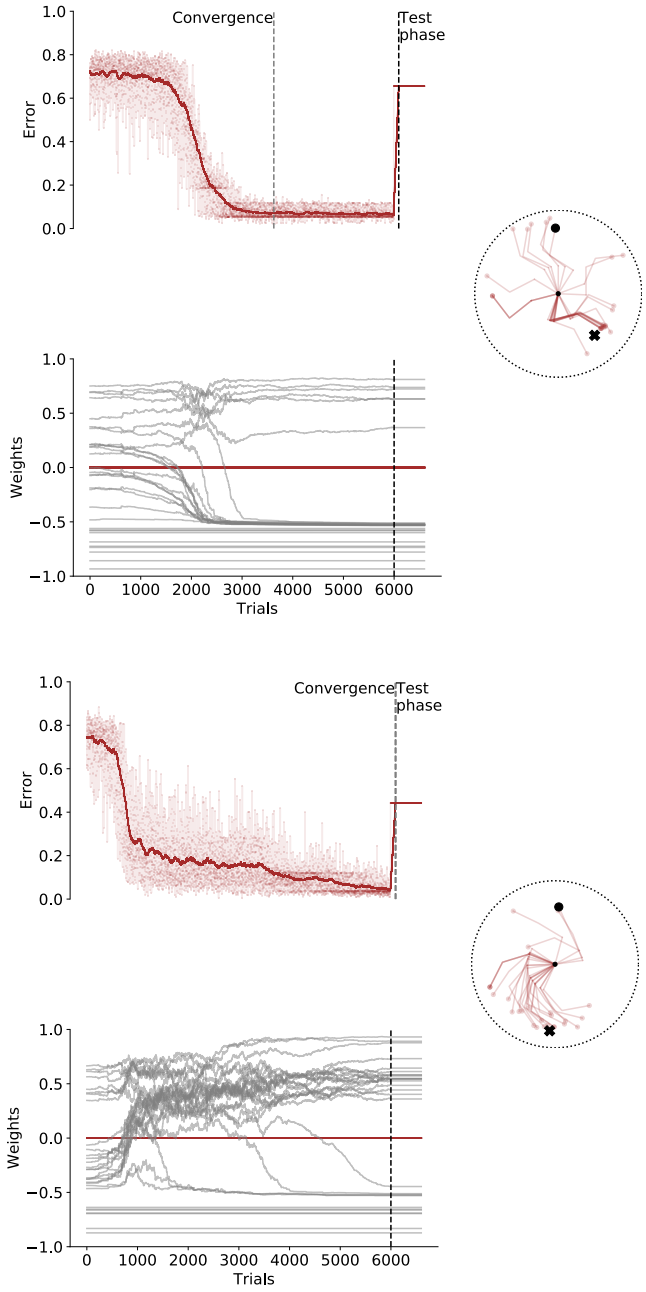


Figure 5.2: Two sample simulations demonstrating the calculation of convergence. To calculate the convergence point, we inactivate path A and observe the number of trials path B takes to reach convergence. The point of convergence is calculated as explained in section 5.1.3, and annotated here with a vertical dotted line. Each subfigure has three panels. The top panel shows the evolution of the error function (red) over the simulation. The bottom panel shows the evolution of the strength of the Path A (red) and Path B (grey) connections. The circular figure, on the right panel, provides a visualisation of the movement of the arm over training, form the initial position (dot) to the target (x). The final configuration of the arm is shown with higher opacity.

the simulation had a better performance in condition 1 than condition 2 is calculated. A paired t-test between this metric for the simulations in condition 1 vs those in condition 2, with corresponding random seeds, is conducted. Additionally, the initial distance between the arm end-effector and the target position is plotted against the final distance, and it is observed that larger the initial distance (i.e. the distance the arm needs to traverse to reach the target), the lower the success in reaching the target in condition 2.

5.1.4 Results

To understand the workings of the model under the two conditions, we look at one example simulation in each condition, as shown in the schema of the simplified model in Fig 5.3. Fig 5.3A illustrates condition 1, i.e. the progression when there is a time lag between the start of the two pathways, while Fig 5.3B illustrates condition 2, the progression of the simulation when both pathways have a simultaneous start. As shown in the bottom panel of Fig 5.3A, when Path A is paused for the initial third of the training process, Path B, using reinforcement learning, begins exploration and is able to find the right direction to move towards the target (top panel). After one third of the training phase, when path A is activated, Hebbian learning mimics the direction chosen by reinforcement learning, by further strengthening the pattern of neuronal activation in RA, selected by path B (bottom panel). This leads to the arm eventually reaching the target (right panel). On the other hand, in condition 2, when there is a simultaneous start of both path A and path B (Fig 5.3B), we observe that the arm is unable to reach the target (right panel). The bottom panel shows us that, before the tutor pathway finds the correct direction (i.e. the correct pattern of neural activation in the RA), path A prematurely strengthens connections to the incorrect set of RA neurons. The strong input to the RA, from path A, leads to path B being unable to sufficiently modulate the incorrect RA activity. This, in turn, obstructs the depotentiation of the required synaptic connections in path A. Hence, the network causes the arm to moves towards an arbitrary position, and ultimately, not reach the target.

During the test phase, the influence of path B is completely removed in both conditions. The delay in condition 1 is insufficient for path B to converge to a solution on its own (when only path B is activated, 90% cases require more than one third of the training phase to converge). Performance in a simulation is measured by the mean normalised error during the test phase. The lower the mean error, the better the performance. Over two hundred such simulations per condition, with varying initial and target positions, we observed[2] that condition 1 (0.12 ± 0.10) has a significantly better performance than condition 2 (0.27 ± 0.21) in reaching the desired target; $t(199)=11.76$, $p<0.001$. In condition 2, the desired target was not reached in a majority of simulations. These results

[2] Here, we have presented the mean $\pm$ standard deviation of the performance metric (mean normalised error in testing phase) for both conditions.

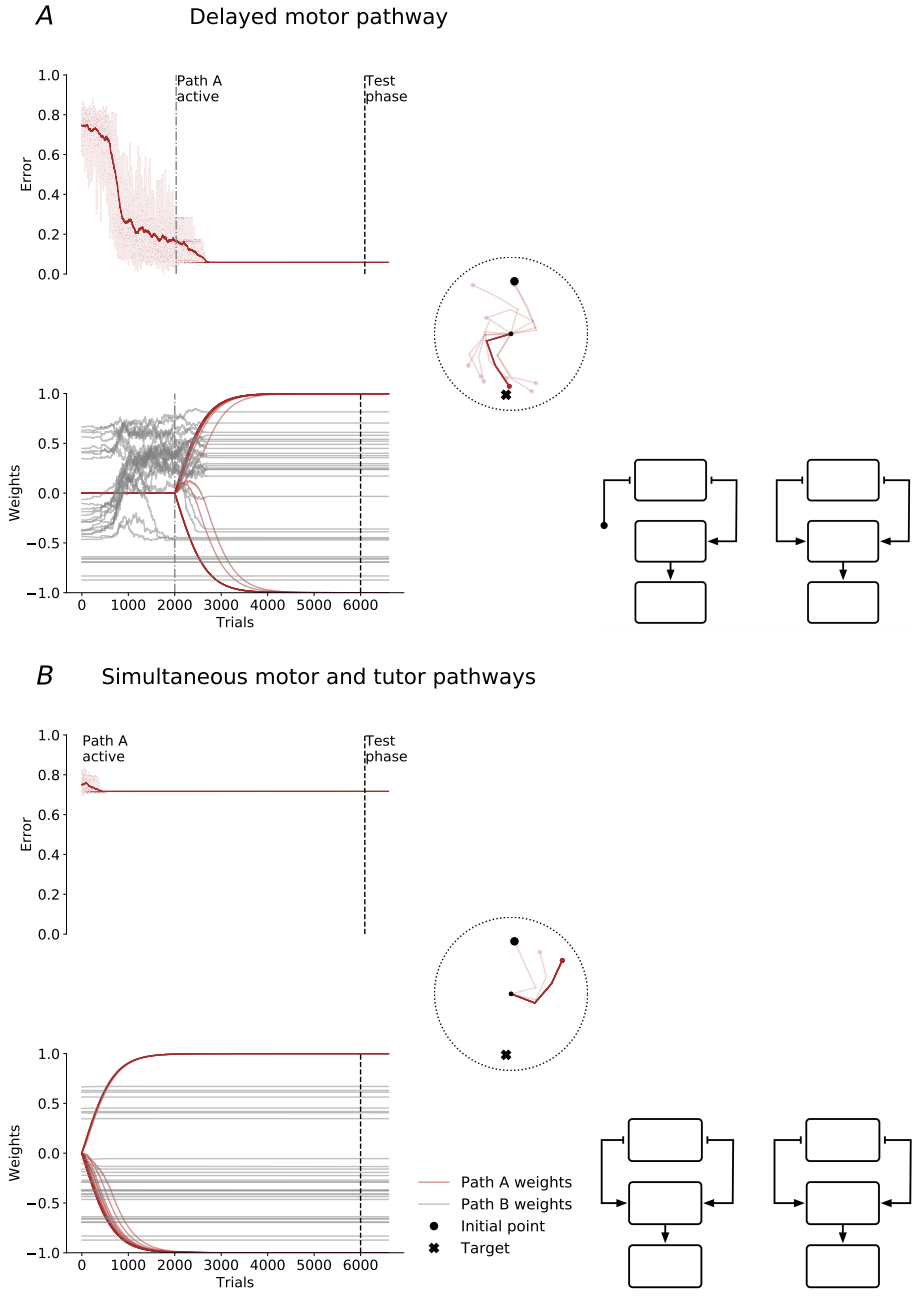


Figure 5.3: A sample simulation demonstrating the advantage in having a delay in the development of the cortical motor pathway. Subfigure A shows the arm successfully converging to the target in condition 1, unlike condition 2, shown in subfigure B. Each subfigure has three panels. The top panel shows the evolution of the error function (red) over the simulation. The bottom panel shows the evolution of the strength of the Path A (red) and Path B (grey) connections. The circular figure, on the right panel, provides a visualisation of the movement of the arm over training, form the initial position (dot) to the target (x). Only one arm configuration per 250 trials has been plotted for clarity. The final configuration of the arm is shown with higher opacity.

suggest that having a delay in the development of path A with respect to to the tutor path B confers a critical advantage in motor learning via sensorimotor exploration (see Figure 5.4).

5.1.5 Discussion

While several advantages of structural plasticity have been noted across the literature, ranging from energy efficiency to optimisation of information storage (as discussed in chapter 2), we demonstrate that structural plasticity, within a specific circuitry, contributes directly to the function of said circuitry. In order to look into the potential direct functional advantages of structural plasticity, we scrutinize a hypothesis that is gaining popularity in the literature. This hypothesis states that the development of automatized skills, such as song production, is driven by the basal ganglia through dopamine-modulated reinforcement learning in order to guide learning in a parallel cortical pathway, which eventually governs the production of the skill (Andalman et al. 2009a; Ashby et al. 2010; Pasupathy et al. 2005a).

The model presented in this section attempts to stress upon the, often overlooked, potential functional role conferred by structural plasticity, using the vocal learning circuitry of songbirds as a framework. This simplified model illustrates why and how the simultaneous development of tutor and motor pathways could be detrimental to learning. From Figure 5.3 A, we observe that when the tutor pathway is granted a period of unregulated exploration, the model is able to test different patterns of RA activation in an unencumbered manner and selectively potentiate the suitable synaptic connections within the cortical motor pathway. Thus, the model is rendered more likely to find the favourable solution, as compared to condition 2 (Figure 5.4). The vocal babbling period of sensorimotor learning in songbirds has been shown to be largely driven by BG. The delayed growth of the cortical pathway, being analogous to condition 2, might thus facilitate an indispensable period of sensorimotor exploration. The simultaneous development of the cortical motor pathway and the BG tutor pathway, limits the influence of the tutor on the motor output to effectively explore the sensorimotor space. Thus, in condition 2, the premature development of the cortical pathway solidifies an undesirable motor control pattern within the RA, rendering the system more likely to be saddled close to the initial configuration, as shown in Figure 5.4B. The tutor pathway being given the opportunity to independently explore, and identify the right direction to move the arm, results in an improved capacity to influence the motor pathway towards the desired target. Thus, structural plasticity, by enabling this differential development of the two pathways in zebra finches, serves a crucial functional purpose of assisting sensorimotor exploration and learning.

In the next section, we continue to investigate sensorimotor learning using a similar framework inspired by the vocal learning circuitry. However, moving one step closer to understanding the biological mechanisms underlying vocal learning, we take into con-

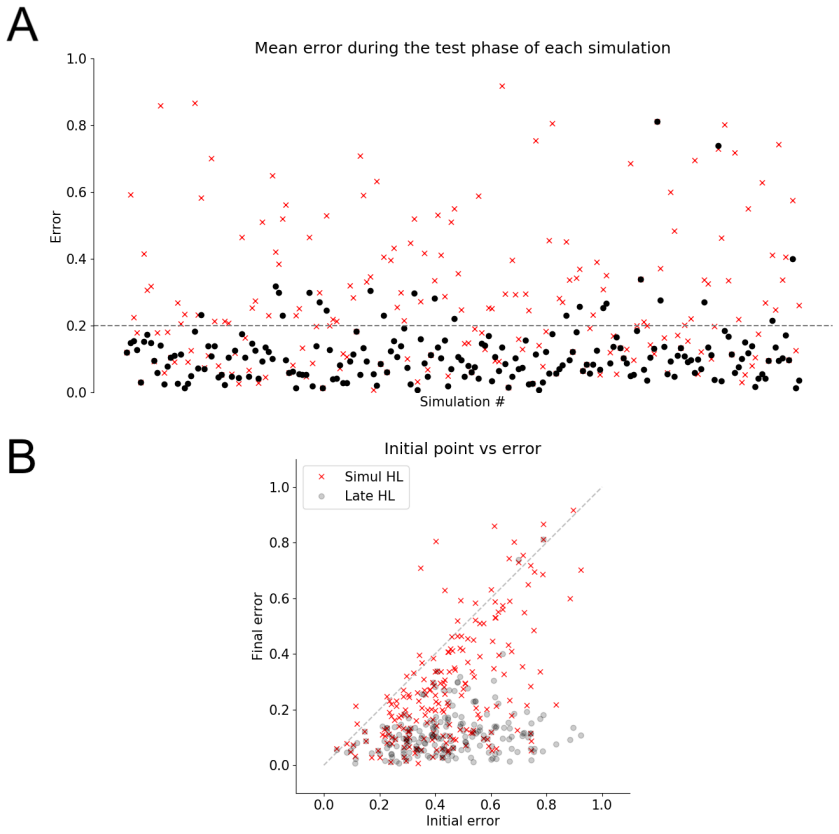


Figure 5.4: Summary of all simulations in condition 1 and condition 2. Panel A shows the mean error during the test phase of each simulation in condition 1 (black) vs condition 2 (red). Panel B shows the comparison between the error at the beginning of the simulation vs the end of the simulation. We observe that in condition 2 (red), the model does not always move away from the initial configuration. In condition 1 (grey), the model is able to land in locations with low error, irrespective of the starting position.

sideration the different behavioural features or constraints of vocal learning exhibited by songbirds, in addition to the architecture of the song system and the structural plasticity therein. Furthermore, in order to better incorporate the complexities of the syrinx, instead of the artificial task of sensorimotor learning by an agent with a multi-segmented robotic arm, we benefit from a biophysical model of the syrinx (Amador et al. 2013). Using this syrinx model, we build a sensorimotor landscape for the dual pathway model to navigate. This model of the songbird syrinx provides us with an improved approximation of the complexities and non-linear nature of the relationship between different neural control patterns and muscular configurations of the syrinx or resulting behaviour. Hence, having established the potential functional role of structural plasticity in this section, in the next section, we study the mechanisms governing sensorimotor learning under more realistic behavioural and physical constraints.

5.2 The dual pathway architecture

The acquisition of motor skills in vertebrates is widely believed to be governed by reinforcement learning (RL) wherein a skill is progressively learned through a series of trials and errors (Boraud et al. 2018b). Each trial in the motor space can be perceived in the sensory space to subsequently provide an evaluation of the performance, which will in turn generate a reward signal to the motor system and help rectify the action. In this view, learning to produce a given sensorimotor target involves finding the global optimum of the reward landscape across all possible motor commands. Reinforcement learning uses exploration to build an estimate of the local reward contour and implements gradient descent in order to maximize the expected reward (Sutton, Barto, and Williams 1992). Previous work has shown how neural circuitry may implement RL (direct policy search) through gradient ascent in the reward landscape to maximize the cumulative reward (Fiete, M. S. Fee, et al. 2007). However, standard RL approaches may result in non-optimal solutions in a continuous action space under uneven reward contours, such as the Himmelblau and the Rastrigin functions. There are several techniques that have been harnessed to optimise the gradient descent underlying RL, especially to evade local optima. Some of these techniques include Newton's method, proximal policy optimisation, momentum-accelerated gradient descent, etc (Vieillard et al. 2020). Other gradient-free algorithms (e.g. simulated annealing) are more efficient in complex uneven reward landscapes, but the mechanisms of their implementation in brain circuits remains speculative (Dhawale et al. 2017; Tsallis et al. 1996). However, these techniques can be computationally intensive and/or do not adhere to biological constraints. We find that the vocal learning behaviour exhibited by songbirds offers a less computationally intensive solution to reach the optimal solution under complex reward contours. Here, we look into the insights offered by the neural circuitry governing vocal learning in songbirds towards an alternate approach for sensorimotor learning.

In order to look into the mechanisms underlying sensorimotor learning, we consider the paradigm of vocal learning in songbirds. Based on their behavior, brain anatomy and physiology, we hypothesise that motor learning is governed by a dual pathway architecture that allows for an efficient learning, mixing reinforcement learning and the regulation of noise. The point we want to make here is not about optimality but rather plausibility. To do so, we adopt an alternative to classical approaches: instead of trying to justify a posteriori the existence of one or the other critical feature (e.g. back-propagation) inside the brain, we start from the raw biological facts and explore different hypotheses as to how the neural circuitry could implement efficient vocal learning.

5.2.1 The vocal learning behaviour and circuitry

Juvenile songbirds learn to imitate the vocalisations of an adult tutor through vocal learning, a form of sensorimotor learning akin to human speech learning. During development, the juvenile's vocalizations progress from highly variable (vocal babbling) to highly stereotyped and accurate imitations of the tutor song through a trial-and-error process indicating reinforcement learning. In zebra finches, song learning lasts around 60 days during which a juvenile produces thousands of vocalisations per day (Derégnaucourt et al. 2005). Vocalisations undergo changes at multiple timescales. While rapid changes are observed across vocalisations produced in a single day, some changes are consolidated on a weekly timescale (Derégnaucourt et al. 2005; Kollmorgen et al. 2020). Moreover, sleep induces a rapid discontinuity in the produced vocalisations with increased variability post-sleep (Derégnaucourt et al. 2005). Overnight changes, daily fluctuations and weekly consolidations are only partially aligned, making the song learning an erratic process (Kollmorgen et al. 2020).

Song acquisition and production is governed by a dedicated neural circuitry that involves two parallel pathways: a cortical (motor) pathway controlling vocal production in adults and a basal ganglia-thalamo-cortical (BG) pathway necessary for vocal learning and plasticity (Figure 5.5a). These pathways connect two main cortical nuclei: the premotor HVC generating song timing (Hahnloser, Kozhevnikov, et al. 2002; Long et al. 2008), and the RA, controlling the syringeal and tracheal musculature in order to produce vocalisations.

Direct axonal projections along the motor pathway develop during the early phase of song acquisition and exert a growing influence on song production during learning. On the contrary, the early-matured BG pathway drives initial vocalizations and variability in the subsequent song production, but has a reduced influence post learning (Mooney and Rao 1994a). The BG pathway receives a performance evaluation signal via strong dopaminergic projections from the midbrain and drives a motor bias that rectifies vocal errors (Bottjer and Johnson 1997b). This motor bias is then gradually consolidated within the cortical motor pathway.

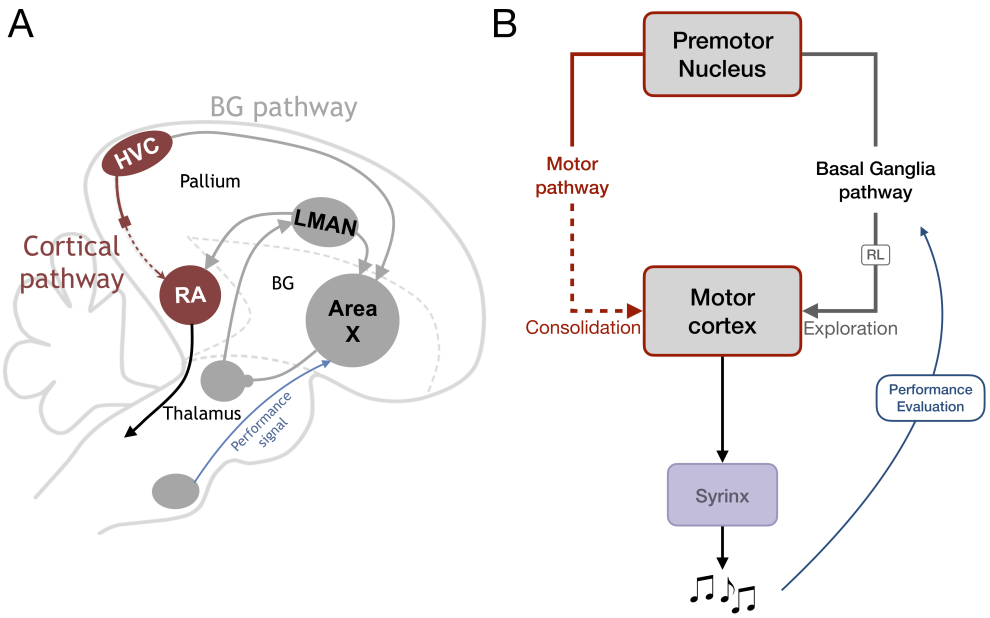


Figure 5.5: Song system in zebra finches and a simplified schema of the dual pathway architecture. **A** The specialised vocal learning circuitry comprises of two pathways: the cortical motor pathway (brown) and the BG-thalamo-cortical pathway (grey). The cortical pathway governs song production and includes the premotor cortical nucleus HVC and the RA. The RA projects to downstream regions which control respiratory musculature. The parallel BG pathway receives performance evaluation from mid-brain dopaminergic neurons and projects to the RA. **B**. The dual pathway architecture inspired by the vocal learning circuitry. The BG pathway (grey) is based on reinforcement learning (RL) and provides a tutor signal, which is consolidated gradually within the parallel cortical motor pathway (brown). The syrinx transforms the combined output of these two pathways into a syllable vocalisation.

Activity-dependent synaptic plasticity is believed to allow RL to be implemented in the BG pathway (HVC-BG synapses (Ding et al. 2004a)), and the consolidation of motor bias within the cortical motor pathway (HVC-RA synapses (Mehaffey et al. 2015)). Following song acquisition, the motor pathway is capable of producing the learnt song without the input from the BG pathway as lesions in the BG pathway do not affect song quality, but reduce the residual song variability.

During song learning, each timing signal in HVC must be associated to the proper muscle configuration to produce a given syllable, i.e. a vocalization with the desired acoustic features. This represents a complex problem to solve in a continuous action space, with a non-linear and redundant relationship between RA neural activity, muscle activation patterns and vocal acoustics (Srivastava et al. 2015).

Gradient descent can lead to sub-optimal solutions under uneven contours in continuous action spaces. Here, we propose to re-interpret the role of the dual pathway architecture underlying sensorimotor learning in birds, which offers a potential solution to the aforementioned limitations of direct gradient descent approaches. We explore the benefits of this architecture by simulating a simplified algorithmic implementation of the vocal learning process in the context of different reward landscapes, both randomly generated and biologically inspired. We demonstrate that the structural (delayed maturation of cortical pathway) and functional plasticity (activity-dependent synaptic plasticity (Ding et al. 2004a; Mehaffey et al. 2015)) observed in a two-pathway framework together help overcome certain shortcomings of standard RL approaches and facilitate a convergence to the global optimum. Further, we propose a novel data-driven algorithmic implementation, governing the functioning of the dual pathways, adhering to biological constraints. Meanwhile, we also draw parallels with traditional machine learning approaches, albeit this comparison does not aim at being exhaustive.

5.2.2 Methods

The goal of the model is to implement a vocal learning process consistent with the behavioural, physiological and anatomical evidence collected in songbirds and thereby gain insight into sensorimotor learning. We theoretically investigate the interplay of the two parallel pathways within the song system and their role in song acquisition. In this manuscript we choose to define an abstraction based on a simplified two dimensional (2D) representation of the dual pathway architecture, in order to perform a more systematic study of its properties in a reduced system.

Architecture

Taking inspiration from the vocal learning circuitry found in songbirds, the model [3] has been designed as a three layered architecture with two major parallel pathways (Figure 5.5b). The first layer (HVC) operates as an input layer which indicates the target syllable to be produced. The second layer (RA) generates a bounded 2D motor output. The third layer mimics the working of an avian syrinx and transforms the low-dimensional motor output from the RA layer into a syllable vocalisation. The HVC and RA layers are connected by two parallel circuits inspired by the song system, the cortical pathway (motor pathway) which drives motor output and the BG-thalamo-cortical circuit (RL pathway) which implements RL and provides a tutor signal to the motor pathway (Figure 5.5). The outputs of the motor and RL pathways are represented by two scalar values, μ_{mtr}, and μ_{rl}, respectively. These values are weighted by the influence of the two pathways, w_{mtr} and w_{rl} that reflect their respective contributions to RA output (Eq 5.10-5.11). RA output, P, is a summation of the contributions of the motor, P_{mtr}, and RL,

[3] The scripts are available at https://doi.org/10.5281/zenodo.6407128.

P_{rl}, pathways (Eq 5.12). The output, P, generates the sampling position of the model in the sensorimotor space, which is further transformed into a syllable vocalisation, as explained in section 5.2.2.

$$P_{mtr} = w_{mtr} * \mu_{mtr} \tag{5.10}$$
$$P_{rl} = w_{rl} * (\mu_{rl} + \xi) \tag{5.11}$$
$$P = P_{mtr} + P_{rl} \tag{5.12}$$

Learning

The RL pathway, μ_{rl}, is governed by reinforcement learning, following the REINFORCE rule (Eq 5.15) (Williams 1992). Local exploratory noise, ξ, is injected directly into the RL pathway, along with the performance (or reward) prediction error, PPE (Eq 5.11, 5.15). Performance prediction error, PPE, here corresponds to the difference between the performance evaluation at a given trial, R_{tr}, and the expected performance evaluation, $\overline{R_{tr}}$ (Eq 5.14). The motor pathway, μ_{mtr}, gradually consolidates the drive from the BG-led exploration, by maintaining a slow trace of the BG contribution, P_{rl}, and eventually learns to produce the desired vocalisation (Eq 5.13).

$$\Delta\mu_{mtr} = \eta_{mtr} * P_{rl} \tag{5.13}$$
$$PPE = R_{tr} - \overline{R_{tr}} \tag{5.14}$$
$$\Delta\mu_{rl} = \eta_{rl} * \xi * H(PPE) \tag{5.15}$$

where tr denotes the current trial, η_{rl}: learning rate within the RL pathway, η_{mtr}: learning rate within the motor pathway, $\bar{R}_{tr}$: running average of recently (100 trials) obtained performance evaluations and $H(x) = \begin{cases} 1 & x > 0 \\ 0 & x <= 0 \end{cases}$.

We supplement the reinforcement learning implemented by the BG pathway with two mechanisms derived from empirical evidences underlying vocal learning in songbirds.

(i) Taking inspiration from the delayed growth of the cortical motor pathway, we simulated the increasing influence of the motor pathway, μ_{mtr}, alongside the decreasing influence of the RL pathway, μ_{rl}. The maturation of the motor pathway is modeled using an asymptotically increasing contribution, w_{mtr}, towards the RA (Eq 5.10,5.16) while the influence of the RL pathway, w_{rl} is modeled using an asymptotically decreasing contribution (Eq 5.11,5.17).

(ii) Drawing on evidences from the studies of (Derégnaucourt et al. 2005) showing a daily post-sleep deterioration of song structure during the sensorimotor period (discussed in section 3.1.3), the output of the RL pathway, μ_{rl}, is shifted each morning. This shift is implemented as a random jitter, ϕ (Eq 5.19) added to the current BG output. These

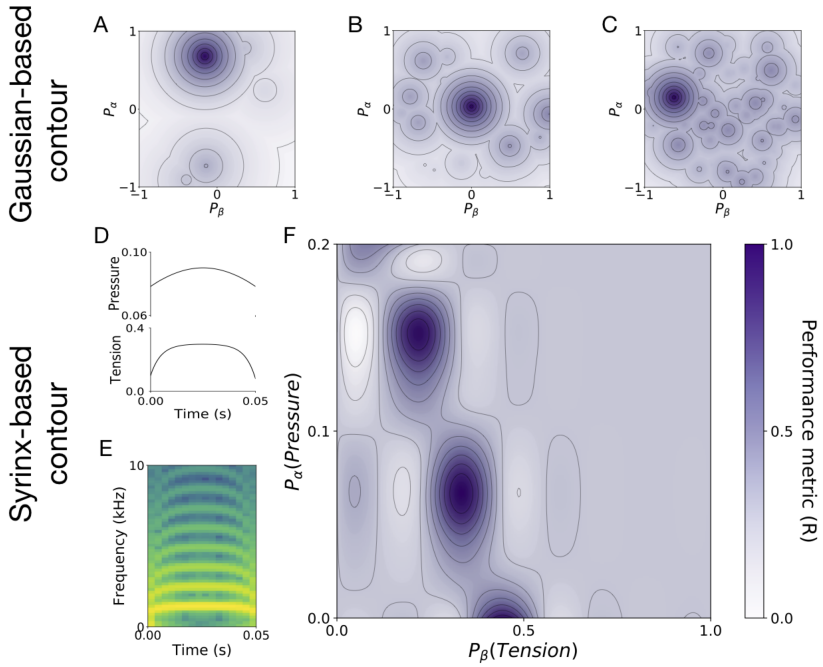


Figure 5.6: Various types of performance landscapes. The concentric circles show equipotential surfaces. **A-C**. Examples of Gaussian-based performance landscapes with 1 global optimum and **A**. 'low' (1-5) **B**. 'medium' (10-20) and **C**. 'high' (30-50) number of local optima. **D-F**. Reward contour generated using a model of the avian syrinx (Amador et al. 2013). **D**. The 50ms waves of tension and pressure used as input to the syrinx model to generate the target syllable. **E**. The spectrogram of a common zebra finch syllable chosen as the target syllable, as generated by the model. **F**. The performance landscape generated using the similarity between the target syllable and vocalisations generated over the parameter range used in (Amador et al. 2013). It has three global optima and several (=11) local optima.

two factors are weighted as per the daily consolidation trace, w^k, determined by the mean rectified performance prediction error experienced in the previous day k (Eq 5.18).

$$w_{mtr} = e^{-.5*c_1/tr} \tag{5.16}$$

$$w_{rl} = 1 - e^{-c_1/tr} \tag{5.17}$$

$$w^k = c_2 * \overline{f(PPE)^k} \tag{5.18}$$

$$\mu_{rl}^{k+1} = w^k * \mu_{rl}^k + (1 - w^k) * \phi \tag{5.19}$$

where c_i denotes scalar constants and $f(x) = xH(x)$.

Performance landscapes

The transformation of RA motor output into produced vocalizations, that are then compared to an auditory template of a tutor song for performance evaluation, represents a complex and largely unsolved problem. We therefore test the model in two different contexts. First, we map the 2D motor space of RA output to a 1D performance space through an arbitrary non-monotonic function artificially denoting the performance quality, detailed below. Second, we generate a biologically realistic performance landscape using an artificial model of the syrinx driven by RA output, detailed in this section.

In order to evaluate the performance of the model, we first create continuous performance landscapes (analogous to reward profiles), by transforming a 2D motor space to a 1D performance evaluation: The performance landscape is a set of Gaussian distributions, giving rise to several local optima along with one global optimum. Each optimum is represented by a 2D Gaussian distribution, and the performance landscape is the maximum of these distributions. We generate several contours by changing the number, position and widths of local optima, awarding a maximum performance evaluation of 60% of that of the global optimum. We categorise these set of contours into three classes with low (1-5), medium (10-20) and high (30-50) number of local optima. Figure 5.6a-c shows an example performance landscape from each class. These randomly generated contours provide us with explicit control over the complexity of the performance landscape, in terms of the number of local and global optima and their relative heights and positions. We take advantage of these properties to test the model's versatility and robustness by simulating the model's performance on different classes of such Gaussian-based performance landscapes.

Further, we test the performance of the dual pathway model with a reward contour that has been generated using a model of the avian syrinx (Amador et al. 2013). The 2D scalar output of the RA layer, P, is transformed to form the input signals for the avian syrinx model as per Eq 5.20, 5.21 (Figure 5.6d). The syrinx layer receives two input signals from the RA layer, corresponding to the air-sac pressure, $\alpha(t)$, and the tension of syringeal labia, $\beta(t)$. These input signals of pressure and tension lead to oscillations in the syrinx and trachea and generate a syllable vocalisation as a pressure wave of 50ms (T) (Figure 5.6e) (Amador et al. 2013). We construct a spectrogram from this oscillatory pressure wave, and choose a target syllable which is similar to a commonly-occurring syllable in zebra finches (shown in Figure 5.6e). We proceed to simulate several vocalisations over a range of input parameters (as described in (Amador et al. 2013)), and thereby, generate a performance landscape using a similarity metric, based on the correlation coefficient between the generated vocalisations and the target template. As shown in Figure 5.6f, the performance landscape has three global optima, as multiple configurations of the syrinx can produce the desired vocal output. Alongside these global optima, there are several (11) shallow local optima across the sensorimotor range. The motor control output of

the RA layer traverses the contour, thus generated, to obtain an evaluation of its performance.

$$\alpha(t) = P_\alpha + 0.04sin((T/2 + t) * \pi * 10) \tag{5.20}$$
$$\beta(t) = P_\beta - 0.2e^{-200t} - 10^{-5}e^{200t} \tag{5.21}$$

where the motor output of the RA layer, $P \propto [P_\alpha, P_\beta]$, T: duration of the vocalisation and $t \in [0, T]$.

Simulation

Each simulation of the model is run over 60 days, which is the typical duration of the sensorimotor phase for a zebra finch, with each day consisting of 1000 trials (Derégnaucourt et al. 2005). The noise level injected by the BG is initialised at 20% of the RA output range. As we reach the crystallisation stage, the RL noise level exponentially decreases to 10% of its initial value (=2%) which is comparable to the variance observed in the pitch of the vocalisations produced by adult birds (Sober, Wohlgemuth, et al. 2008).

Metrics

We measure the 'terminal performance' as the mean performance evaluation obtained during the last five days of a simulation. For the Gaussian-based performance landscape, the global optimum has an associated reward level of 1 while all other local optima have a maximum associated performance evaluation of 0.6. Therefore, we consider a simulation to be 'successful' if it achieves a performance metric above 0.6. For the syrinx-based performance landscape, we observed that the highest peak outside the global optimum has an associated performance evaluation of 0.55 while the global optima has an associated reward of 1. Therefore, we maintain 0.6 as the threshold above which a simulation is considered to be successful. Finally, the 'success rate' of the model for a given scenario is the proportion of successful simulations compared to the total number of simulations.

Benchmark algorithms

In order to compare the performance of the dual pathway model with established approaches, we build a framework with a single pathway architecture, implementing variants of RL. The single pathway injects exploratory noise, receives performance evaluation and governs motor output. In essence, we lesion the motor pathway allowing the RL pathway to be in sole control of the motor output.

First, we test the performance of the single pathway framework using the standard reinforcement learning approach (StdRL). Here, the learning rule is based on gradient descent, akin to the former RL pathway (Eq 5.15). Under this scenario, the variability of the output does not reduce over time as the influence of the BG remains intact due to the

absence of a motor pathway. Second, we compare the performance of the dual pathway system with a modified standard RL approach, i.e. RL with decreasing noise (DevRL). Here, we exponentially reduce the noise injected into the RL pathway, as learning progresses (Eq 5.22).

$$\xi_{tr} = \xi_0(1 - e^{-c_1/tr}) \tag{5.22}$$

where ξ_0 denotes the initial noise level injected into the pathway. Third, we compare the performance of the dual pathway system with simulated annealing (SA), a probabilistic technique, used to find the global optimum in discrete search spaces (Tsallis et al. 1996). This technique uses an explicit acceptance function M, to control the exploitation-exploration trade-off, or more specifically, the probability to move to lower rewarding positions. This acceptance probability is determined by an exponentially-decreasing temperature parameter Γ, weighted by the difference in performance evaluation between successive iterations ΔR (Eq 5.23-5.25). Note, the range of exploration available to the system at each timestep is not explicitly altered.

$$\Gamma_{tr} = 1 - e^{-c_1/tr} \tag{5.23}$$

$$\Delta R_{tr} = R_{candidate} - R_{current} \tag{5.24}$$

$$M(tr) = e^{\frac{\Delta R_{tr}}{\Gamma_{tr}}} \tag{5.25}$$

where ΔR_{tr} refers to the difference in performance evaluation between a randomly chosen candidate motor output, $R_{candidate}$ and the current motor output $R_{current}$. The candidate motor output is chosen in a range corresponding to $\eta_{rl} * \xi$ (Eq 5.15) to maintain an equivalent step size with the previous algorithms.

5.2.3 Results

In this section, we test the proposed data-driven algorithm on our simplified implementation of the dual pathway architecture. We verify the robustness of the model using different types of reward contours and modifying exploratory parameters. We, further, compare the performance of the model with established reinforcement learning approaches.

Sensorimotor learning by the dual pathway model

We simulate the vocal learning process using the algorithm described above governing the dual pathway architecture, on a Gaussian-based performance landscape with a 'medium' number of local optima (17) and 1 global optimum (Figure 5.7). Figure 5.7a shows that the motor output of the model P explores several local optima, including the global optimum. Each day, the RL pathway output is shifted, following which it performs gradient ascent to find the nearest local optimum, over the course of the day. Meanwhile, over

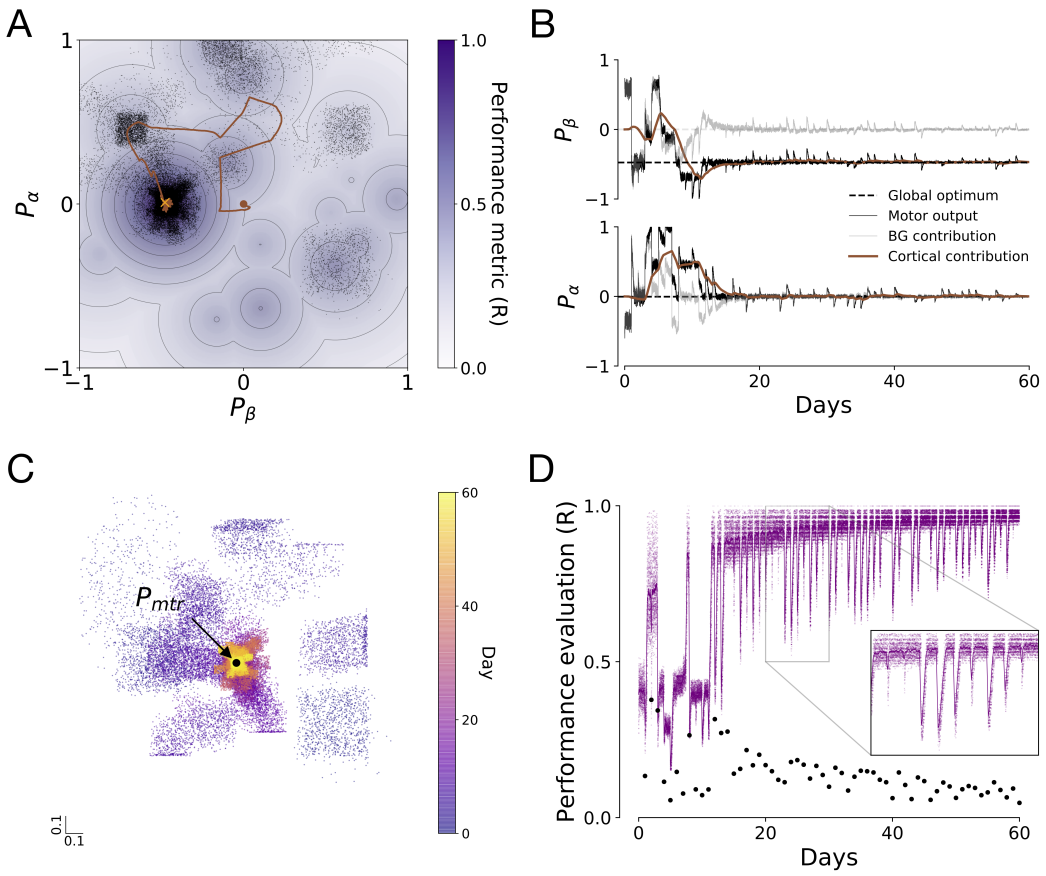


Figure 5.7: Simulation of the dual pathway model on a Gaussian-based reward contour with medium number of local optima (17) and 1 global optimum using 20% initial BG variability. **A** The cortical motor pathway, in brown, follows the BG-led exploration to several local optima on the performance landscape before coverging at the global optimum. The black dots denote the total motor output. **B** Initially, the contribution of the RL pathway P_{rl}, in grey, drives a strong bias in the motor output P, in black. As the contribution of the motor pathway P_{mtr}, in brown, reaches the global optimum, the BG contribution recedes. **C** The range of BG-led exploration, around the motor pathway, shrinks with development. Each dot represents the bias driven by the RL contribution P_{rl} at a given trial. **D** Performance evaluation, in purple, fluctuates over the coarse of learning on both daily (inset) and weekly timescales. The daily BG consolidation trace w^k, in black, determines the shift on the following day.

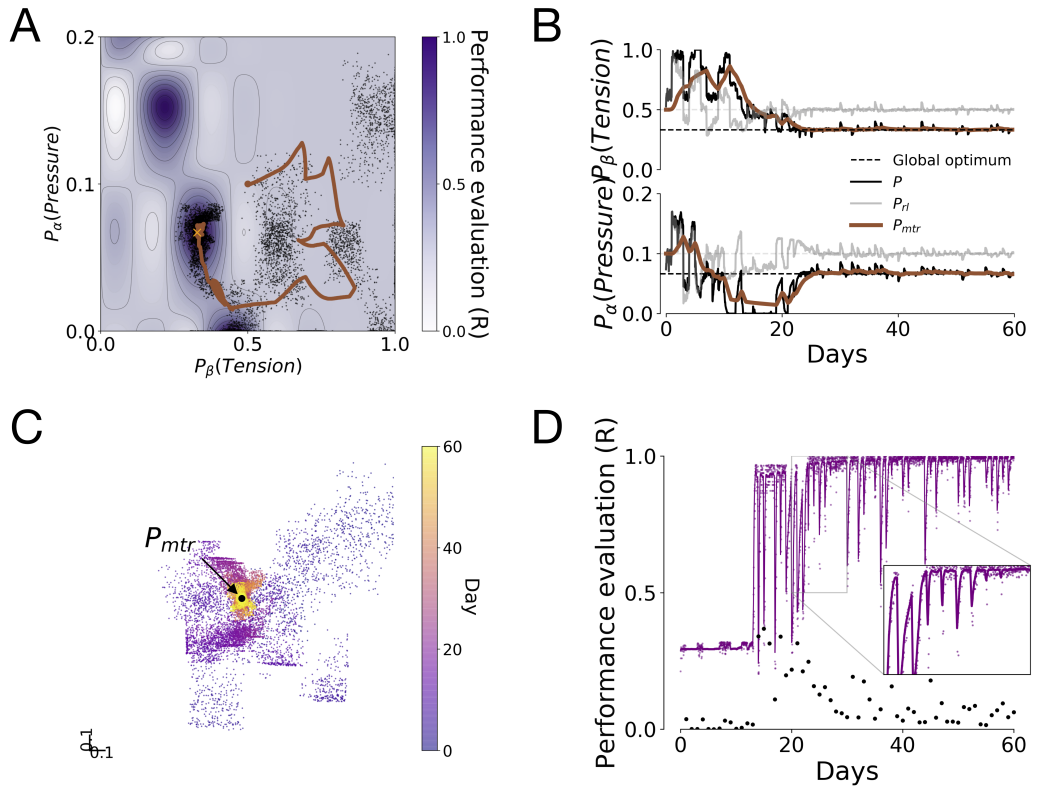


Figure 5.8: A demonstration of the dual pathway system on the syrinx-based reward contour with 3 global optima using 20% initial noise injected into the BG. **A** The cortical motor pathway, in brown, follows the BG-led exploration to several local optima on the performance landscape before coverging at the global optimum. The black dots denote the total motor output. **B** Initially, the contribution of the RL pathway P_{rl}, in grey, drives a strong bias in the motor output P, in black. As the contribution of the motor pathway P_{mtr}, in brown, reaches the global optimum, the BG contribution recedes. **C** The range of BG-led exploration, around the motor pathway, shrinks with development. Each dot represents the bias driven by the RL contribution P_{rl} at a given trial. **D** Performance evaluation, in purple, fluctuates over the coarse of learning on both daily (inset) and weekly timescales. The daily BG consolidation trace w^k, in black, determines the shift on the following day.

the weekly timescale, the motor pathway consolidates more information from the motor outputs that are produced more often. Figure 5.7a-b shows that the motor pathway does follow the model output to certain local optima. However, the dual pathway model is able to successfully evade it and eventually converge at the global optimum.

Figure 5.7b shows that the output is governed increasingly by the contribution of the motor pathway as the model approaches the global optimum. The contribution of the RL pathway induces a strong bias and high variability in the early stages. As learning progresses, the target information is consolidated into the motor pathway and the bias and variability induced by the RL pathway recedes. This helps the model to converge at the chosen peak. Figure 5.7c shows that in the initial days of learning, BG-led exploration ranges over a large area of the sensorimotor space around the contribution of the motor pathway, P_{mtr}. Eventually, the growth of the motor pathway curbs the influence of the RL pathway, resulting in a reduced range of exploration as well as increased exploitation tendency. Thus, despite the same amount of noise being injected into the BG, the variability in the motor output decreases, leading to the convergence of the vocal production at the position consolidated within the cortical pathway. Progress in learning is accompanied by a non-monotonic increase in the average performance evaluation obtained by the model, as shown in Figure 5.7d. The daily shift in the RL pathway output induces dips in the performance evaluation, which rapidly improves within the time course of a single day, as the RL pathway finds the best local optimum within the exploration range. The variability in performance evaluation reduces as learning proceeds and as the influence of the BG-injected noise reduces over time.

We further demonstrate the versatility and robustness of the dual pathway framework by testing the system on a relatively more biologically plausible performance landscape generated using a model of the avian syrinx, as explained in section 5.2.2 (Amador et al. 2013). Figure 5.7 demonstrates a simulation using the dual pathway architecture on a sample artificially generated reward contour with 10 local optima and 1 global optimum. Similar to the previous case, Figure 5.8a shows that the sampling position of the model P, shown in black, explores several local optima, including the global optima. Each day, the BG locus is displaced to a different position and it then performs gradient descent to find the nearest local optimum. When the variance in exploration is less than the width of the optimum, the gradient descent leads the system to get stuck at local optimum, while the motor pathway gradually tracks the BG-led exploration. Now, each day a new region is explored due to the daily BG displacement, which helps guide the motor pathway away from the local optimum. Meanwhile, the motor pathway consolidates more information from the optima that are sampled more often. When the reward obtained in a neighbourhood is high, the daily BG displacement would be slightly biased towards its vicinity. Hence, the global optimum is more likely to be sampled, which in turn increases the probability of the motor pathway tracing its position. Figure 5.8a-b shows that the motor pathway does follow the model output to certain local optima.

Figure 5.8c shows that in the initial days of learning, exploration ranges over a large area of the sensorimotor space. However, over time as the influence of the BG pathway reduces, the exploration gets restricted to a region around the locus of the motor pathway. Figure 5.8b shows that as the model approaches the global optimum, the output is governed increasingly by the contribution of the motor pathway (in brown). While the contribution of the RL pathway induces a strong bias and high variability in the early stages, as learning progresses and the target information is consolidated into the motor pathway, the bias and variability induced by the RL pathway recedes. This helps the model to converge at the chosen peak. Progress in learning is accompanied by an increase in the average reward obtained by the model, as shown in Figure 5.8d. The daily displacement of the BG pathway induces dips in the performance evaluation, which rapidly improves within the time course of a single day, as the BG pathway finds the best local optimum within the neighbourhood/exploration range. The variability in reward reduces as the model approaches the global optimum and as the influence of the BG-injected noise reduces over time (shown in grey). Thus, we observe that the dual pathway framework enables the system to evade local optima and converge at the global optimum.

Robustness

Having observed a demonstration of the dual pathway architecture, in this section we proceed to demonstrate the versatility and robustness of the dual pathway framework. We test the system under various types of scenarios and compare the dual pathway model with a set of benchmarks using a single pathway framework.

First, we verify the stability of the model under identical conditions (20% initial RL noise level) on the Gaussian-based performance landscape consisting of "medium" (10-20) number of local optima. We simulate motor learning within this scenario using 100 different seeds for the random number generator resulting in varying performance landscapes and observe that the model is successful in finding the global optima in 76% cases (success defined as per section 5.2.2) (Figure 5.10a). Moreover, in successful simulations, the model consistently achieves a high terminal performance (above 0.9).

Second, we verify the robustness of the model under different types of performance landscapes. We test whether the number of local optima has an impact on the model performance. In order to do this, we randomly generate 100 different performance landscapes each with "low" (1-5) and "high" (30-50) number of local optima in addition to the global optimum. We simulate sensorimotor learning on these performance landscapes and observe that the model is successful in 92% and 64% cases, respectively (Figure 5.10a). Further, in successful simulations, the model has a terminal performance higher than 0.9. Thus, we observe that the model success rate decreases with increasing number of local optima (Figure 5.10a), however the terminal performance of successful simulations is unaffected.

Third, we verify the performance of the model under a biologically-inspired performance landscape, built using a model of the avian syrinx (described in section 5.2.2). Under identical levels of BG-induced noise (20% initial RL noise), we simulate vocal learning using 100 different seeds for the random number generator. We observe that the model is successful in 92% cases and reaps a terminal performance above 0.9 in these cases (Figure 5.10b). Thus, the model is capable of sensorimotor learning under biologically realistic performance landscapes as well.

Fourth, we test the stability of the model when higher levels of noise are injected into the RL pathway. We simulate the model with different higher initial levels of noise (30%, 40% and 50%) injected into the RL pathway. We observe that the increase in the RL noise level leads to an improvement in the success rate of the model (Figure 5.10b). On the other hand, there is a slight decrease in the terminal performance of successful simulations with increase in initial RL noise levels.

Benchmarks

Now, we compare the performance of the dual pathway architecture with that of the single pathway framework, as described in section 5.2.2, when governed by different learning rules.

First, we test the performance of the single pathway framework using the standard reinforcement learning approach (StdRL), as described in section 5.2.2. We observe that, over 100 simulations, each at 20% RL noise level using the syrinx-based performance landscape, the dual pathway framework (92% success rate, n=100, median terminal performance=0.96) performs significantly better (Mann–Whitney U=9437, $p<0.01$) than the single pathway framework with StdRL (55% success rate, n=100, median terminal performance=0.62), as shown in Figure 5.10b. Now, increasing the noise level injected in the pathway does facilitate the model to escape local optima and find the global optimum. However, this leads to a drawback where the high variability leads to a low terminal performance (n=100, median terminal performance=0.4 at 50% noise) being harvested by the model post learning (when RL noise is above 20%, terminal performance is below 0.6 even at global optimum).

Second, in order to address the aforementioned disadvantage of StdRL, we compare the performance of the dual pathway system with a modified standard RL approach, RL with decreasing noise (DevRL), as described in section 5.2.2. Here, we explicitly reduce the noise injected into the single pathway, as learning progresses. We observe in Figure 5.10b that the dual pathway system achieves a higher success rate than the single pathway system with the DevRL approach at all noise levels (Mann–Whitney U=3816, $p<0.01$, n=100 at 50% noise). Now, increasing the level of noise injected into the pathway does help resolve this issue, and drastically improves the success rate, however this is at the cost of high vocal variability and reduced terminal performance.

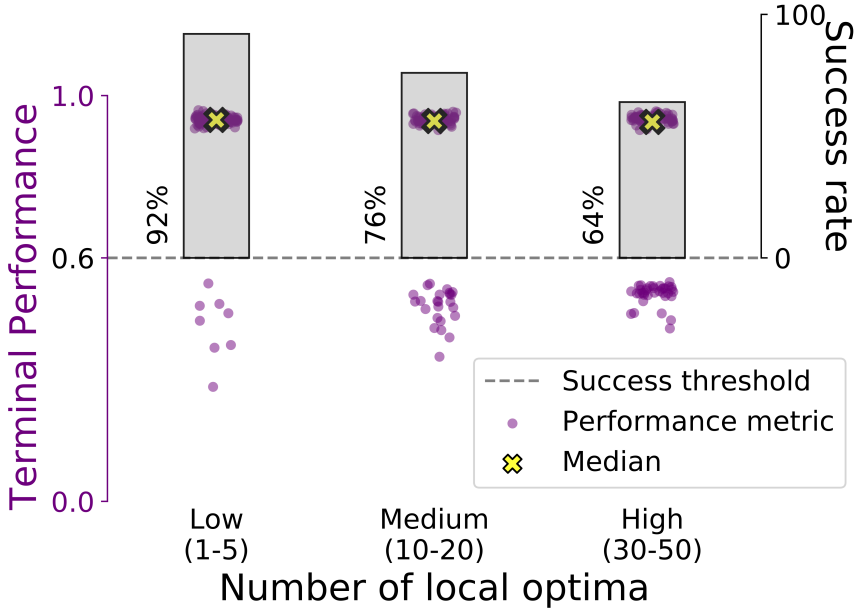


Figure 5.9: Performance of the dual pathway architecture on the Gaussian-based performance landscapes with 'low', 'medium' and 'high' number of local optima, at 20% initial RL noise. The grey bar and the percentage value next to it denote the success rate, i.e. the proportion of simulations with a high terminal performance (above 0.6). The purple dots represent the terminal performance of individual simulations, i.e., the mean performance evaluation obtained in the last five days. The opacity of the dots denotes the number of simulations that received a similar terminal performance. Yellow crossed markers represent the median terminal performance in each scenario.

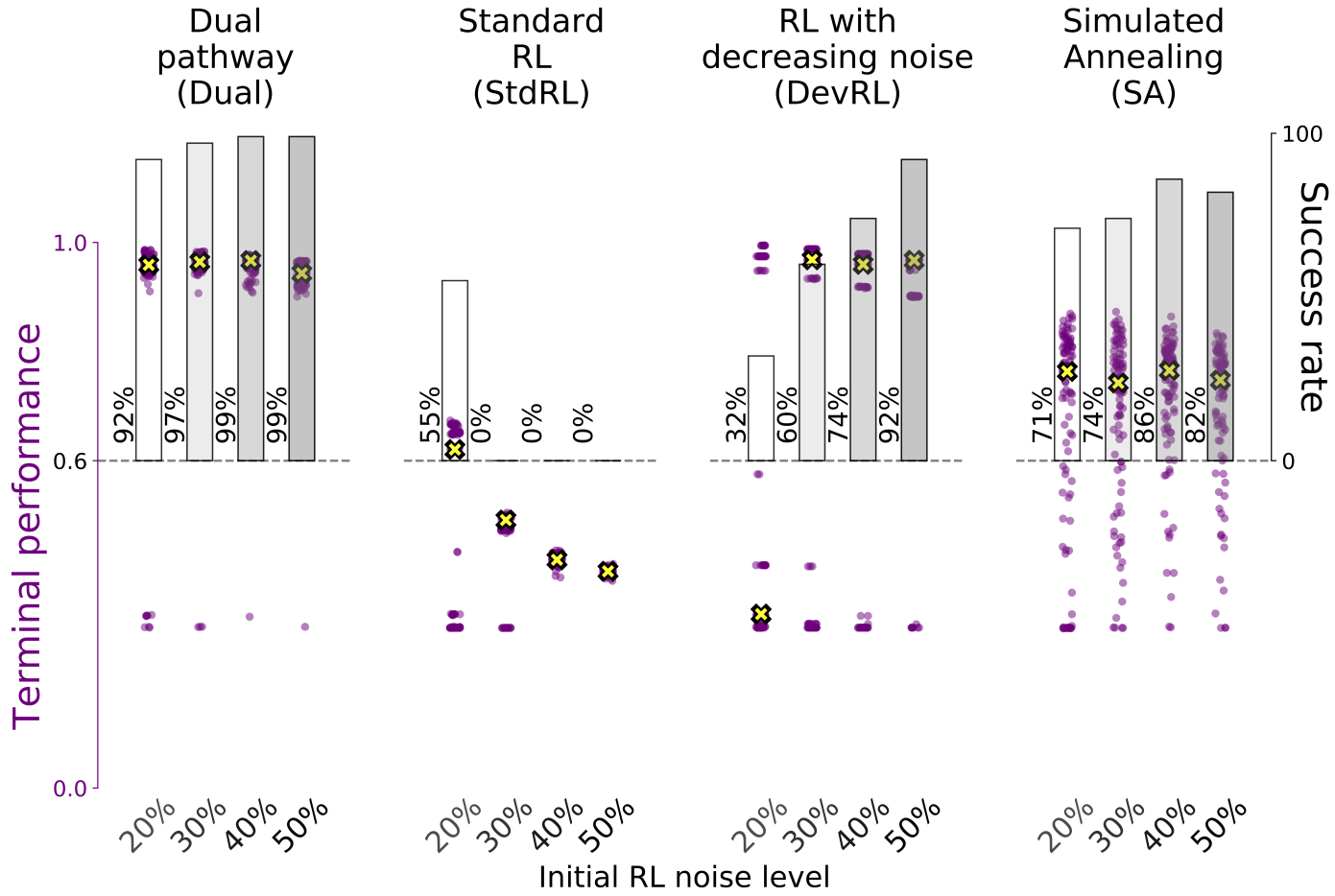


Figure 5.10: Performance of the dual pathway architecture on the syrinx-based performance landscapes at different noise levels and comparison with benchmarks. The grey bar and the percentage value next to it denote the success rate, i.e. the proportion of simulations with a high terminal performance (above 0.6). The purple dots represent the terminal performance of individual simulations, i.e., the mean performance evaluation obtained in the last five days. The opacity of the dots denotes the number of simulations that received a similar terminal performance. Yellow crossed markers represent the median terminal performance in each scenario.

Third, we compare the performance of the dual pathway system with a single pathway system implementing simulated annealing (Tsallis et al. 1996). When lower levels of noise are injected into the pathway, simulated annealing performs better (71% success rate) on the syrinx-based performance landscapes than the standard RL (55% success rate) approach due to its superior ability to escape local optima. With increasing noise, this approach yields improved success rates. On the other hand, the dual pathway framework obtains a higher success rate (92%) as well as significantly higher terminal performance metrics (Mann Whitney U=9306, $p<0.01$) than the simulated annealing approach (71%) at 20% RL noise as well as higher noise levels (Figure 5.10b).

Thus, the dual pathway framework, with its advantages and shortcomings, provides a viable approach for sensorimotor learning, even under low noise conditions.

5.2.4 Discussion

Inspired from the dual pathway architecture of the vocal learning circuitry in songbirds and by the large overnight changes undergone by juvenile vocalizations during early phases of learning (Derégnaucourt et al. 2005), we proposed a new algorithmic implementation for efficient sensorimotor learning in the face of uneven performance landscapes with multiple optima.

In songbirds, a BG pathway drives early vocalizations and rectifies vocal output through RL guided by a dopaminergic signal denoting performance evaluation (Gadagkar, Puzerey, et al. 2016). In parallel, a cortical motor pathway consolidates BG-driven changes, utilising activity-dependent plasticity at the HVC-RA synapses (Mehaffey et al. 2015). Similarly, in the model, a trace of the RL pathway output is gradually consolidated within the motor pathway (Eq 5.13). First, to reflect on the large day-to-day changes in juveniles vocalizations that do not necessarily align with the long-term improvement of the song (Kollmorgen et al. 2020), the initial state of the RL pathway is reset every day with a partial copy of the previous day's final output (Eq 5.19). As synapses in the nervous system are known to be very volatile (Holtmaat et al. 2005; Roberts et al. 2010), the daily shift in the BG output could reflect large overnight changes in the HVC-X synapses, inducing a partial loss of the memory formed by RL in the previous days. In the model, this synaptic fluctuation has been simulated using the daily consolidation trace w^k. Such a biologically plausible mechanism for partial forgetting in the RL pathway facilitates the exploration of a new region of the performance landscape each day, which helps guide the motor pathway away from the local optimum. Thus, while the partial forgetting within the BG does not directly contribute to the formation of strong synaptic connections and is detrimental in the short timescale of a few hours, in this case, partial forgetting is beneficial in the long term as it helps unlearn sub-optimal performances. Second, the cortical motor pathway exhibits a delayed maturation during sensorimotor learning (Mooney and Rao 1994a). It has been hypothesised that the development of the HVC-RA synapses plays a role in reducing the RA sensitivity towards LMAN input and thereby, gradually suppresses the

BG-induced bias in motor output (Garst-Orozco et al. 2014a). Correspondingly, the relative influence of the two pathways in the model to the motor output changes over development (Eq 5.16, 5.17). As the RL pathway contribution to the global output decreases over development, so does the variability of the vocal output due to the daily shift and the noise in RL pathway output, consistent with the progressive decrease in song variability during learning in songbirds (Derégnaucourt et al. 2005). Interestingly, a parallel can be drawn between the effect of these two mechanisms and simulated annealing, a technique used to optimise stochastic gradient descent (Tsallis et al. 1996). The daily shift in BG output facilitates escape from local optima, akin to discrete fluctuations implemented within simulated annealing. The change in relative influence exerted by the two pathways plays a strikingly similar role as temperature in simulated annealing. Indeed, the exploitation-exploration trade-off is initially skewed towards exploration with high trial-by-trial variability and daily changes, analogous to the high temperature condition. As learning progresses, it leans towards exploitation akin to low temperature scenarios in annealing, as mentioned in (Derégnaucourt et al. 2005).

It is to be noted that such a heuristic does not convey any guarantees concerning the optimality of the solution. Other optimization techniques used in machine learning such as simulated annealing may succeed in pathological landscapes that pose difficulties to the model proposed here. They are however less biologically realistic as they rely on a long-term maintenance of the memory of all explored options. Alternatively, developmental regulation of variability in song through sexual hormones is known in songbirds and may provide a biologically realistic mechanism to partially overcome the drawbacks of gradient descent (as implemented above using DevRL) (Sizemore et al. 2011).

Using the conceptual model presented above as a basis, in the upcoming section, we further test the viability of the proposed dual pathway framework towards sensorimotor learning, by building an analogous network composed of rate-coded sigmoidal units representing the activity of neuronal populations.

5.3 Neural implementation of dual pathway model

In the previous section, we looked into the idea of interplay between two parallel pathways, using a highly simplified abstract model. We would like now to explore whether a neural implementation of this dual pathway is feasible when taking into account the anatomical constraints of the song system in zebra finches, both at structural and functional levels.

In songbirds, the cortical pathway connecting HVC and RA is necessary for the production of song (F. Nottebohm, Stokes, et al. 1976). RA-projecting HVC neurons have been shown to burst sparsely at single, precise timepoints during song (Hahnloser, Kozhevnikov, et al. 2002). As a population, these HVC neurons could form a representation of timing

within the song sequence. Further, as song learning progresses, RA neurons fire in a more stereotyped, sparse and bursty pattern, as opposed to a variable and distributed firing pattern during early subsong (Ölveczky, Otchy, et al. 2011). Aronov et al. (2008) have shown that the variable firing pattern during babbling is driven by the parallel BG-thalamo-cortical pathway. On the anatomical side, the pathways within the AFP which project to the RA and from RA to downstream areas controlling vocalisation are topographically organised (Luo, Ding, et al. 2001; Vicario 1991). Moreover, there are around 13 to 15 neurons in the RA in each hemisphere in the early sensorimotor phase (Bottjer, Miesner, et al. 1986; Herrmann and Arnold 1991). During this period, between 5 to 10 thousands neurons from LMAN project to project to RA (Bottjer, Glaessner, et al. 1985). 23-40 thousand HVC neurons exist in one hemisphere, during the sensorimotor period, however only a fraction of these project to RA and are active during song ((Herrmann and Arnold 1991)). In this section, we attempt to take into consideration these features of the song system and investigate biologically plausible mechanisms which could support the analytical model, discussed in section 5.2.

5.3.1 Methods

Architecture

From the empirical observations outlined above and building the analogy with the model presented in section 5.2, we designed a model comprising four layers representing respectively HVC, RA, BG (reduced model of AFP) and MC (motor control). These structures are connected as follows (see Figure 5.1a): the HVC and the RA are connected via one direct pathway and a parallel pathway though the BG. The RA projects to downstream regions which control respiration and syringeal musculature, represented here as the motor control layer (MC). The model has been scaled according to the proportions of the respective number of neurons in the corresponding nuclei of the song system introduced in previous section. This leads to the HVC and RA having 100 rate-coded units while the BG layers has 50 rate-coded units. Each unit represents a population of inhibitory and excitatory neurons. The RA neurons are modeled using very steep sigmoidals for activation functions, which bound the RA firing rate between -1 and 1. The firing rates of the BG neurons are linearly bound between -1 and 1. The HVC and RA are fully connected with each other, as well as the HVC and BG. The weights between these layers are plastic, based on the activity-dependent plasticity found at the RA and BG synapses. The BG-RA and RA-MC pathways are topographically connected in clusters, with fixed positive weights (path C).

$$J_{RA} = sig(\frac{J_{HVC} \cdot W_A}{n'_{HVC}} + 2 * \frac{J_{HVC} \cdot W_B}{\sum^{n_{BG}} W_B}) \tag{5.26}$$

$$sig(x) = \frac{1}{1 + e^{-1*(x-sm)*ss}} * 2 - 1 \tag{5.27}$$

$$J_{BG} = \frac{J_{HVC} \cdot W_B}{n'_{HVC}} + \xi \tag{5.28}$$

$$J_{MC} = \frac{J_{RA} \cdot W_C}{\sum_0^{n_{RA}} W_C} \tag{5.29}$$

where n' denotes the number of active neurons in a layer at a given timepoint, p the summation of fixed weights projecting to a neuron in the layer, J the activity of a layer and ξ represents the exploratory noise injected into the BG layer. The steepness and displacement of the sigmoidal function is determined by ss and sm. The sum of the fixed weights in the BG-RA and RA-MC pathways serve as a normalisation factor to maintain the input activation between -1 and 1.

Learning

The plastic weights in the BG pathway, W_B, are modeled using reinforcement learning, specifically the REINFORCE rule (Williams 1992). Local exploratory noise, ξ, is introduced within the BG layer. Additionally, the BG layer receives a relative evaluation of performance (R). The plastic weights in the HVC-RA pathway, W_A, are modeled using Oja's rule, a normalised variant of Hebbian learning, where the co-activation of two neurons in both layers, potentiates the mutual synapse (Oja 1982). All plastic weights are bound between 1 and -1.

$$\Delta W_B = \eta_B * J_{HVC} * J_{BG} * (R - \bar{R}) \tag{5.30}$$

$$\Delta W_A = \eta_A * J_{RA} * (J_{HVC} - J_{RA} * W_A) \tag{5.31}$$

where η_A and η_B represent the learning rates of their corresponding pathways.

The reinforcement learning driven by the BG pathway is supplemented with behavioural features, drawn from empirical evidences (Derégnaucourt et al. 2005). More specifically, a daily deterioration in song performance has been shown to occur post-sleep in juvenile birds during the sensorimotor period. This post-sleep deterioration is implemented with 1000 iterations of random potentiation and de-potentiation of the HVC-BG synapses overnight, which results in a slight jitter of the BG output the following day. Unlike the

analytical model, the change in the influence of the two pathways, due to the delayed innervation of RA by the HVC axons, is not explicitly modeled here.

$$\Delta W_B = \eta_B * J_{HVC} * \xi * (1 - w_k) \tag{5.32}$$

where w_k represents the daily consolidation trace, which is the cumulative potentiation experienced by the HVC-BG synapses over a given day, k.

Performance landscapes

The model is tested on the same set of performance landscapes, as described in the previous section 5.2.2. Different set of contours are generated using a superimposition of Gaussians, giving rise to several local optima and one global optimum, with varying positions and depths. The output of the model is used to denote the position on the performance landscape.

Simulation

Within each simulation, the model attempts to learn a motif of 4 syllables, over a 60-day sensorimotor period. Each day comprises 1000 trial motifs. Each motif consists of the 4 syllables in a fixed sequence. Each syllable corresponds to a different performance landscape within the same motor range. The local exploratory noise injected into the BG is maintained at 20% of the BG activity range. After a 60 day sensorimotor period, we we make the equivalent of a lesion in LMAN outputs by silencing BG outputs to the RA, i.e. the input from the BG pathway to the RA is inactivated. We can then observe the performance of the model when governed solely by the cortical pathway.

Metric

Described in section 5.2.2.

5.3.2 Results

Figure 5.11 demonstrates the progression of sensorimotor learning for each individual syllable within a song motif. In the example simulation shown in Figure 5.11, we observe that the motor output traverses several regions of the sensorimotor space, over the course of a few weeks. Each day, the model explores a new local optimum and converges eventually to the global optimum. The search for the optimal solution of the four syllables is carried at different rates resulting in different times for finding the optimal one.

Figure 5.12A shows the progression in the quality of each syllable vocalisation over time. The performance evaluation for each syllable changes at different rates. Some syllables are learnt faster than others. At the beginning of each day, there is a varied dip

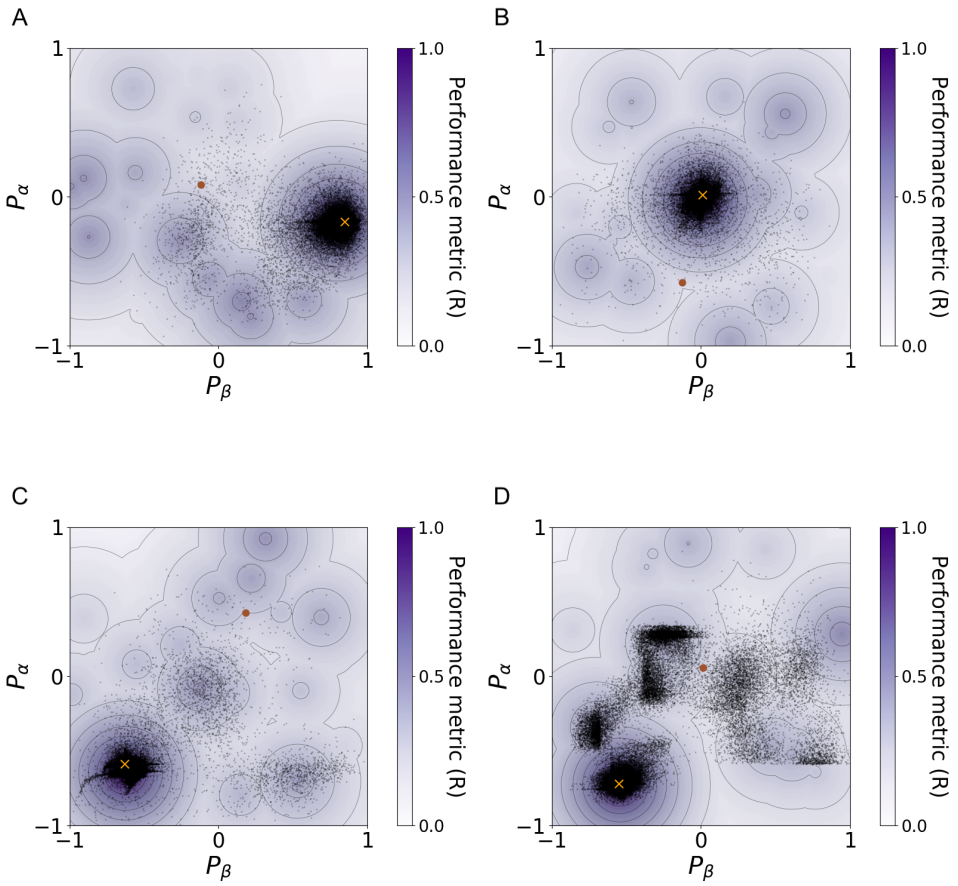


Figure 5.11: A demonstration of the dual pathway model with rate-coded neurons on several Gaussian-based reward contour with a medium number of local optima and 1 global optimum using 20% BG variability. The motor output (black dots) is driven to different regions of the sensorimotor space, due to the daily jitter experienced within the BG pathway. Within the day, the BG pathway helps the system find the local optima, while the motor pathway maintains a trace of this exploration. Over the course of several weeks, the BG pathway explores several such local optima, with the cortical motor pathway gradually consolidating this information and ultimately converging at the global optimum. The initial and final motor outputs are shown in brown and yellow, respectively.

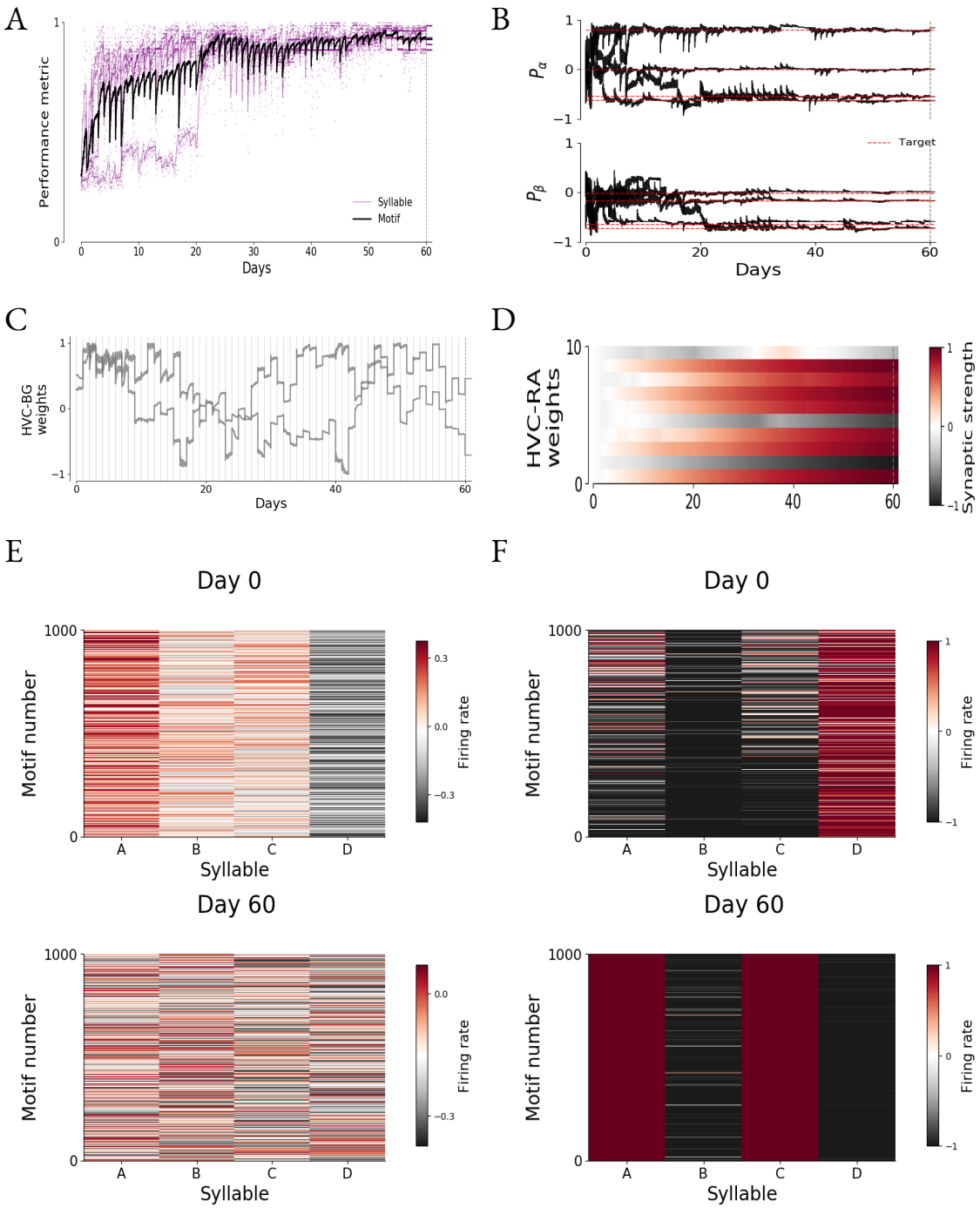


Figure 5.12: Progression of sensorimotor learning by the dual pathway model on a Gaussian-based reward contour with medium number of local optima and 1 global optimum. **A** Performance evaluation fluctuates over the coarse of learning on both daily and weekly timescales. The performance evaluation for each syllable is shown in purple, while the average performance evaluation over the motif is shown in black. **B** Progression of motor output, corresponding to the 2-D position on the performance landscape. Initially, the motor output, in black, is highly variable due to being primarily driven by the BG pathway. As learning progresses, the variability of motor output reduces. The target output is shown in red. **C** The weights of HVC-BG synapses remain variable across learning, and experience overnight discontinuous changes. **D** The weights of HVC-RA synapses develop slowly and ultimately saturate. **E** The activity (firing rate) of BG units remains variable across learning. **F** The activity (firing rate) of RA units is highly variable in the beginning of learning. As the HVC-RA synapses grow, the RA activity develops a bursty pattern. **A, B, C, D** The vertical dotted black line represents the point of BG lesion i.e. inactivation of the BG inputs to RA. A decrease in variability of output occurss post BG lesion.

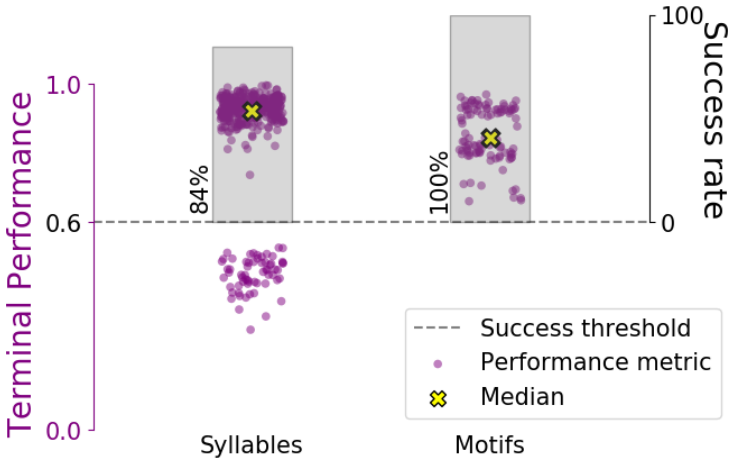


Figure 5.13: Robustness of the dual pathway model. The purple dots represent the terminal performance of each syllable, on the left, and of each motif, on the right. The grey bar denotes the success rate, i.e. the number of simulations with a terminal performance over 0.6. In over 100 simulations, with a target motif of 4 syllables each, the model successfully learns 84% of the target syllables. Moreover, in most successful simulations, the model achieves a terminal performance greater than 80%.

in performance evaluation of each syllable. This deterioration in song quality reduces as the song reaches crystallisation. The overall quality of the song motif improves non-monotonically over the sensorimotor period. Post lesioning of the BG pathway, the variability in performance evaluation reduces. As shown in Figure 5.12B, at the beginning of sensorimotor learning, the variability in motor output is relatively high and decreases as learning proceeds. The trajectory from the initial motor output to the target output, for each syllable, is not monotonic. Post BG lesion, the variability in motor output further reduces. The HVC-BG synaptic weights experience activity-dependent plasticity during the day time and undergo a sudden change overnight (Figure 5.12C). The firing rate of the BG neurons fluctuates across the learning period. Figure 5.12D shows the slow, gradual growth of the HVC-RA synaptic weights. Correspondingly, the neurons in the RA display a variable firing rate in the beginning of sensorimotor learning. The firing rate of these neurons towards the crystallised phase demonstrates a more binary nature, indicating bursts of spiking activity, especially post BG-lesion.

Further, we test the robustness of the dual learning framework, by testing the model on different performance landscapes, while learning different song motifs. In over 100 simulations of the model attempting to learn a four-syllable song, the system successfully produces the target syllable in more than 84% cases (considering 400 individual syllables). Moreover, in successful simulations, the model reaps a high reward and achieves a terminal performance of greater than 60%, as shown in Figure 5.13. Similarly, when con-

sidering the quality of the motif as a whole, 100% of the simulations were able to yield a terminal performance of above 60%.

5.3.3 Discussion

The dual pathway model presented in this section serves as an extension to the analytical model presented in section 5.2. The model presented here tests the hypothesis proposed by the analytical model under the constraints of a biological neural circuitry, using an analogous network with rate-coded neurons. Similar to the analytical model, the BG layer, here, implements reinforcement learning and serves as a tutor for the cortical motor pathway.

Figure 5.11 and 5.12 show the dual pathway system learning one song motif comprised of four syllables, over a 60 day sensorimotor period. We observe that the BG pathway initially drives exploration leading to a highly variable motor output and performance evaluation, as shown in Figure 5.12a, d. Meanwhile, the cortical motor pathway gradually consolidates the BG-led exploration within its synapses, using activity dependent plasticity. Figure 5.12a shows the non-monotonic increase in performance evaluation. As the influence of the cortical pathway increases, the variability in motor output reduces and the post-sleep deterioration in motor output decreases. Due to non-overlapping syllable encoding within the HVC, learning one syllable doesn't interfere with other syllables.

Gradient-descent based reinforcement learning often tends to get stuck in local optima. Inspired from the post-sleep deterioration in song quality observed in songbirds, we posit that the displacement in motor output helps escape local optima. We model this displacement as an effect of randomised overnight volatility of spines within the BG pathway. This volatility is partially modulated by the cumulative potentiation experienced over the day by the HVC-BG synapses. The more the change in performance evaluation (signalled by dopamine) over the course of the day, the stronger the potentiation experienced by the BG pathway and weaker the random synaptic fluctuation during the night. This helps the system evade local optima and converge at the global optimum. Over the course of every day, the BG pathway guides the dual pathway system to the nearest local optimum, under a gradient descent protocol. At the beginning of each day, the overnight volatility experienced by the HVC-BG synapses, leads a new region of the sensorimotor space to be explored. This overnight volatility of the BG pathway is modulated partially by the daily consolidation trace, w_k, i.e., the amount of potentiation experienced over the course of the day. The more the potentiation experienced in a given day, the less the the BG synapses are affected during nighttime, and less the post-sleep song deterioration. Consequently, over the sensorimotor period, the cortical motor pathway gradually consolidates the BG-led exploration and converges at the global optimum.

On the other hand, the RA activation pattern is steered by the BG pathway, in order to produce the desired motor output. The HVC-RA synaptic weights grow over time using the Hebbian learning rule which reinforces the RA activation pattern determined by the

BG. Over the sensorimotor period, the cortical motor pathway grows steadily and can exert a growing influence over the motor output. This results in a less variable motor output as well as performance evaluation as the target output is learnt. With the growth of the HVC-RA synapses, the RA neurons(modeled using steep sigmoidals as the activation function) get saturated by strong excitatory input from the HVC, leading to firing rates indicative of burst-like spiking activity, as seen in songbirds within the crystallised phase. This saturation of RA neurons leads to the influence of the BG inputs being suppressed. This is in accordance with the hypothesis by Garst-Orozco et al. (2014b) that the strengthening of the excitatory HVC projections to the RA leads to an increased inhibitory nature of the RA, which, in turn, suppresses the largely NMDAR-mediated glutametargic inputs from the LMAN. Unlike the analytical model, we do not explicitly model the change in influence between the two parallel pathways, here. The growth of the HVC-RA weights, in itself, leads to a suppression of the BG influence and decrease in behavioural variability.

Thus, the dual pathway framework, incorporating the delayed maturation of the cortical pathway and overnight synaptic volatility within the BG pathway, provides a robust and biologically plausible mechanism to evade local optima and facilitate sensorimotor learning.

5.4 Discussion and perspectives

In this chapter, we exploited anatomical and behavioural features of avian vocal learning to investigate mechanisms underlying sensorimotor learning. While in the first section, we look into a particular role of the initial babbling phase of vocal learning, typically before 40 days post hatch, in the second and third sections, we further explore the interplay between the subcortical and cortical structures between 30-90 days post hatch. We exploit empirical evidences concerning the cortical and BG-thalamo-cortical pathways to build a dual pathway framework, such that while one pathway is responsible for song production, the other functions as a tutor. Morphological studies showed the delayed innervation of RA by HVC axons, as opposed to the completion of the anterior forebrain pathway, or avian BG loop. This provides a biologically realistic mechanism to potentially control the exploration exploitation trade off. It also provides a plausible explanation as to why juveniles learn poor imitations of tutor song when deprived of tutor interactions in the early sensorimotor period. Thus, we confirm our hypothesis concerning a potential functional role played by structural plasticity in facilitating sensorimotor learning.

Moreover, the findings of (Derégnaucourt et al. 2005) reveal a pronounced deterioration in song structure after night-sleep, to be regained after intense singing in the morning. We interpret this apparent deterioration as an opportunity to explore a novel region of the landscape, implementing a variation of simulated annealing. Within simulated annealing, new candidate solutions are discontinuously explored, while maintaining a memory of the best solution. However, such algorithms are computationally intensive and do

not adhere to biological constraints. The dual pathway framework provides a biologically plausible alternative wherein one pathway conducts partial discontinuous exploration while the parallel pathway maintains a trace of this exploration. The dual pathway presented here, implements this discontinuous exploration by exploiting spine volatility within the BG. Although spine volatility during juvenile vocal learning in the area X of zebra finches hasn't been studied, dendritic spines have been shown to be transient even within the timescale of a day, in rodents, and a large turnover of dendritic spines in the HVC has been associated with sensory learning, as discussed in chapter 2. Additionally, simulated annealing uses an explicit temperature parameter to control the trade-off between exploration and exploitation. We propose that the delayed maturation of the cortical pathway could indirectly play a key role in controlling such an exploration-exploitation trade off.

Vocal learning involves several aspects in addition to the learning the appropriate neural code to produce a desired vocalisation in response to an internal signal. For instance, vocal learning in songbirds also involves the learning of the desired sequence of vocal gestures. It is important to note that in this study, we do not investigate the learning of sequence structures or the neural encoding of time, in terms of generating vocalisations at specific timings and duration, as well as, the delated evaluation of recently produced and perceived vocalisations.

From a theoretical perspective, the avian song system can be considered as a dual learning system, one is reactive and dependent on instantaneous reward (reinforcement learning) while the other is much slower and independent of reward but can strongly bias the output. Reinforcement learning builds an approximation of the local error contour via its exploration techniques, and can implement gradient descent in its pursuit to maximize the expected reward (or minimise the error) (Sutton, Barto, et al. 1998). However, as discussed earlier, under uneven contours in continuous action spaces, reinforcement learning using vanilla gradient descent can result in sub optimal solutions. In this chapter, we have encountered one non-gradient based alternative to circumvent this issue. Simulated annealing allows you to evade being stuck in local optima by exploring the candidate solutions in a discontinuous manner. Although this approach directly is biologically unrealistic, we have discussed a possible mechanism for birds to reap similar benefits, as those conferred by simulated annealing, in an analogous manner. On the other hand, machine learning has often supplemented gradient descent with algorithms, such as, Adagrad, Adam, RMSprop, etc (Ruder 2016). On a core level, these algorithms take advantage of the fundamental notion of momentum in the field of machine learning. Momentum is a simple and popular technique in supervised learning for improving stochastic gradient descent and to escape local minima (Goh 2017). This technique averages the last few gradients in order to maintain a consistent direction of the gradient. This has been proven to work better and faster than pure stochastic gradient descent. Very recently, such usage of momentum has been adapted to the framework of reinforcement learning in the case of the value iteration algorithm (Vieillard et al. 2020) but the technique can be adapted to

any reinforcement learning algorithm, including actor-critic architectures. For this latter case, it is a matter of replacing the critic by the average of successive critics.

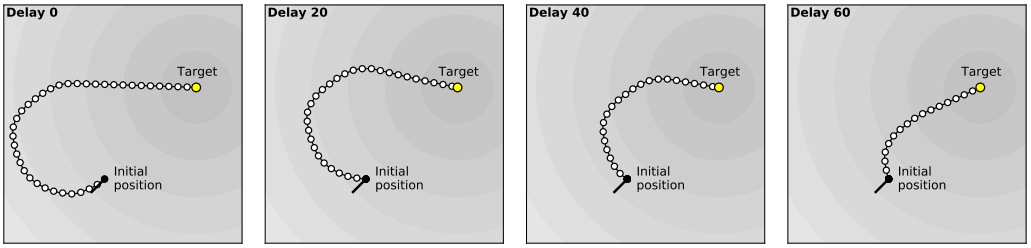


Figure 5.14: Momentum guided reinforcement learning can be understood using an analogy of velocity and acceleration. In this illustration, velocity corresponds to Hebbian learning, which represents the momentum of the trajectory, while acceleration corresponds to reinforcement learning and can modify the velocity vector up to a certain degree (10% in the figure). The illustration simulates the trajectory from a fixed initial position to a target, when the direction of the velocity is continually modified by the acceleration. In the figure, four trajectories with increasing initial delays (left to right) are shown. These delays correspond to the moment when velocity is activated/allowed to initiate the trajectory. Before that, the acceleration influences the direction of the velocity, however without any displacement. When there's no delay, the initial velocity vector can potentially drive the trajectory away from the target (as shown in this specific example), while with longer delays, the acceleration can explore and identify a more conducive direction in order to exhibit a more direct trajectory to the target.

Drawing an analogy with momentum-guided reinforcement learning, the BG pathway biases the motor output towards the target syllable using reinforcement learning, akin to the role of acceleration in the aforementioned illustration, as both play the role of modulating the direction of the system. The growth of the slower and more stable cortical pathway can be compared to the role of velocity, where both maintain a steady direction in the learning phase. They provide a central locus around which acceleration/the BG loop are able to explore locally and bias the output. The local changes made by acceleration/the BG loop are gradually incorporated and consolidated within velocity/the cortical pathway. Within this analogy, we extend the findings of the illustration to the dual pathway system. We have observed that the slowly-modulated velocity parameter can govern the direction of the system, and can thus provide the momentum required to escape local minima by countering the effect of the acceleration parameter that would preferentially guide the system towards the local minima. This overshooting of shallow local minima in an attempt to reach the global minimum could be a potential role fulfilled by the development of the cortical pathway and the transfer of information from the subcortical to cortical pathways. Moreover, in the absence of momentum (as simulated using

initiation delays), acceleration can exercise comparatively greater influence over velocity, which could, in turn, help find a more conducive direction for the system to begin its trajectory. Thus, during the early babbling stage of vocal learning, where the BG loop is formed, but the cortical pathway isn't yet fully formed, the avian BG can explore the sensorimotor space and find a desirable direction with respect to the local contour. Such early exploration would result in the system moving more rapidly in the chosen direction, eventually, when momentum kicks in or, in this case, when the HVC axons innervate the RA. Considering the cortical motor pathway to implement an equivalent of the notion of momentum, it might be necessary for this pathway to be active only after the initial direction of the gradient is determined. The alternative being a momentum, in a random initial direction, that may cause delays in learning, or ultimately drive the learning away from the actual target. Hence, the developmental delay facilitates an initial period of high exploration, which could be hindered by the presence of a comparatively change-resistant cortical pathway. The delay in the development of cortical pathway, thus, might have a crucial role in controlling the exploration-exploitation trade-off by allowing for an initial predominantly exploratory phase, while additionally assisting in escaping local minima in the later stages of sensorimotor learning. Providing this auxiliary support to gradient descent could be another potential functional role for the otherwise unaccountable delay observed in the innervation of RA neurons by HVC axons.

While this hypothesis might be difficult to test at the experimental level, it is much more feasible to simulate a comparison between a random initial momentum and a late gradient-directed momentum. Our prediction is that in the former case, learning would be deeply hindered, as illustrated in Figure 5.14, using an accelerated velocity analogy. It is to be noted that the illustration has been made in 2D and may misleadingly suggest that the initial velocity vector has a fair chance of pointing in the right direction. However, in higher dimensions, this probability would be much lower and therefore, the delay in the activation of the Hebbian-based pathway would be functionally critical [4].

The comparison of the structural plasticity within the cortical motor pathway with the technique of momentum is also in agreement with a prevalent hypothesis (discussed in section 3.3.1) that the innervation of RA by the HVC axons causes an increased inhibitory nature of the RA, which in turn diminishes the influence of the mainly NMDA receptor-based LMAN inputs to the RA (Garst-Orozco et al. 2014b; Ölveczky, Otchy, et al. 2011). The development of the cortical pathway, thus, slowly diminishes the exploratory LMAN influence and, over time renders the vocal production to be more stereotypical in nature.

The dual pathway architecture we've introduced and illustrated in the songbird is, in fact, a widespread cerebral organization among vertebrates, including birds, rodents and primates (Boraud et al. 2018b). A rough description of this architecture is that the BG pro-

[4] The scripts are available on https://github.com/rsankar9/Review-momentum-illustration and archived in a Zenodo repository https://doi.org/10.5281/zenodo.4063714 (Sankar, Rougier, et al. 2021)

vides the necessary motor variability for exploration while the cortex (pallium for some taxa) provides stability and late exploitation. Once acquired, skills are expressed solely by the motor cortex without the need for the BG. This constitutes a generic and powerful mechanism for the acquisition of sensorimotor skills that departs from modern machine learning techniques. Given the architectural, behavioral and physiological constraints we've introduced, the dual pathway model constitutes a plausible approach to sensorimotor learning that is strongly rooted in neuroscience and behavior.

6 Neural correlates of transfer of learning between cortico-striatal circuits

It has widely been hypothesised that reinforcement learning underlies sensorimotor learning. Vocal learning in songbirds provides a tractable system to investigate this hypothesis further. The avian song system consists of a direct cortical pathway, capable of controlling song production, and a BG-thalamo-cortical loop, which contributes to learning an imitation of the tutor song. The goal of this chapter is to conduct a theoretical and electrophysiological investigation into the song-related circuitry in zebra finches, and verify the feasibility of reinforcement learning within this circuitry. We being by elucidating behavioural and anatomical features of vocal learning, discussed in chapter 3, and compiling them to formulate a hypothesis to investigate the transfer of BG-led reinforcement learning to cortical motor pathways. We conduct a primary verification of the hypothesis using a computational model, and then build an experimental protocol to test the predictions that emerge from the model. We present the preliminary data collected from a pilot electrophysiology study, using this protocol, and discuss methods to analyse the data in the context of our hypothesis.

6.1 Hypothesis: Transfer from the avian BG-cortical loop to the cortical motor pathway

As discussed in chapter 3, the vocal learning behaviour of songbirds is quite similar to human speech acquisition. Juvenile songbirds learn to sing by imitating adult tutors. They initially listen to the adult vocalisations, and build a neural representation. This is followed by sensorimotor exploration, wherein young birds start to produce babbling sounds and gradually improve their vocalisation of the tutor song. Eventually, vocal exploration decreases and the produced song is more stereotyped, as we approach the crystallised phase. However, this entails an erratic process which is not easily tractable. For instance, it is not straightforward to deduce which tutor song syllable the juvenile bird is attempting to imitate at a given moment, in the early babbling phase (Kollmorgen et al. 2020; Tchernichovski et al. 2001). Thus, the trajectory of song improvement over the early

sensorimotor phase is not easily quantifiable. On the other hand, post-crystallisation, the bird produces more stereotypical and thus, tractable songs, consisting of usually the same set of syllables in a stable sequence, or 'motif'. Hence, in our study, we focus on vocal learning during the crystallised phase, in adult songbirds.

Vocal learning in songbirds is governed by a system of cortical and subcortical structures that coordinate the required muscle and respiratory activity, as discussed in chapter 3. This song system is solely responsible for producing and learning vocalisations. It consists of two parallel pathways. The cortical motor pathway includes the premotor cortical nucleus, HVC, and the robust nucleus of the arcopallium (RA), which is the primary site of motor control. This pathway is required for normal song production. In parallel, a BG-thalamo-cortical loop, called the anterior forebrain pathway (AFP), indirectly connects the HVC and RA. This consists of three nuclei connected in a loop: the BG homologue Area X, the dorsolateral thalamic nucleus (DLM), and the lateral magnocellular nucleus of the anterior nidopallium (LMAN). The pre-motor nucleus, HVC, is involved in generating the timing and sequencing of song and drives the RA activity pattern, which controls downstream muscle activity for song production. Meanwhile, similar to the cortico-basal ganglia circuitry in mammals, the AFP plays a crucial role in motor skill learning and plasticity (Boraud et al. 2018a). Within this framework of two parallel pathways, tutor signals from sub-cortical circuits aid the primary cortical pathway in motor skill acquisition.

To elucidate the putative function of the AFP further, lesions in the LMAN in juvenile birds, disrupts song development, but does not affect production of stable song patterns, in adult birds (Bottjer, Miesner, et al. 1984). This suggests that juvenile vocal learning is driven by the BG circuit, which is distinct from that which produces adult vocalisations. LMAN lesions in adult birds also reduce trial-to-trial variability in song rendition, suggesting a role for LMAN in introducing variability into the circuitry. Microstimulation of the LMAN during song induces acute, specific changes in the learned features of song indicating the ability of LMAN activity to modulate ongoing motor performance (Kao et al. 2005). Meanwhile, the Area X receives strong dopaminergic innervation from the midbrain, providing an online evaluation of song quality. Moreover, there is experimental evidence showing activity-dependent synaptic plasticity at RA and area X synapses. This suggests that all the elements required to implement reinforcement learning (exploratory variability, reward-related information and plasticity), are available within the AFP.

Several studies have hypothesised that reinforcement learning underlies sensorimotor learning, and specifically, vocal learning (M. Fee et al. 2011; Wickens et al. 2003). As discussed in chapter 3, and summarised above, the AFP seems to be an ideal candidate to implement RL within the avian song system. Further, during adult plasticity, studies have shown that the role of LMAN (output of AFP) in the expression of learned motor bias decreases over time (Andalman et al. 2009b). Several theoretical studies (Mehaffey et al. 2015; Teşileanu et al. 2017) provide additional support to this view that the transfer of learning within a

dual pathway framework can enable the BG to guide behavioural adaptations, until it is eventually consolidated into cortical networks following extensive training.

If reinforcement learning (RL) does indeed govern vocal learning in songbirds, then the LMAN (the output nucleus of the AFP and a source of variability) should be able to drive variability in vocalisations during song learning. This suggests that the trial-to-trial variability of multi-unit activity in LMAN should show a correlation with the acoustic features of resulting vocalisations, during learning. Further, within the two-step learning paradigm discussed above, the exploration conducted by the AFP would need to ultimately be consolidated into the cortical motor pathway. This would imply the existence of a neural framework that allows LMAN to influence RA activity. If the LMAN (a source of variability generation in motor exploration), drives RA (the locus of motor control in the song pathway) activity, a correlation will be found between the trial-to-trial variability in the multi-unit activity of these two nuclei, during song learning. We, therefore, seek to investigate the role of the LMAN, the output nucleus of the AFP, in modulating RA activity, during vocal learning, in this chapter. To this end, we gather information from relevant literature on adult vocal plasticity, and design an experimental protocol to test the aforementioned hypothesis.

Studies have shown that adult birds, when deafened, show a gradual deterioration in their song quality (Brainard et al. 2000). However, a similar deterioration is not observed when projections from the LMAN are lesioned, post deafening. This indicates that the AFP continues to help maintain song quality even after the song has crystallised. Furthermore, songbirds retain the ability to implement minor modifications to their song, even post-crystallisation, well into their adulthood. It has been established that when provided with auditory feedback contingent on a given syllable, the bird is able to alter the corresponding syllable in order to avoid the feedback (Tumer et al. 2007). Within this protocol, researchers use a closed-loop system to target a particular syllable of the bird's song. Whenever the bird sings, an acoustic feature of the target syllable is computed, for example, the pitch, duration, or spectral entropy. Depending on this feature, a feedback is delivered to the bird immediately. This feedback can be visual or auditory, for e.g. a playback of the bird's syllable or white noise (Zai et al. 2020). When a short pulse of white noise is used regularly to provide feedback to the bird, it attempts to evade this feedback by altering its song accordingly. For e.g., if the feedback is delivered when the bird sings a lower range of syllable pitch, the bird will attempt to escape the feedback by increasing its pitch, or more generally, shifting its pitch. This pitch shift observed in songbirds is a tractable demonstration of adult plasticity. When the AFP outputs to the RA are lesioned, adult birds lose their ability to respond similarly to such behavioural training protocols. Andalman et al. (2009a) show that the AFP influence is required to shift the pitch according to such external feedback, and this bias is consolidated within a period of a day in the cortical motor pathway. Within the pitch-modulation regime, the pitch of a target syllable can be incrementally modified by increasing the range at which feedback is provided. We harness this ability of songbirds, to observe the interplay between LMAN

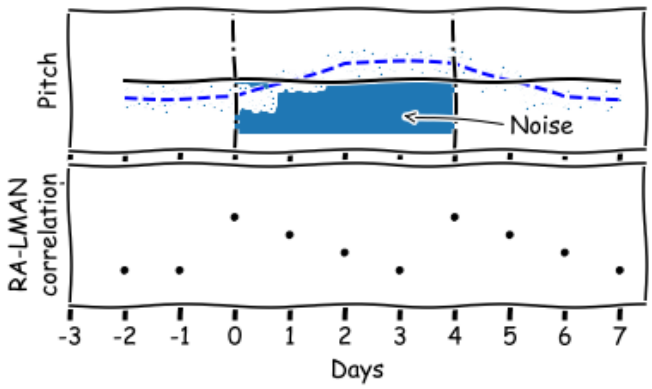


Figure 6.1: Sketch illustrating the proposed hypothesis. A protocol using conditional auditory feedback can be used to shift the pitch of a particular syllable. We hypothesise that on days where the bird is learning to either increase or decrease the pitch of a target syllable, the correlation between the trial-to-trial variability in the activity of the LMAN and the RA, in the premotor period, will increase. Correspondingly, the correlation in the trial-to-trial variability between the activity in LMAN and pitch of the produced vocalisations would also increase.

and RA activity during learning using the more controlled and tractable setting of adult plasticity, instead of juvenile learning.

We test our hypothesis by inducing adult plasticity using the aforementioned pitch shift paradigm with conditional auditory feedback. More specifically, we posit that if the LMAN does indeed drive RA, during adult plasticity, then Post-consolidation within the cortical pathway, this correlation would subside, as the LMAN would no longer need to alter the RA activity pattern. Following the end of the feedback protocol, the bird's song has seen to return to its original form. In this case as well, we expect the correlation between RA and LMAN activity would increase during the transition, but subside once the song has returned to baseline. More specific details about the experimental protocol are specified in section 6.3.

To scrutinise this hypothesis, we explore vocal learning using a computational framework resembling a simplified song system, with sigmoidal rate-coded neurons. We use this model to make quantitative predictions about the correlation between RA and LMAN activity within the protocol. Post learning, we simulate the conditional feedback protocol and test the model's response to it, in order to study the influence of the BG-cortical pathway towards adult plasticity. Based on the prediction made by the model, we, further, investigate our hypothesis using experimental approaches. We design an experimental protocol and use electrophysiology to collect preliminary data of the neural activity of

RA and LMAN in freely-behaving birds, when subjected to the distorted feedback protocol. To summarise briefly, we posit that if the LMAN (a source of variability generation in motor exploration) drives RA (the locus of motor control in the direct pathway) activity during learning, then a correlation will be found between the variability in acoustic features of a syllable, and the premotor activity in these neural regions.

6.2 Theoretical analysis of hypothesis

In this section, we attempt to gain further insights into the interplay between the BG-thalamocortical pathway (the AFP) and the cortical motor pathway during vocal learning. We build a simplified computational framework inspired by the dual pathway framework, using rate-coded neurons. We, then, simulate learning in juvenile and adult conditions, and observe the activity of the neurons in the different layers during learning. We use this model to make quantitative predictions about the correlation between the activity of LMAN and RA, during plasticity, and verify our hypothesis in a two-step learning paradigm within a dual pathway framework.

6.2.1 Methods

Architecture

The model's network comprises of three main layers: the HVC, RA and BG layer. Each layer is composed of thirty rate-coded sigmoidal neurons, each. The HVC and the RA layers are connected via two parallel pathways, with one of the pathways passing through the BG layer. The RA layer projects to downstream layers that produce vocalisations, which is represented here using an output layer, which produces a scalar value, denoting the pitch of a syllable. The weights between the HVC and RA layer, as well as those between the HVC and the BG layer are plastic. Meanwhile, the BG is connected to the RA layer, and the RA to the output layer using fixed synaptic weights. The neurons are modeled using steep sigmoidals for activation function, which bound the firing rate of the neurons between -1 and 1 (Eq 6.3).

Learning

The two pathways are governed by Hebbian learning and reinforcement learning, respectively. The BG layer (J_{BG}) functions as a source for variability (ξ), and receives feedback about performance quality (R). It uses a covariance learning rule to update its synaptic weights (W_{RL}), as shown in Eq 6.1. In parallel, the direct motor pathway (W_{HL}) between HVC (J_{HVC}) and RA (J_{RA}) is governed by a modified version of Hebbian learning, known as Oja's rule, such that co-activation of an HVC and RA neuron leads to the potentiation of their shared synapse, in a more stable manner (Eq 6.2) (Oja 1982).

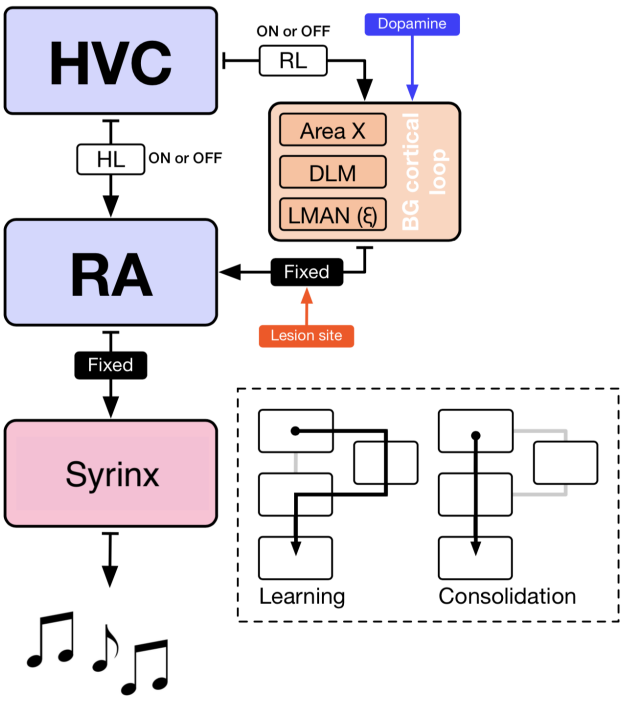


Figure 6.2: Simplified schema of the dual pathway architecture. The cortical motor pathway connects the HVC and RA and is built using Hebbian learning. The components of the parallel AFP is represented using a BG layer. The synaptic connections between HVC and BG are updated using reinforcement learning, while the connections between BG and RA are topographic and fixed. The RA output is transformed into a scalar value denoting the pitch of the desired vocalisation. Post learning, in order to test the effect of LMAN lesions, we inactivate the inputs from the BG to the RA.

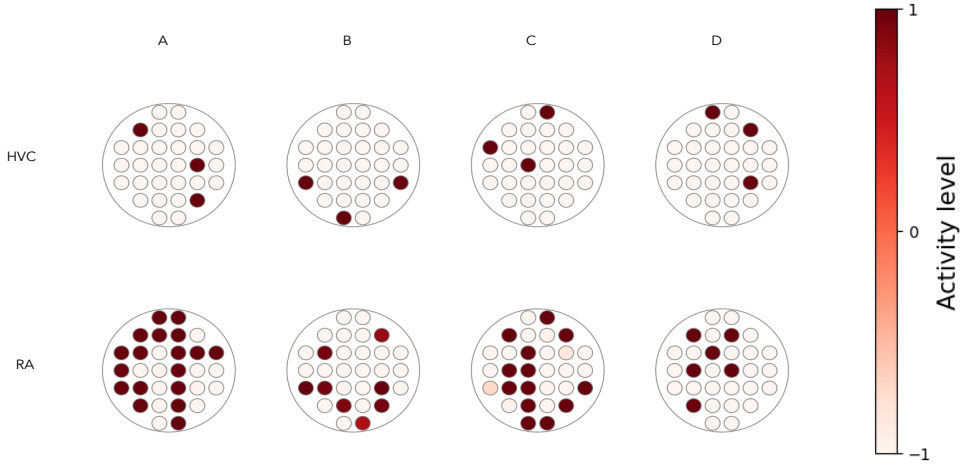


Figure 6.3: Activity pattern within the HVC and the RA layer corresponding to each syllable in the song motif. Syllables are encoded within the HVC layer, in a sparse and non-overlapping manner. This syllable encoding functions as the input to the system signalling the target syllable to be produced. The RA layer gradually learns the activity pattern corresponding to the desired pitch value for a given syllable. Hebbian learning within the HVC-RA pathway gives rise to the binary pattern of firing rates observed here in the RA, signalling bursts of spiking activity within the RA.

$$\Delta W_{RL} = \eta_{RL}.J_{HVC}.J_{BG}.(R - \bar{R}).\xi \tag{6.1}$$

$$\Delta W_{HL} = \eta_{HL}.J_{RA}(J_{HVC} - J_{RA}.W_{HL}) \tag{6.2}$$

$$sig(x) = \frac{1}{1 + e^{-1*(x-sm)*ss}} \tag{6.3}$$

where η refers to the learning rate of a pathway, such that $\eta_{HL} = 0.01 * \eta_{RL}$.

Encoding

The model has to learn a song motif of four sequential syllables (A, B, C, D) using the above-described dual pathway framework. Each syllable is associated with a scalar value denoting the pitch of the syllable. The HVC encodes a discrete sparse representation for each syllable in a song, and drives RA activity. The model must explore and find the right RA activation pattern, to produce the desired song motif, as shown in Figure 6.3. The HVC encoding used here is reminiscent of the findings of Hahnloser, Kozhevnikov, et al. (2002).

Plasticity protocol in the model

Once the target motif has been learnt, we proceed to test the influence of the BG-cortical pathway in song production. We simulate lesioning of LMAN, by inactivating the inputs from the BG pathway to the RA. We observe the effect of such lesioning of the BG pathway on the variability of the pitch distribution generated by the model.

Post learning the target song motif, as a simplified analogy of juvenile song learning, we study the role of the BG-cortical pathway in adult sensorimotor learning. We induce adult plasticity, by simulating the conditioned auditory feedback protocol, described above, and observing the model's response to it. Under this protocol in songbirds, even a single specific syllable of the song can be targeted to induce plasticity. Once a song motif has been learnt by the model, we specifically distort the perception of the lower range of the pitch distribution. We do this by artificially increasing the performance error associated with a particular syllable, when the pitch produced by the model lies in the lower range of its distribution, as illustrated in Figure 6.5. We observe the change in BG and RA activity in response to this distorted feedback protocol.

We take a deeper look at the change in BG and RA activity in response to adult plasticity, by simulating an artificially exaggerated scenario. Here, instead of inducing a minor change in the syllable pitch, we task the model with learning a new target pitch for the corresponding syllable. We then observe the relationship between the BG and RA activity.

6.2.2 Results

The model described here provides a signal to produce a syllable, using the HVC as an input layer. The HVC uses a sparse non-overlapping representation for each syllable in a song. As shown in Figure 6.4, the model learns to produce a song motif as a sequence of four syllables, where each syllable is represented by a randomly chosen target pitch value. Initially, similar to the babbling stage, the generated pitch distribution is highly variable, and there is no sequence structure. As the model approaches the crystallised stage, the variability in the pitch distribution of each syllable reduces and the renditions increasingly resemble the target song. At this stage, we simulate lesioning the LMAN outputs, by inactivating the influence of the BG pathway. This leads to the song continuing to be produced in a stable manner, albeit with reduced variability.

Now, we further investigate the influence of the BG-cortical pathway towards plasticity. To do this we emulate a common experimental protocol of distorted auditory feedback, generally used to induce adult plasticity in birds. As shown in Figure 6.5, after the song has crystallised, we selectively penalise the production of lower half of the pitch distribution for syllable C. This results in an upward shift in the pitch distribution of syllable C generated by the model.

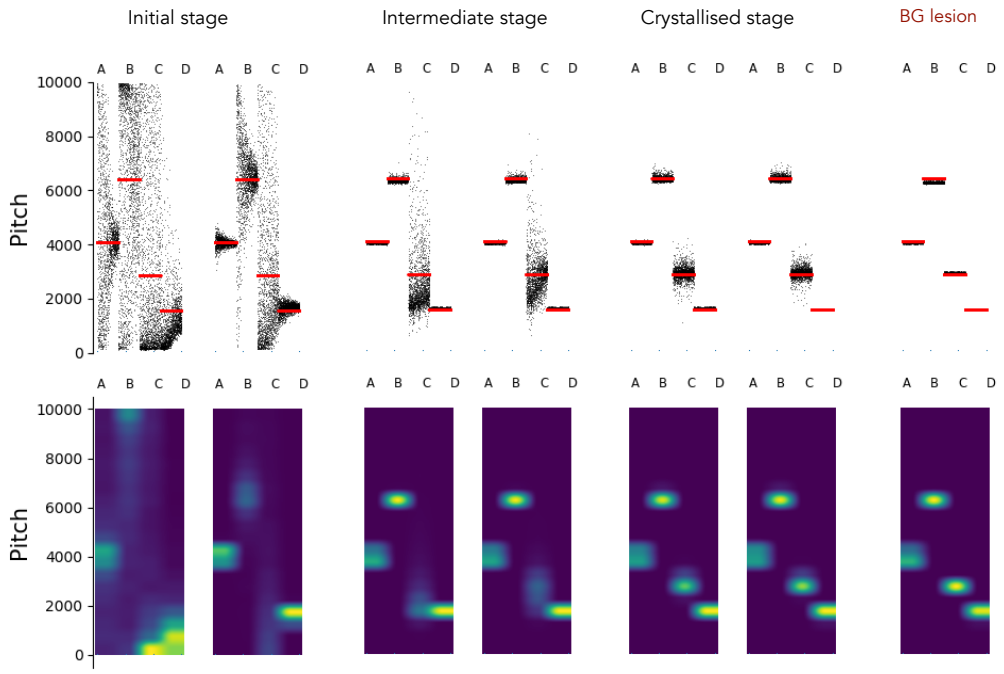


Figure 6.4: Effect of lesion post vocal learning. The top panel groups multiple renditions within each stage of vocal learning. The red horizontal bar denotes the target pitch for each syllable. The black dots denote each production of syllable pitch. The bottom panel shows the spectrogram corresponding to song in each stage. In the initial stage, the vocalisations are highly variable. As training proceeds, the vocalisations are more stereotyped for each syllable. Post lesion of BG outputs, the variability reduces further, as shown on the left vertical panel.

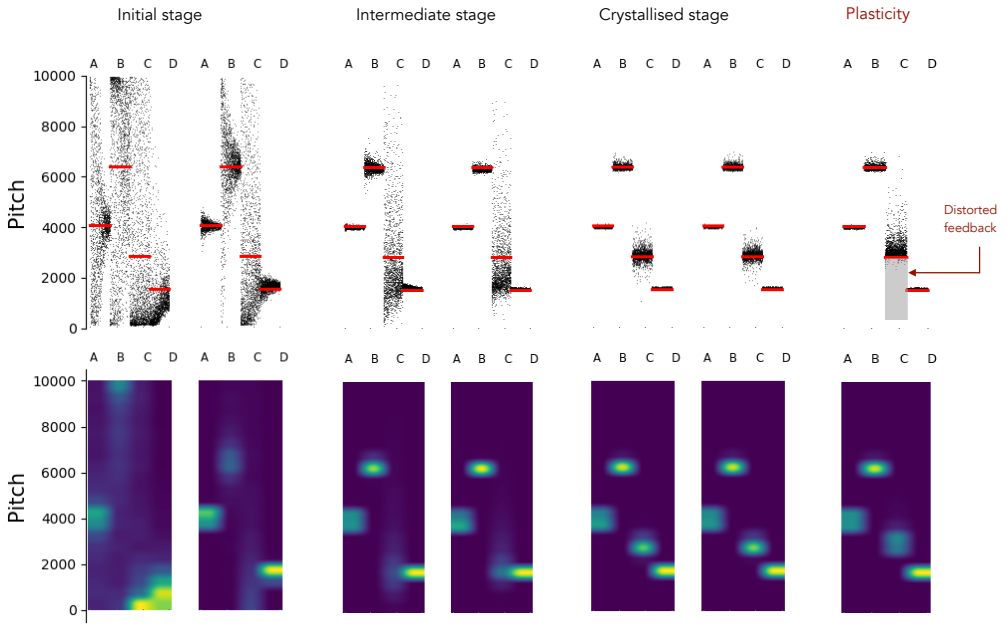


Figure 6.5: Effect of the CAF protocol post song crystallisation. The top panel groups multiple renditions within each stage of vocal learning. The red horizontal bar denotes the target pitch for each syllable. The black dots denote each production of syllable pitch. The grey shaded region shows the presence of distorted feedback. The bottom panel shows the spectrogram corresponding to song in each stage. In response to the CAF protocol on syllable C, the pitch distribution of the target syllable undergoes an upward shift.

To explore the neural activity during this transition more closely, we simulate plasticity in a more explicit scenario by exaggerating this change in motor action. We artificially alter the target template for the syllable pitch, and observe the BG and RA activity in response to this modified reward profile. Note that songbirds are not capable of such a drastic change post crystallisation. In Figure 6.6, we take a closer look at syllable C in isolation. When the target pitch for syllable C is altered, there is a drop in performance evaluation. This leads to a renewed exploration of the motor space by the BG-cortical pathway. This is reflected in the increased variability of the pitch distribution. As training progresses, the performance improves and variability decreases. We look into the influence of the BG pathway during this learning process, by tracing the correlation between the LMAN and RA activity. We observe that the correlation between the activity of these two regions increases in the exploration phase when a new target has to be learnt and decreases as the the song is controlled more and more by the cortical pathway.

6.2.3 Discussion

It has been shown that online perturbations of auditory feedback of the pitch generates adaptive changes in the pitch of syllable. The response of the model to distorted feedback is similar to songbirds where within a day, the bird attempts to evade the distorted auditory feedback by preferentially producing pitches from the non-distorted range of its pitch distribution (Andalman et al. 2009a). This could be by gradually raising or lowering its pitch, depending on the direction of the distorted feedback. The fast timescale of learning within the AFP can support such rapid adaptation, while the slower learning mechanisms within the cortical pathway can consolidate information on a longer timescale. This mechanism is akin to the one hypothesised by several researchers, including Boraud et al. (2018a) and Hélie et al. (2015), where fast RL-based BG trial-and-error learning facilitates slower Hebbian learning mechanisms in the cortex for the storage and expression of automatic skills.

In Figure 6.6, we see that the correlation between the activity of the LMAN and the RA increases in the exploration phase, and decreases as the the song is controlled more and more by the cortical pathway. This serves as an indication that the LMAN drives neural activity in the RA, during sensorimotor exploration, but has diminished influence post the consolidation in the cortical pathway. Based on the model prediction, we then hypothesise that when subjected to such a conditioned auditory feedback protocol, the shift in pitch of a syllable will possibly be accompanied by an initial increase in the correlation between the variability in LMAN and RA activity from trial-to-trial. The model also predicts that such an increase in correlation would diminishe, as the induced change is consolidated within the cortical pathway.

In this case, we then hypothesise that when adult songbirds are subjected to such a distorted auditory feedback protocol, the shift in pitch will possibly be accompanied by an initial increase in correlation between LMAN and RA variability. This would be followed

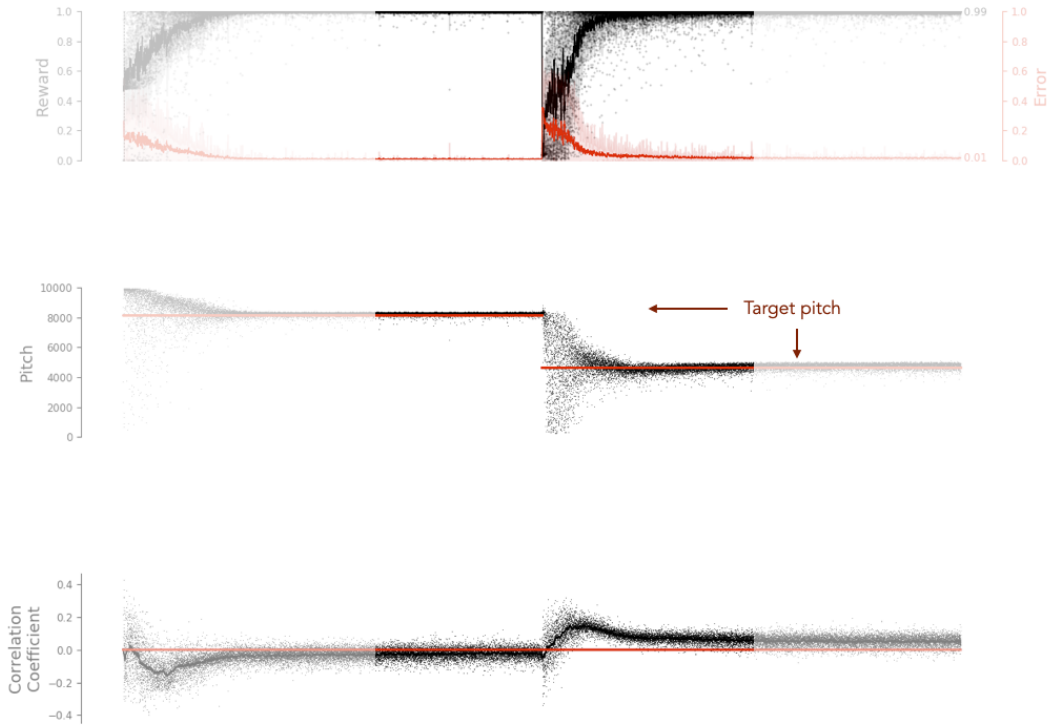


Figure 6.6: Correlation between RA and BG layer during learning. The top panel shows the progression of the produced error (red) in pitch and perceived reward (black). The middle panel shows the target pitch for syllable C in red and the generated pitch distribution of syllable C in black. The bottom panel shows the change in correlation coefficient between the activity of the RA and BG layers during the simulation. When the target for syllable C is changed, there is an increase in the correlation coefficient between activity of the RA and BG layers.

by a corresponding decrease in correlation, as the shift in pitch gradually gets consolidated in the cortical pathway.

6.3 Interaction between LMAN and RA during learning

In this chapter, we presented a preliminary theoretical study about the role of the dual pathway architecture in sensorimotor learning, in the avian context. We hypothesise that reinforcement learning based sensorimotor learning within a dual pathway system, as seen in the song system, would require the exploration of the RL pathway to be consolidated within the motor pathway, implying the LMAN (a source of variability generation in motor exploration) would drive the RA (the locus of motor control in the direct pathway) activity during learning. If this is the case, then, a correlation would be found between the variability in acoustic features of a syllable, and the premotor activity in these neural regions.

We proceed to use experimental approaches to investigate the prediction emerging from the model, described in the previous section. We conduct behavioural tests to study the response of a songbird to distorted feedback of vocalisations, contingent on syllable pitch, and observe the shift in pitch, described earlier. We, then, use electrophysiology to record the neural activity of RA and LMAN in freely-behaving birds, when subjected to this distorted feedback protocol. Here, we primarily design an experiment in order to test the prediction made by the models in the previous two sections, and conduct a pilot study using the designed protocol. Thus, within the avian context, we test the hypothesis that if the LMAN drives RA activity during exploration, then a direct correlation will be seen in the variability in acoustic features of a syllable, with the premotor activity in these neural regions.

6.3.1 Methods

Bird care

Adult (>100dph) Zebra finches were bred in an open-air aviary and housed with their parents until more than 90dph. Post 100dph, birds were isolated and housed in sound attenuating chambers with food and water provided ad libitum. Birds were housed in a natural photoperiod, in both the aviary and sound-attenuating chambers. Here, they were tested for suitability for experiments, as per the following criteria: i) weight ii) readiness to sing iii) no. of syllables in song iv) presence of a high pitch syllable v) ability to shift pitch during a CAF protocol. Birds which satisfied these criteria were selected as candidate birds. A micro-drive was surgically placed on the skull of candidate birds (M. S. Fee and Leonardo 2001). After micro-drive placement, birds were isolated and housed in sound-attenuating

chambers. All recordings presented in this manuscript are from undirected singing. All procedures were in accordance with the animal care protocols approved by University of Bordeaux.

Micro-drive

A lightweight motorised micro-drive was built and used for electrophysiological recordings from the bird (M. S. Fee and Leonardo 2001). Each micro-drive contained a customised array of three high-impedance ($10 - 20M\Omega$) electrodes and 3 silver wires (for EEG, ground and reference). The micro-drive was built in the lab by attaching electrodes to a motor, which allowed for controlling the depth of the electrodes inside the brain. Three tubes were placed on the drive, corresponding to three neural regions, the RA, the area X and LMAN, according to Table 6.1. Thus, the former tube was placed on the opposite side of the motor of the two latter tube, with a distance of 5.2mm between them. This corresponds to the average distance between the RA and LMAN on the anterior-posterior axis at $40°$ implantation angle, according to our preliminary studies with anatomical tracers. Electrodes were placed within the tubes such that the relative distances from the exit of the tubes were according to the relative depths between their corresponding neural regions (accounting for the curvature of the skull between the anterior and posterior electrodes). The electrodes were further connected using silver wires to a PCB which could transmit the signal to the recording apparatus by Neuralynx.

Surgery

Before the surgery, birds were food-deprived for 30 minutes. An analgesic (meloxicam, 5mg/kg) was administered at the start of the procedure. Birds were anaesthesised (induction with 4% isoflurane and maintained with 0.5-1.0% isoflurane) while placed on the stereotaxic apparatus at a head angle of $40°$. Local anaesthetic (lidocaine) was applied under the skin before opening the scalp (incision site shaved and cleaned with betadine), and re-applied on the open wound every 20 minutes. Small craniotomies was performed above the midline reference point, the bifurcation of the midsagittal sinus and the above the structures of interest. Stereotaxic coordinates in the anteroposterior and mediolateral axis were determind as per the sinus junction. The stereotaxic coodinates used for each site has been specified in Table 6.1. Coordinates were determined using anatomical tracers, in advance. The previously described microdrive was positioned stereotaxically over RA, LMAN and X in one hemisphere and secured to the skull using dental cement. An EEG wire was placed above the dura posterior to LMAN. The silver wires corresponding to reference and ground were placed in the cerebellum. Birds were sub-cutaneously injected with glucose prior to the surgery, and fed glucose orally for two days post surgery. Birds were injected (intra-muscular) with meloxicam for two days post surgery. After surgery, the birds recovered and resumed singing within 5-10 days.

Structure	Antero-post (mm)	Medio-lateral (mm)	Depth (mm)
LMAN	4.5	1.8	2.3
Area X	4.5	1.5	3.3
RA	-0.7	2.5	2.1

Table 6.1: Stereotaxic coordinates summary. Anteroposterior and medio-lateral coordinates are expressed in millimeters from the sinus junction, and depth coordinates in millimeters from the surface of the brain.

Electrophysiological data collection

After singing resumed, each day the electrodes were lowered into the brain and extracted back at the end of the day, for a period of ten days. Recordings were made in the Area X, LMAN and RA, as shown in Figure 6.7. Acquisition of the signal was done using apparatus built by Neuralynx, at a sampling frequency of 32000 Hz. Putative RA recording sites were identified by the presence of characteristic changes in activity associated with production of songs and post-singing inhibition. Putative LMAN recording sites were identified by increase in baseline activity during singing and the presence of spontaneous bursts during and outside singing. Putative Area X recording sites were identified by the presence of tonically active neurons. Multi-unit activity at each site was recorded until 100+ songs were produced by the bird. At the end of the experiment protocol, electrolytic lesions will be made at one recording site, each, for RA and LMAN. Recording sites will be confirmed using post-hoc histological confirmation of the trajectory of each electrode array (coated with a fluorescent tracer), as well as the location of the electrolytic lesion.

Conditional auditory feedback protocol

For the first day of recording, the bird was allowed to sing without any interventions. Post the baseline day, for five days, distorted auditory feedback (white noise of 100ms duration) was provided using external speakers, contingent on the pitch of the chosen target syllable. The amplitude of white noise was set manually to be higher than the amplitude of the bird's song. For the first three days (post baseline), the threshold for auditory feedback was adaptively increased or decreased. The new threshold was set to be the 60th percentile of the previous 300 songs, if the protocol intended to induce an upward shift in pitch, and 40th percentile, if a downward pitch shift was intended. On days 4-5 post baseline, the threshold was maintained at the mean of the last 300 songs of day 3 post baseline. After 5 days of CAF, no more feedback was provided on the subsequent days

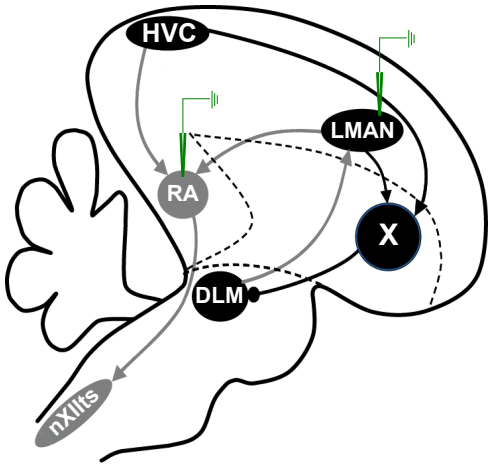


Figure 6.7: Recording sites for electrophysiology. A micro-drive was placed on the skull of a bird which contained a customised array of three high-impedance ($10-20M\Omega$) electrodes and 3 silver wires (for EEG, ground and reference). The micro-drive was built in the lab by attaching electrodes to a motor, which allowed for controlling the depth of the electrodes inside the brain. Three electrodes were placed on the drive, corresponding to three neural regions, the RA, the area X and LMAN, according to Table 6.1. Simultaneous recordings were made from the RA and LMAN sites.

until the bird re-learnt to produce the syllable with the original pitch distribution. The CAF protocol was implemented using a closed-loop system run by a customised version of the RTXI software (Skocik et al. 2013).

The recordings were classified into three categories according to the CAF protocol the bird was subjected to. The recordings from days where feedback was provided with a gradually increasing threshold (day 1-3 post baseline), inducing an upward shift in the distribution of pitches of the target syllable produced by the bird, were included in the category 'learning'. The recordings from the days immediately after feedback was stopped (after day 5 post baseline), leading to the bird gradually regaining its original pitch distribution of the target syllable, i.e. extinction, were also included in the category 'learning'. The recordings from the days where CAF was provided below a fixed pitch threshold (day4-5 post baseline) were included in the 'maintenance' category. The recordings from days were no feedback was provided, and the bird produced its original distribution of pitches of the target syllable were classified as 'baseline'. Before the commencement of the CAF protocol, the first day of recording with no feedback (day 0) was considered baseline. Post the end of the CAF protocol, the baseline day is determined by observing the pitch distribution. When two consecutive days show no significant difference as per the Kruskal Wallis test, we consider the vocalisations of the target syllable to be outside of the learning protocol, and is included in the 'baseline' category.

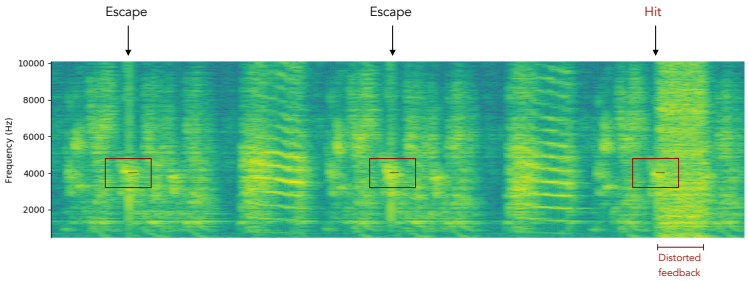


Figure 6.8: Instance of the CAF protocol. The fundamental frequency of a high pitched syllable is computed. As shown in the first two song motifs, if the calculated pitch is in the top forty percentile of the pitch distribution of the target syllable, no feedback is provided. As shown in the third song motif, if the calculated pitch is in the lower sixty percentile of the pitch distribution of the target syllable, a distorted auditory feedback of 100ms is provided.

Acoustic feature

For each syllable, we defined a measurement time relative to the syllable onset that corresponds to a well-defined spectral feature, at 15ms post syllable onset. Syllable onsets were defined based on amplitude threshold crossings. Pitch is defined as the fundamental frequency of the signal. Pitch was computed by combining several FFTs over a 25ms sample (ending at the measurement time) of the target syllable. Each FFT was performed with a decreasing number of sample points, such that the resulting cumulative FFT has a finer resolution. The pitch is determined by the frequency of the FFT peak within a manually pre-determined frequency range. To observe the effect of the CAF protocol, the distribution of all vocalisations of the target syllable was plotted across all days of the protocol.

Data analysis - spiking statistics

Units were sorted manually with the Spike2 (CED, UK) software, using principal component analysis of spike waveforms. A unit was considered for analysis only if the average signal-to-noise ratio was higher than 1 during baseline and if the spike shape looked stable and salient, based on visual inspection. A unit was classified as multi-unit or single-unit depending upon the distribution of its inter-spike intervals. From here on, unless specified, the term unit refers to a multi-unit cluster. For further analysis, units with low average firing rate under the time-window pertaining to the corresponding analysis, were discarded. The peri-stimulus time histograms (PSTH) were calculated for all recorded units, with respect to syllable durations.

Pre-motor neural activity

In this chapter, neural activity of a unit in a specified time window denotes the mean detrended firing rate. The instantaneous firing rate (IFR) was calculated for the spike train for each unit. The IFR was used to detrend the signal. The number of spikes in the desired time window (pre-motor, in this case) was retrieved by calculating the area under the curve. This results in the mean detrended firing rate over the specified time window. It is determined by building the instantaneous firing rate of the unit over the entire recording, detrending the resulting signal and calculating the area under the curve for the specified time window. Premotor neural activity denotes the mean detrended firing rate within the "premotor window" corresponding to a duration before the measurement time of acoustic features of a syllable. This "premotor window" for all RA units was chosen to be 40ms to reflect the latency at which RA activity influences the song structure, determined based on previous studies (Sober, Wohlgemuth, et al. 2008). The premotor window for all LMAN units was chosen to be 5ms before the premotor window for RA units, to account for transmission lag from LMAN to RA (White et al. 1999).

Correlation analyses

First, for each RA and LMAN unit, we computed the linear correlation between average firing rate in the premotor window for each syllable and the corresponding acoustic feature (pitch), to investigate the relationship between neural activity and behaviour. Before computing correlations, we discarded outlier syllable renditions with acoustic feature measured at greater than 4 times the standard deviation from the mean feature across all renditions in a recording session. Any unit with <1 spike on average within the premotor window was discarded.

Second, for each simultaneously recorded RA-LMAN multi units, we computed the linear correlation between the trial-to-trial variance in the average firing rate in premotor window for each syllable. This helps us investigate the relationship between the neural activity of these two regions, during vocalisation.

Third, we account for the possibility that the LMAN may be influencing RA neural activity on a longer timescale than the pre-motor window. To look into this, we consider the a larger time window, from motif onset until the measurement point of each syllable (henceforth referred to as motif onset window). Now, for each simultaneously recorded RA-LMAN multi units, we computed the linear correlation between the trial-to-trial variance in the average firing rate in the motif onset window for each syllable.

Further, we performed a bootstrap analysis for each condition, as explained in Sober, Wohlgemuth, et al. (2008), to build a distribution of the number of spurious significant correlations that can arise from our dataset by chance.

6.3.2 Preliminary analysis from pilot study

Here, we present the data from a pilot study on one male zebra finch using the above experimental protocol. We induce adult plasticity in the adult bird by implementing the conditional auditory feedback protocol explained above. Vocalisations with a pitch from the lower 60 percentile of the distribution were penalised with a distorted auditory feedback. In response to this training, we observe an upward shift of 4.43% over three days, as compared to the baseline day, where no feedback was provided (Figure 6.9). Post training, when the conditional feedback was stopped, the vocalisations gradually regained their initial pitch distribution in four days (shown as the extinction phase (Kruskal-wallis test between first and last day of protocol: K=4.75, p>0.01) in Figure 6.9). The Kruskal wallis test over the last two days show that they the pitch samples could be drawn from the same distribution (Kruskal-wallis test between the last two days of protocol: K=4.75, p>0.01). Thus, we infer that the pitch of the target syllable has returned to baseline productions, and conclude the protocol on day 10. From day 0 to day 3, there is an average increase of 57.8Hz each day, with an average standard deviation of 193.8Hz each day. [1].

Simultaneously, electrophysiological recordings were obtained from the RA and LMAN of the zebra finch. As seen in a sample neuron in Figure 6.10, the recordings confirms that units in the RA show a clear pattern of activation across song renditions, followed by a post-song inhibition. 49 putative RA multi-units displayed tonic activity out of the singing context, with a mean firing rate of 32 Hz during rest. The recordings were classified within 'baseline', 'maintenance' and 'learning' categories, as explained in the previous section.

As shown in Fig 6.10 for a sample RA multi-unit, we observed a bimodal distribution of the ISIs of the RA unit during singing (corresponding to intra- and inter-burst firing rates). The PSTH for the multi-unit shows modulation of neural activity both during motifs, unaffected by the distorted feedback, and during motifs, which received distorted feedback. Here, the syllable 'b' was targeted to induce an upward pitch shift using the CAF protocol. The PSTH confirms that this RA unit exhibits song-timing locked bursts during singing, followed by an inhibition of their resting tonic activity.

32 putative LMAN multi-units showed spontaneous bursts, with increased baseline activity during singing. As shown in Fig 6.11 for a sample LMAN multi-unit, we observed a unimodal distribution of the ISIs during singing. The PSTH for the multi-unit shows modulation of neural activity both during motifs, unaffected by the distorted feedback, and during motifs, which received distorted feedback. The modulation of LMAN neural activity during song is less pronounced than RA multi units.

[1]Note: The bird was trained on the CAF protocol twice with a gap of 4 days between the two sessions. Owing to the dearth of simultaneously recorded RA and LMAN unit pairs, we discard the first 4 days of the first session. The second session could not be continued to the end due to the drive being dislodged by the bird. Therefore, in this section, we present the two sessions of the CAF protocol.

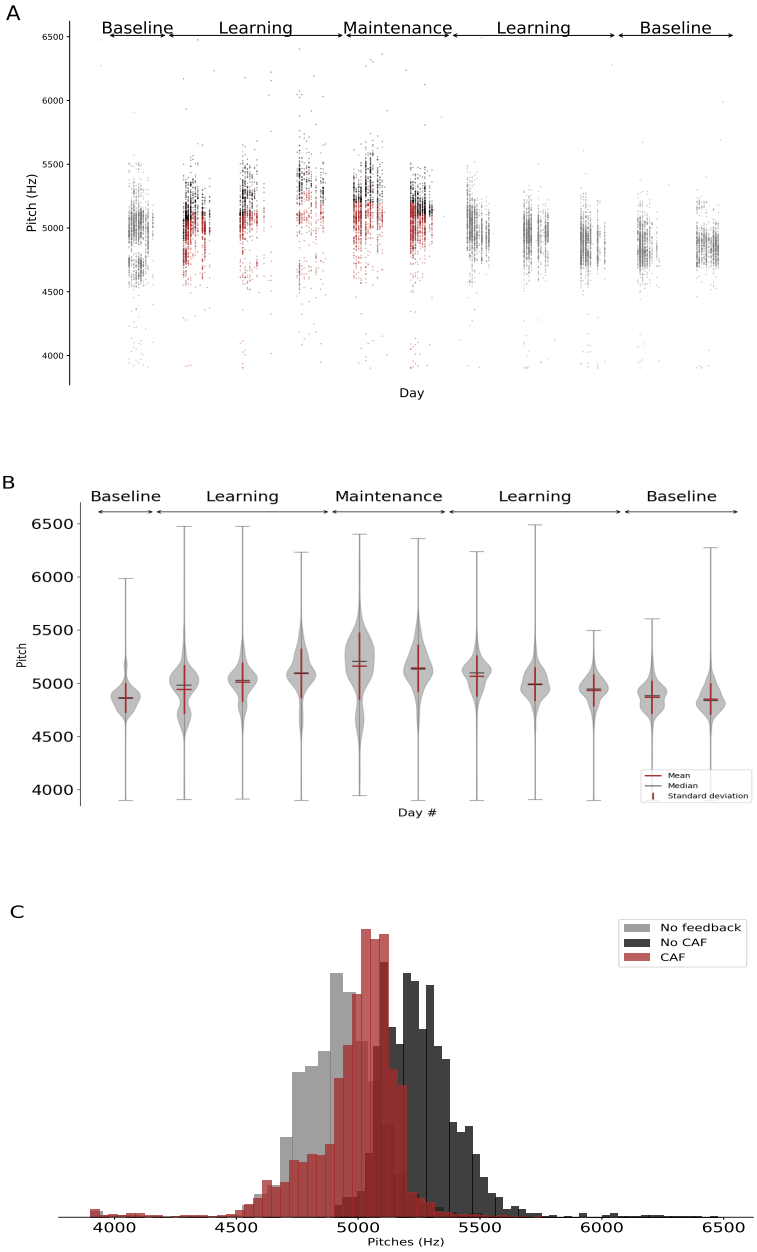


Figure 6.9: Effect of CAF protocol on behavior. Panel A shows the change in the pitch distribution of the target syllable during the CAF protocol. The renditions receiving distorted feedback within the CAF protocol are shown in red. The renditions receiving no distorted feedback within the CAF protocol are shown in black. The renditions in absence of the CAF protocol are in grey. Panel B shows the mean and median of the pitch distribution across days. Panel C shows the shifted distribution of pitches on days where CAF is present (red+black) versus the days where no CAF is provided (grey).

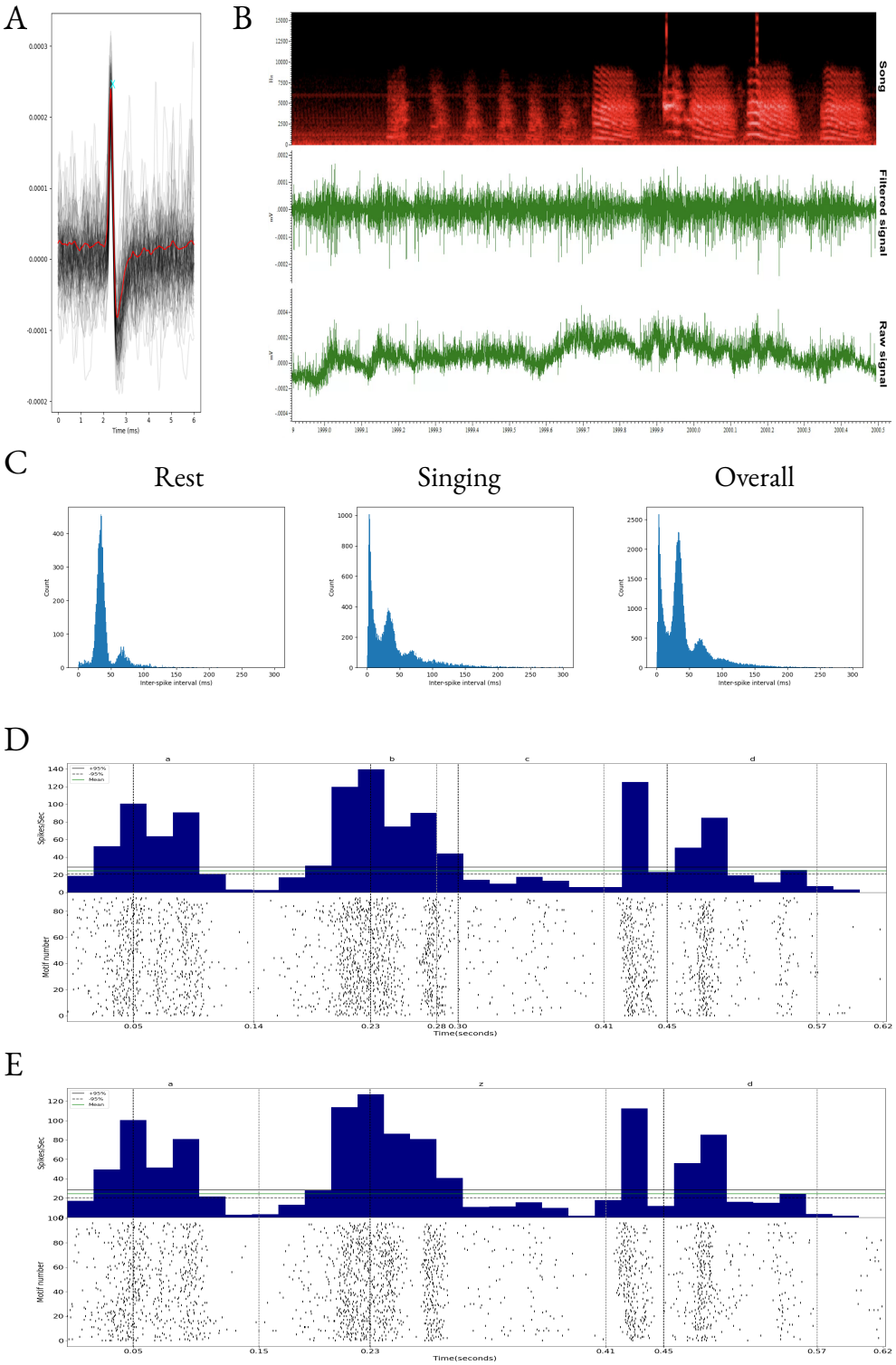


Figure 6.10: A sample RA unit. A. Superimposed spike shapes during singing (black) along with the mean spike shape (red). B. Raw and filtered signal of the sample RA unit over 1s during singing. C. The distribution of ISIs during rest, singing and over the entire recording. D. PSTH during motifs with no distorted feedback. E. PSTH during motifs with distorted feedback. Note: The distorted feedback ('z') begins during the target syllable 'b' and lasts until the next syllable 'c'.

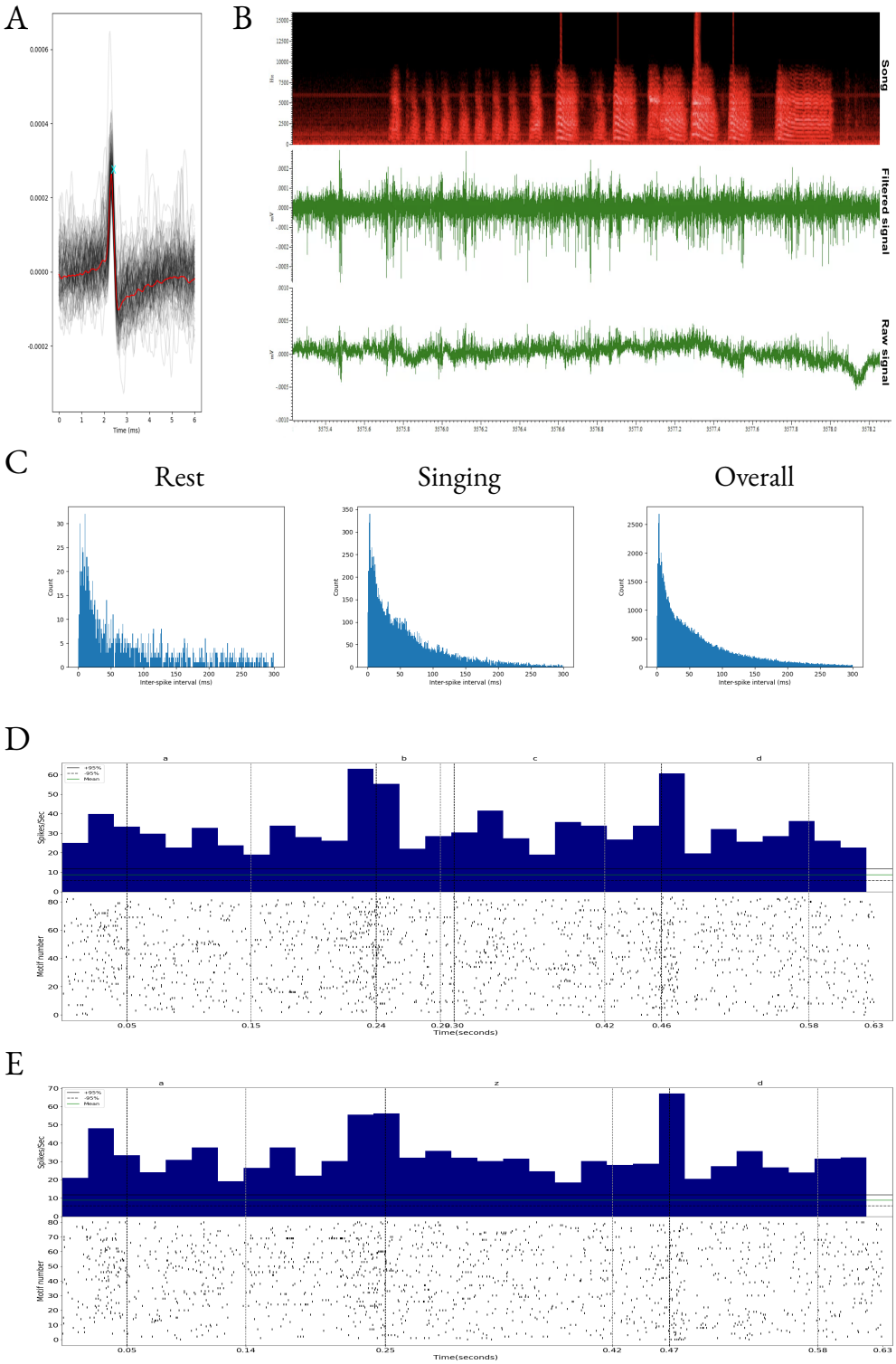


Figure 6.11: A sample LMAN unit. A. Superimposed spike shapes during singing (black) along with the mean spike shape (red). B. Raw and filtered signal of the sample LMAN unit over 1s during singing. C. The distribution of ISIs during rest, singing and over the entire recording. D. PSTH during motifs with no distorted feedback. E. PSTH during motifs with distorted feedback. Note: The distorted feedback ('z') begins during the target syllable 'b' and lasts until the next syllable 'c'.

Correlation between unit and behaviour

The main question we aim to study here is if the LMAN can drive RA activity such that it has an effect on behaviour. To look into this, we first look at the effect of LMAN and RA neurons on behaviour. We verify if a portion of behavioural variation can be explained by the trial-to-trial variation in LMAN and RA activity. We look into this by measuring the correlation in trial-to-trial variation between pitch and premotor neural activity of LMAN and RA. Owing to the song-modulated activity of neurons, not all units were active in the premotor period of all syllables. When comparing neural variability and behavioural variability, we discarded cases where the multi-unit was not sufficiently active in the premotor period of any syllable. For each recorded RA and LMAN unit, we compute the premotor spiking activity and the acoustic feature (pitch) of the corresponding syllable. Figure 6.12 shows a sample RA unit whose premotor neural activity was significantly correlated with the pitch of syllables 'a', 'b', and 'c', along with a sample LMAN unit whose premotor neural activity was significantly correlated with the pitch of syllable 'b'.

Table 6.2 shows the number of RA and LMAN multi-units whose premotor activity was significantly correlated with the pitch of the subsequent syllable in each category of the protocol, as compared to the total number of units recorded within the category. Bootstrapping shows that the number of significant correlations between the premotor activity of an RA unit and the pitch of syllable 'a', 'c' or 'd', in the three categories, is more than as expected by chance. Bootstrapping shows that the number of significant correlations between the premotor activity of an RA unit and the pitch of syllable 'b', in the 'learning' category, is more than as expected by chance. However, this is not the case for the 'baseline' and 'maintenance' categories. For the LMAN units, the number of significant correlations between their premotor activity and the subsequent syllable was not more than as expected by chance. 18% of RA units showed a signification correlation between their premotor activity and the pitch of the subsequent syllable.

Correlation between LMAN and RA units

We continue to investigate if the LMAN can drive RA activity such that it has an effect on behaviour. After seeing the relationship between individual LMAN and RA units and behaviour, we look into the relationship between LMAN and RA units in a given context. To see if LMAN can drive RA activity during learning, we measure the correlation between premotor neural activity of LMAN and RA. We consider the premotor period of an LMAN unit to be 5s before that of a RA unit, to account for a transmission delay from LMAN to RA. For each recorded RA and LMAN unit, we compute the premotor spiking activity corresponding to each syllable. Figure 6.13 shows a sample RA-LMAN unit pair whose premotor neural activity for syllable 'b' was significantly correlated with each other. Table 6.3 shows the number of pairs of RA-LMAN multiunits showing a sig-

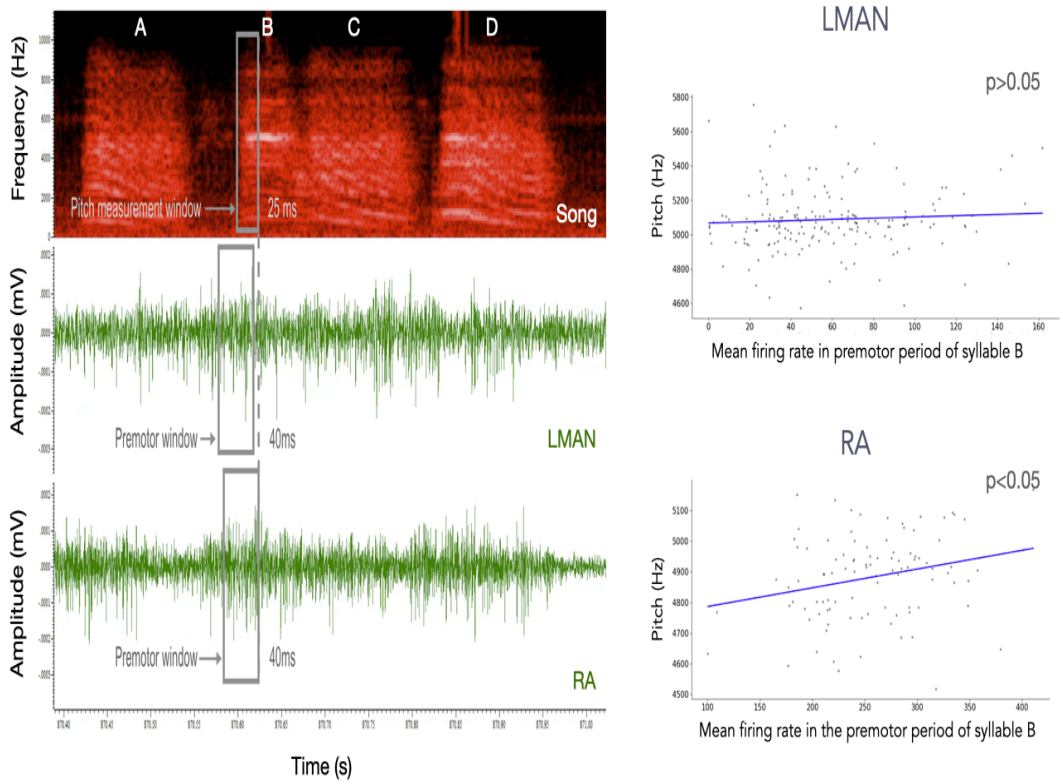


Figure 6.12: Sample correlation plot between the trial-to-trial variability in the premotor neural activity (MDFR) of a unit and the pitch of the corresponding syllable (grey). Left: Sample simultaneous recording of an RA and LMAN multi-unit during singing. The premotor windows with respect to the pitch measurement point of the target syllable is depicted in grey. The time window for LMAN is displaced with respect to that of RA to account for transmission delays. Right: Sample correlation plot between the trial-to-trial variability in pitch of a syllable and the neural activity (MDFR) of an RA and LMAN multi-unit, respectively, within the corresponding premotor window. Here, the LMAN unit does not show a significant correlation ($p > 0.05$) in its activity in the premotor window for a syllable, but the RA unit shows a significant correlation ($p < 0.05$) in its activity in the motif onset window for the syllable. The blue line denotes the linear regression across the data points. MDFR: Mean detrended firing rate.

	RA		LMAN		
	ACD	B	ACD	B	
Baseline	23	8	17	7	Total
	9	0	1	1	Significant
Learning	80	28	42	14	Total
	12	5	1	2	Significant
Maintenance	33	11	16	6	Total
	7	0	0	0	Significant

Table 6.2: The table shows the number of significant correlations between the variability in the premotor activity of an RA or LMAN multi-unit and the pitch of the subsequent syllable. The syllables are divided into two groups, the target group with syllable 'b' and the control group with syllable 'a', 'c' and 'd'. For each category, we display the number of significantly correlated unit-behaviour pairs vs the total number of pairs recorded. The color denotes the result of bootstrapping. 'Green' denotes that the number of significant correlations in the dataset is not as expected by chance. 'Red' denotes that the number of significant correlations in the dataset is within chance levels.

	Premotor window		Motif onset window		
	ACD	B	ACD	B	
Baseline	23	7	40	16	Total
	1	0	3	1	Significant
Learning	51	19	111	42	Total
	2	0	19	6	Significant
Maintenance	27	11	37	15	Total
	1	2	2	3	Significant

Table 6.3: The table shows the number of significant correlations between the variability in the activity of an RA and LMAN multi-unit in the premotor window and motif onset window. The syllables are divided into two groups, the target group with syllable 'b' and the control group with syllables 'a', 'c' and 'd'. For each category, we display the number of significantly correlated RA-LMAN unit pairs vs the total number of pairs recorded. The color denotes the result of bootstrapping. 'Green' denotes that the number of significant correlations in the dataset is not as expected by chance. 'Red' denotes that the number of significant correlations in the dataset is within chance levels.

nificant correlation between their premotor neural activity corresponding to a syllable, as compared to the total number of simultaneously recorded pairs which were active in the premotor period of the corresponding syllable. Bootstrapping shows that the number of significant correlations obtained in this dataset is not outside chance levels.

To account for a wider range of timescale in which the LMAN may potentially drive RA, we compute the correlation between LMAN and RA spiking activity within the motif onset window, as described in the previous section. Figure 6.13 shows a sample RA-LMAN unit pair whose neural activity for syllable 'b' and 'd' in the corresponding motif onset window was significantly correlated with each other. Table 6.3 shows the number of pairs of RA-LMAN multiunits showing a significant correlation between their neural activity in the motif onset window of a syllable, as compared to the total number of simultaneously recorded pairs. Bootstrapping shows that the number of RA-LMAN unit pairs whose neural activity were significantly correlated to each other within the motif onset window of either syllable group was higher than expected by chance, within the 'learning' category. This was not the case for the 'maintenance' and 'baseline' categories.

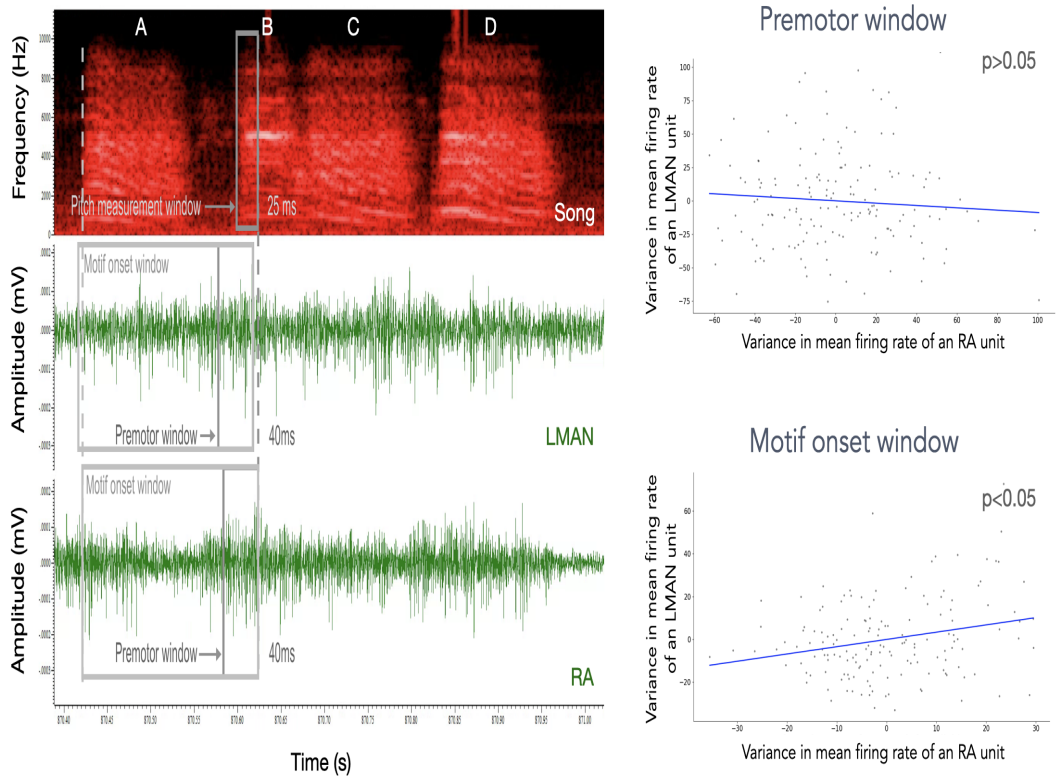


Figure 6.13: Sample correlation plot between the trial-to-trial variance in the neural activity (MDFR) of an RA and LMAN multi-unit within different time windows. Left: Sample simultaneous recording of an RA and LMAN multi-unit during singing. The premotor and motif onset windows with respect to the pitch measurement point of the target syllable are depicted in grey. The time window for LMAN is displaced with respect to that of RA to account for transmission delays. Right: Sample correlation plot between the trial-to-trial variance in the neural activity (MDFR) of a RA and LMAN multi-unit within the premotor and motif onset windows. Here, the units do not show a significant correlation ($p > 0.05$) in their activity in the premotor window for a syllable, but they show a significant correlation ($p < 0.05$) in their activity in the motif onset window for the syllable. The blue line denotes the linear regression across the data points. MDFR: Mean detrended firing rate.

6.3.3 Discussion

It has been shown that the song system has an influence on acoustic features, such as pitch (Tchernichovski et al. 2001). Tchernichovski et al. (2001) showed that three features (amplitude, spectral entropy and pitch) explained majority of the trial-by-trial variation in each syllable (Srivastava et al. 2015). Thus, in our study, we study the correlation between neural activity of nuclei in the song system and one of these acoustic features, the pitch, of the generated vocalisation. Here, we successfully induce adult plasticity in an adult male zebra finch by using the CAF protocol described by several studies (Skocik et al. 2013; Warren et al. 2011). The bird incrementally increases the pitch of the target syllable in order to escape aversive feedback. On absence of feedback post training, the bird gradually returns to producing the target syllable with a pitch distribution similar to pre-training days.

The glutamatergic projections from LMAN are topographic in nature. This implies that the influence of LMAN on RA might not necessarily be evident when looking at single units from each region, if the two single units do not belong within the same topographical cluster. Moreover, if populations of neuronal units in the LMAN drive RA activity, it would be more useful to scrutinise the correlation between multi units. Hence, in this study, we look at the correlations between the activity of multi units in both regions.

Multi-units in RA have shown significant correlations between their premotor activity and the acoustic features of their resulting vocalisations (Darshan et al. 2017; Sober, Wohlgemuth, et al. 2008). Similar to Sober, Wohlgemuth, et al. (2008), we observe approximately 15% of significant correlations between the premotor activity of RA units and the acoustic features of the subsequent syllable. For the target syllable in the CAF protocol, we see that the number of significant correlations between premotor activity of RA units and pitch is outside chance levels in the 'learning' category. As this is not the case for the 'maintenance' and 'baseline' categories, it could be an indicator towards increased influence of RA towards behaviour during learning.

To the best of our knowledge, there has been no studies reporting on the correlation between LMAN activity and vocal behaviour. In our dataset, the number of significant correlations obtained between the premotor activity of LMAN multi units and the pitch of the corresponding syllable is within chance levels. This resembles unpublished data from our team which also do not show any evident correlation between LMAN activity and vocal behaviour. Moreover, the hypothesis posits that, during learning, the premotor activity of LMAN would be more correlated with behaviour, i.e. acoustic features of corresponding syllable. However, the data presented here shows no such indicators. This is interesting in the context of the widely discussed theory of BG-driven reinforcement learning governing vocal learning in songbirds (M. Fee et al. 2011). As per this theory, the AFP functions as a tutor, and provides signals to the RA, the locus of motor control, via LMAN, the output of the AFP, in order to influence behaviour. If this is indeed the case, one would expect a correlation between variability in LMAN activity and behaviour. Its

absence would question the viability of the BG-driven reinforcement learning paradigm. However, the data collected here is insufficient to make any inferences about the correlation between LMAN activity and behaviour with certainty.

Next, we study the correlation between the neural activity in the premotor window of a syllable of two nuclei in the song system, the RA and LMAN. The number of correlations found between the neural activity, in the premotor window of any syllable group, of the RA and LMAN were within chance levels. Thus, we do not make inferences based on the pilot data at this time scale. To account for a wider range of timescale in which the LMAN may potentially drive RA, we compute the correlation between LMAN and RA neural activity within the motif onset window, as described in the previous section. We see that the number of significant correlations between the neural activity of RA and LMAN in the motif onset period for the two syllable groups is outside chance levels in the 'learning' category, but not so in the 'maintenance' and 'baseline' categories. This may be an indicator that the LMAN influences the RA on a longer timescale (100ms), while the RA exerts a temporally specific control on the generated motor action.

Due to the insufficient number of birds and electrophysiology recordings included in this pilot study, we only look for directions to continue the study and do not make a comparison with the hypothesis with certainty. The hypothesis stated that if the LMAN drives trial-to-trial variations in RA to influence behaviour, the activity of LMAN and RA would be more correlated during the 'learning' period than during the 'baseline' period. While the data presented here is insufficient to make conclusions, it provides indicators that the two nuclei may be more correlated during learning, on a timescale longer than the premotor window.

6.4 Appendix

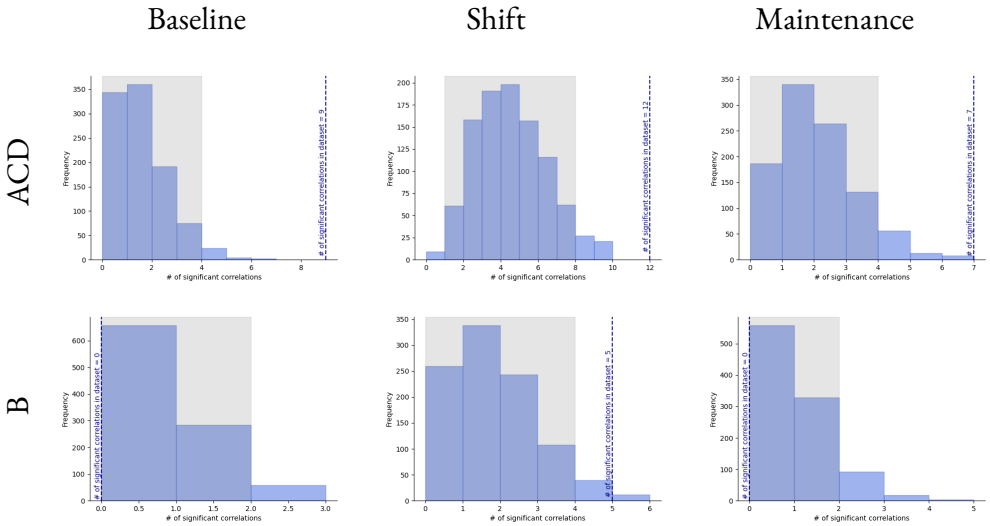


Figure 6.14: Bootstrapping results over the dataset presented in Table 6.2 for RA units across all three conditions for both syllable groups. The 95% confidence interval is shaded in grey.

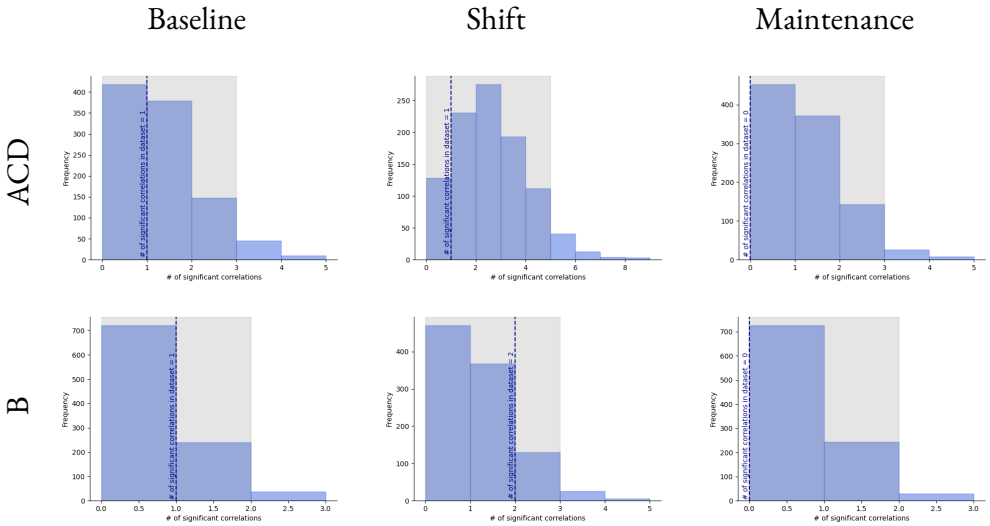


Figure 6.15: Bootstrapping results over the dataset presented in Table 6.2 for LMAN units across all three conditions for both syllable groups. The 95% confidence interval is shaded in grey.

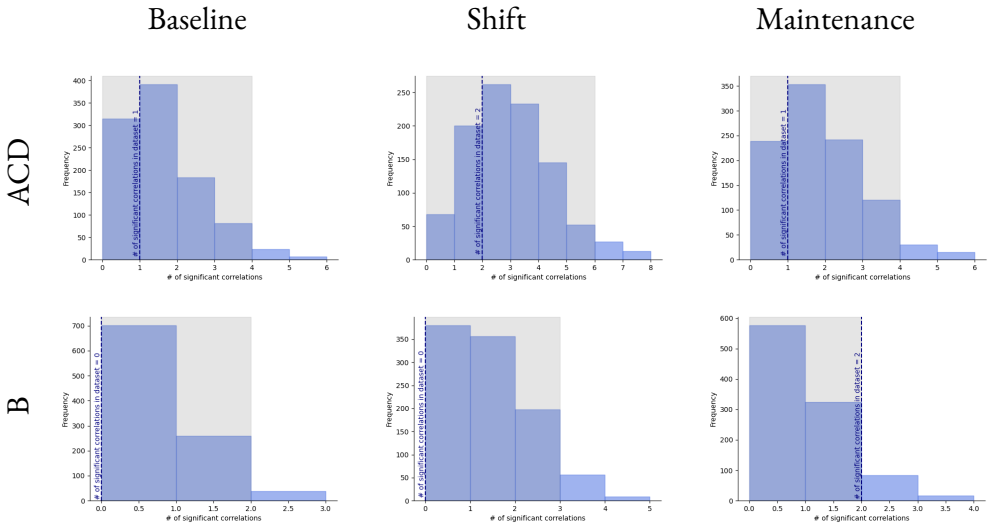


Figure 6.16: Bootstrapping results over the dataset presented in Table 6.3 for RA-LMAN multi unit pairs across all three conditions for both syllable groups in the premotor window. The 95% confidence interval is shaded in grey.

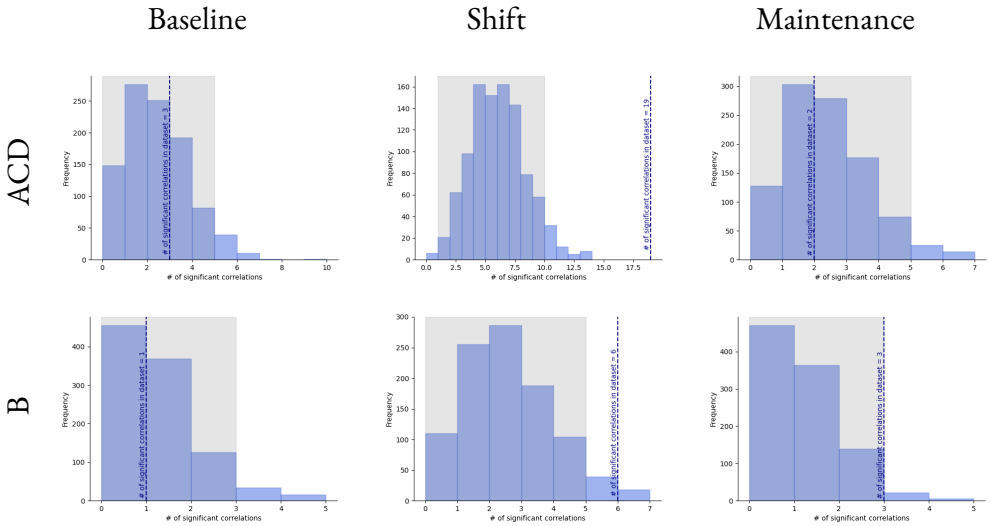


Figure 6.17: Bootstrapping results over the dataset presented in Table 6.3 for RA-LMAN multi unit pairs across all three conditions for both syllable groups in the motif onset window. The 95% confidence interval is shaded in grey.

7 Conclusion

It has been widely hypothesised that reinforcement learning underlies sensorimotor learning. The exact mechanism of how reinforcement learning is implemented within neural circuitry is unclear. In this manuscript, we use the vocal learning behaviour of songbirds to study the neural circuitry underlying sensorimotor learning.

The dual pathway model we present in chapter 5 demonstrates that the parallel pathways within the song system allows the songbird to take advantage of both reinforcement learning and Hebbian learning, while possibly avoiding the shortcomings of both mechanisms. This has benefits akin to the sensorimotor model built by Pyle et al. (2019), where two parallel pathways utilise two different learning algorithms, one based on supervised learning, while the other uses reward modulated Hebbian leaarning. The model harvests the benefits of both algorithms, with the former being better at maintenance of the learnt timeseries, while at the same time avoiding their pitfalls, with the former needing a supervisory signal with complete error information, unlike the latter. Similarly, a circuit with reinforcement learning alone could lead to highly variable output even post learning. On the other hand, a circuit with solely Hebbian learning might not be flexible enough to navigate the complex, uneven sensorimotor landscape of controlling the avian syringeal musculature, especially in the presence of a low dimensional error signal.

In chapter 5, we establish that structural plasticity could indeed contribute to the function of a neural circuit. Using a model of the song system in birds, we show that structural plasticity can facilitate vocal learning by helping to modulate the exploration-exploitation trade-off. It has been observed that the cortical pathway develops with a delay with respect to the parallel BG-thalamo-cortical pathway. The model presented in section 5.1 takes a look into the early sensorimotor phase (until 35 dph) and suggests that this delay in cortical maturation allows for unrestrained exploration by the BG. Such exploration can perhaps help form an inverse model of the neural patterns of the RA and the resulting vocalisations, the utility of which has been demonstrated by multiple studies (Giret et al. 2014; Hanuschkin et al. 2013).

Taking a step further into later stages of sensorimotor learning, section 5.2 and 5.3 look into the role of structural plasticity in vocal learning, beyond 30dph as well. Our models show that as the cortical pathway gradually develops, it can gradually suppress the influence of the BG-thalamo-cortical pathway on motor output. This inhibitory effect is in line with the study by Ölveczky, Andalman, et al. (2005) which posits that strong excitatory projections from HVC can saturate the RA, leading to a decrease in the ef-

fect of the NMDA-mediated glutamatergic projections from LMAN, the output of the BG-thalamocortical pathways in birds. Given that the role of inducing exploratory variability is generally attributed to the LMAN (Andalman et al. 2009a; Ölveczky, Andalman, et al. 2005), by diminishing its influence, the growth of the cortical pathway slowly tilts the exploration-exploitation tradeoff within the system towards exploitation. This effect is similar to that of the temperature parameter in simulated annealing, an optimisation technique used in machine learning (Tsallis et al. 1996). In simulated annealing, temperature controls the exploration-exploitation trade-off such that moving against the gradient is less likely as the temperature decreases. Thus, we demonstrate one potential mechanism through which structural plasticity can play a crucial role towards the function of a neural circuit, including sensorimotor learning.

Derégnaucourt et al. (2005) observe a post-sleep deterioration in song similarity to tutor song during sensorimotor learning. This deterioration was associated with an eventual superior quality of imitation of tutor song, during adulthood. In section 5.2 and 5.3, we present a potential mechanism that could implement a similar post-sleep deterioration and interpret its role as a facilitation factor in sensorimotor learning. We speculate that a similar daily drop in performance can be caused by overnight spine volatility within the HVC-X synapses, and use a theoretical model to investigate its effects. In our theoretical model, the daily deterioration in song quality post sleep corresponds with the exploration of a different region of the sensorimotor landscape by the BG pathway, which can facilitate evading being trapped in local optima. While detrimental to song quality on the shorter time scale, this could eventually lead to converging at the globally optimal solution, akin to simulated annealing. Simulated annealing conducts a discrete exploration of the available solutions, leading to a non-monotonic trajectory towards the target (Tsallis et al. 1996). The non-monotonic trajectory in song quality or its acoustic features, towards a better imitation of tutor song, as observed in the model, has been indicated in Derégnaucourt et al. (2005) as well as in instances of juvenile birds engaging in pitch doubling by juvenile birds (Tchernichovski et al. 2001). Such a trajectory is also partially in accordance with Kollmorgen et al. (2020). Using a novel technique of reservoir dating on the vocalisations in later stages of the sensorimotor phase (beyond 45dph), they observe the changes in acoustic features were partially misaligned on the daily and weekly timescales. Such a misalignment can be due to the comparatively poor performance at certain timepoints, while the song continues to improve on a longer weekly timescale.

In effect, as shown in chapter 5, the presence of two parallel pathways could allow the bird to use an interesting strategy to navigate complex performance landscapes. It could potentially use the cortical pathway to maintain a memory of past exploration, while allowing unconstrained exploration by the subcortical pathway on a daily basis, without a need to keep track of its exploration on a longer timescale. This has parallels to the discontinuous exploration conducted by simulated annealing, but eliminates the need for maintaining a memory of the best candidate solution yet encountered. Thus, our theoretical model suggests that one, the spine volatility within HVC-X dendrites can be a possible in-

ternal mechanism for post-sleep deterioration of song quality during early sensorimotor learning, two, the daily deterioration in song quality can lead to an eventual improvement in song quality by facilitating the search for the global optimum, and three, the delayed maturation of the cortical pathway facilitates early sensorimotor exploration, the fruits of which can be exploited in the subsequent stages of learning.

Our model of vocal learning uses gradient descent based learning led by the BG which perturbs RA nodes, similar to Fiete, M. S. Fee, et al. (2007). However, direct gradient descent approaches tend to get stuck in local optima in complex, uneven landscapes (Bottou et al. 1991; Gori et al. 1992). To address this shortcoming, model proposes a biological plausible strategy to optimise gradient descent. Moreover, our model is tested on a performance landscape derived from a simplified mathematical model of the avian syrinx (Amador et al. 2013). Vocal learning in songbirds is likely more complex and higher dimensional problem to solve. Benureau (2015) demonstrate that with increasing dimensionality (referred to as the curse of dimensionality in machine learning circles), random exploration of the motor space is insufficient to induce a uniform exploration of the sensory space. Both our model and Fiete, M. S. Fee, et al. (2007) use random exploration of the motor space to form an internal model of the landscape for gradient descent. While the optimisation suggested in chapter 5 could help address this issue, it remains to be tested on a higher dimensional task. Either a comprehensive model of the syrinx, or the analogy of a multi-segmented arm (used in section 5.1) can be an appropriate task to test the feasibilty of the optimisation suggested by our dual pathway model. It is important to note that our model does not account for temporal components of a motor gesture and reduces a syllable from a complex temporally varying acoustic signal to a scalar feature at a single time point. We also do not account for delays in feedback signal with respect to the production of a vocal output. Moreover, song learning involves several components in addition to learning how to produce a desired vocalisation, which are not addressed by our model. For instance, it also involves sequence learning, which is provided as an input to our model, but needs to be learnt by songbirds, as simulated by Yildiz et al. (2011). While we use a rate-coded neurons and look at the average firing rate of a neuron within a time window, there are models of vocal learning, such as Farries and Fairhall (2007) which take a step further towards biological plausibility implementations of reinforcement learning by incorporating spike timing dependent plasticity rules to train a neural network.

Moving to the second part of our study, we look into viability of the avian BG-thalamocortical loop (AFP) to implement reinforcement learning, as per the hypothesis proposed by M. Fee et al. (2011). In chapter 6, we designed an experimental protocol to verify the feasibility of BG-driven reinforcement learning within the song system in male zebra finches, and its eventual consolidation in the cortical motor pathway. First, we posit that if BG drives reinforcement learning, then the LMAN, the output nucleus of the AFP, would influence vocal output during learning. In order to test this hypothesis, we conduct pilot electrophysiological studies to record the activity of the LMAN and RA in a freely-moving bird engaged in vocal learning. Our recordings from the RA show correlation

levels in the premotor period with vocal behaviour, similar to those reported by Sober, Wohlgemuth, et al. (2008). We posit that, if the LMAN influences behaviour during vocal plasticity, the biased variability injected by LMAN would also show a correlation with the acoustic features of the produced vocalisations. However, we see no such correlation between the LMAN in the premotor window and resulting vocalisations in our preliminary dataset. This is in agreement with unpublished results from our lab. The lack of correlation between premotor LMAN activity and vocal behaviour, during learning, can raise serious questions about the circuit implementation of the BG-driven learning hypothesis. However, keeping in mind the insufficient data collected, more experiments would be needed to conclusively discuss this issue.

Further, in order to consolidate the tutor signals from such BG-driven reinforcement learning, the output nucleus of the AFP, LMAN, will need to influence the activity of RA, the locus of motor control within the cortical motor pathway. Within this hypothesis, the LMAN would influence behaviour by driving RA activity. We first build a model to make qualitative predictions about the interaction between RA and LMAN, based on this hypothesis. Then, we conduct electrophysiology recordings from the LMAN and RA, during vocal learning. From the pilot dataset that we collect, we see no clear evidence of correlation between trial-to-trial variability within RA and LMAN activity within the premotor window (40ms). However, our analysis indicates the presence of such a correlation on a much longer timescale of 100-300ms between trial-to-trial variability in neural activity of RA and LMAN. This raises interesting questions regarding the mechanisms through which the LMAN influences RA activity. Further electrophysiological studies with simultaneous recording of RA and LMAN during vocal learning are necessary to speak with certainty about the timescale at which LMAN drives the RA. Also, further analysis of the data is required. For instance, analysis of the correlation between activity of LMAN and RA during the entire song motif can give us further insight into the timescale at which LMAN influences the RA. Further, analysis of the local field potential can give us a better idea of the population-level interaction between the two nuclei.

Thus, the dual pathway architecture underlying sensorimotor learning, investigated in this study, incorporates the widely discussed theory of transfer of learning from subcortical to cortical structures. Additionally, we propose biologically realistic optimisations of BG-driven learning which is gradually consolidated into cortical pathways, within a two-step learning paradigm. Finally, we design and test an experimental protocol to investigate the feasibility of such a two-step learning paradigm within the song system of zebra finches.

Appendix

Here, we include our replication study of a sensorimotor learning model by Pyle et al. (2019), which simulates the transfer of sensorimotor learning from subcortical to cortical networks. It harvests the dynamics of a sparsely connected recurrent neural network to generate a spatio-temporal signal. We adapted this model for an additional verification of the hypothesis presented in chapter 6, within a dynamical system.

RESCIENCE C

Replication / Computational Neuroscience

[Re] A Reservoir Computing Model of Reward-Modulated Motor Learning and Automaticity

Remya Sankar[1,2,3], Nicolas Thou[1], Nicolas P. Rougier[1,2,3] Arthur Leblois[3]

[1]INRIA Bordeaux Sud-Ouest, Bordeaux, France – [2]LaBRI, Université de Bordeaux, Institut Polytechnique de Bordeaux, Centre National de la Recherche Scientifique, UMR 5800, Talence, France – [3]Institut des Maladies Neurodégénératives, Université de Bordeaux, Centre National de la Recherche Scientifique, UMR 5293, Bordeaux, France

Edited by
Benoît Girard

Reviewed by
Daniel Schmid
Robin Gutzen

Received
12 March 2020

Published
22 November 2021

DOI
10.5281/zenodo.5718075

A replication of pyle2019.

1 Introduction

Pyle and Rosenbaum [1] introduce a novel learning algorithm to the reservoir computing framework, which harnesses the dynamics of a recurrently connected network to generate time series. Most existing algorithms are built on fully supervised learning rules (e.g. FORCE [2]), which limits their potential applications, or the more biologically-realistic reinforcement learning techniques (e.g. RMHL [3]) which unfortunately fail to converge on complex spatio-temporal signal generation tasks. Pyle and Rosenbaum [1] use the advantages of these two learning rules, while averting their individual shortcomings, by combining the two algorithms to form the SUPERTREX model. The workings of this model are aligned to the theory of motor learning involving the basal ganglia, from rodent and songbird literature [4]. This hypothesises that a cortical pathway works in tandem with the basal ganglia for motor skill acquisition, wherein the basal ganglia pathway functions as a tutor, providing guiding signals that would ultimately be consolidated in the primary cortical pathway in charge of production of the motor commands [5]. Here, the basal ganglia pathway, which uses reward-modulated exploration based learning, akin to the RMHL algorithm, works in parallel with the cortical pathway, modeled using the fully supervised FORCE algorithm. The SUPERTREX model uses both these pathways in parallel, with the RMHL-based pathway providing the supervisory signal that the FORCE-based pathway requires.
In this article, we provide a modular and user-friendly Python re-implementation of the model presented by Pyle and Rosenbaum [1]. We were able to successfully reproduce the model performance in Python, for two tasks out of the three presented in the original article. For the third task, we were able to do so with limited robustness. We address this by introducing some modifications, and discuss how their inclusion vastly improves the robustness as well as scalability of the model.

Terminology –

- Original scripts: the MATLAB scripts used by the authors to produce the results presented in [1].
- Python adaptation: our Python adaptation of the original MATLAB scripts.
- Python re-implementation: an improved version of our Python adaptation.

Correspondence should be addressed to Remya Sankar (Remya.Sankar@inria.fr)
The authors have declared that no competing interests exists.
Code is available at https://github.com/rsankar9/Reimplementation-SUPERTREX/releases/tag/v3.0 – DOI https://doi.org/10.5281/zenodo.4596425.
Open peer review is available at https://github.com/ReScience/submissions/issues/50.

1.1 Framework

Pyle and Rosenbaum [1] proposes a model for sensorimotor learning using the framework of reservoir computing. The model is based on two existing reservoir computing techniques: FORCE and RMHL.

$$\tau \frac{d\mathbf{x}}{dt} = -\mathbf{x} + J\mathbf{r} + Q\mathbf{z} \tag{1}$$

$$\mathbf{r} = \tanh(\mathbf{x}) + \epsilon \tag{2}$$

FORCE (or first-order reduced and controlled error) is a fully supervised learning rule, which is widely used within the reservoir computing framework [2]. A recurrently connected reservoir, composed of rate-coded neurons is trained to produce a target time series by modifying the readout weights between the reservoir and the output layer (Eq 3, 4). The output, in turn, interacts with the reservoir by providing feedback (Eq 1, 2). FORCE can accurately generate complex dynamical target time-series. However, the model must have explicit knowledge of the target function, as FORCE requires a fully supervisory signal of the correct output in order to compute the error during training.

$$\mathbf{z}_1 = W_1\mathbf{r} \tag{3}$$

$$\tau_{w_1} \frac{dW_1}{dt} = -\mathbf{e}\mathbf{r}^T P \tag{4}$$

RMHL (or Reward-Modulated Hebbian Learning) is built on the concept of reinforcement learning, and uses only a scalar error signal indicating reward, allowing it to be applicable in a wider range of scenarios than FORCE [3]. RMHL introduces perturbations in the performance of the model, and uses the information gained from this exploration to find the target (Eq 5, 6). This is akin to dopamine-dependent Hebbian learning in the basal ganglia. However, RMHL fails to converge to an accurate solution on several complex tasks. Moreover, it has been observed in songbirds that while the basal ganglia provides a tutor signal in the early stages, learning is eventually consolidated in a parallel cortical pathway, which is primarily responsible for motor activity [5]. RMHL cannot account for such empirical observations.

$$\mathbf{z}_2 = W_2\mathbf{r} + \Psi(e)\boldsymbol{\eta} \tag{5}$$

$$\tau_{w_2} \frac{dW_2}{dt} = \Phi(\hat{e})\hat{\mathbf{z}}\mathbf{r}^T \tag{6}$$

SUPERTREX (Supervised Learning Trained by Reward Exploration), the model proposed by the authors, tries to merge the advantages of both of these algorithms by combining both models. The more wide-ranged applicability of RMHL, owing to its usage of a one dimensional error signal, is used to train the model, while the superior maintenance ability of FORCE is recruited to consolidate the tutoring of the RMHL pathway. This could also potentially support the empirical evidence showing the basal ganglia and cortical pathways working in tandem for motor skill acquisition, discussed above. Thus, the SUPERTREX model consists of two parallel pathways, one based on RMHL (exploratory) and one based on FORCE (mastery), each consisting of its own set of weights. The mastery pathway uses the output of the exploratory pathway as its supervisory signal (Eq 7, 8, 9).

$$\mathbf{z} = \mathbf{z}_1 + \mathbf{z}_2 \tag{7}$$

$$\tau_{w1}\frac{dW_1}{dt} = (\mathbf{z} - \mathbf{z}_1)\,\mathbf{r}^T P \tag{8}$$

$$\tau_{w2}\frac{dW_2}{dt} = \Phi(\hat{e})\hat{\mathbf{z}}\mathbf{r}^T \tag{9}$$

where $\mathbf{x}$ denotes the reservoir dynamics, J the recurrent connectivity matrix, Q the feedback weights and $\mathbf{r}$ the reservoir activity. The output $\mathbf{z}$ of the SUPERTREX model is the combination of the outputs $\mathbf{z}_1$ and $\mathbf{z}_2$ of the FORCE and RMHL pathways, respectively. W_1 and W_2 denote the readout weights of the two pathways, respectively, e denotes the squared distance between the output and the target trajectory, τ is the corresponding timescales for learning, η is the exploratory noise, ϵ is a small noise term and P is a running estimate of the inverse of the correlation matrix of rates. Ψ and Φ are two sub-linear functions that serve to damp runaway oscillations during learning and control weight update, respectively. $\hat{x}$ is a high-pass filtered version of x, which represents the recent changes in x.

1.2 Task

The authors test the SUPERTREX model on three motor tasks, with increasing difficulty, and compare its performance to those of FORCE and RMHL. The target of each task is to learn to produce a given spatio-temporal signal, under different constraints. Task 1 tests the performance of the model when the target output is known. This task requires the spatio-temporal signal to be produced directly by the model. Thus, the error signal is a direct indicator of the change required in the output of the model, i.e. fully supervisory. Task 2 and Task 3 use the paradigm of exploration by a multi-segmented arm, pivoted at a point.

Task 2 tests the performance of the model when the target output is unknown, and only an indirect error signal is provided. The task requires the angles between the arm segments to be generated by the model, which would in turn produce the trajectory of the target spatio-temporal signal. In this case, the error signal is not a direct indicator of the change required in the output of the model. A non-linear inverse transformation of the trajectory would be required to compute the desired angles between the arm segments. It is, thus, not a fully supervisory signal, but simply a reinforcement signal. In Task 3, the movement of the arm segments are penalised variably. This creates the need to choose one from multiple candidate solutions by optimising the cost of changing the angles between the arm segments.

The simulation for each task includes a training phase and a testing phase. In the training phase, for ten periods, the time-series is generated by the model while the weights are being updated according to the current error feedback. After this, in the testing phase (lasting five periods), the readout weights are frozen and the time-series is generated using these frozen weights, without any further feedback-based update. In the SUPERTREX model, the exploratory pathway is also deactivated. It is also worth noting that the authors use teacher-forcing in the testing phase, which considerably improves the model performance by limiting the dependence on the stability of the learned solution (refer to Section "State information provides stability of learned output" in [1]).

Disclaimer — Pyle and Rosenbaum [1] proceed to test the model under variations of the above tasks, including disrupted learning and with additional state information. However, these variations have not been replicated by us. We only test the performance of the three learning rules on the three tasks, specified above.

2 Comparison with Python Adaptation

In this section, we compare the results presented in the paper [1] with the MATLAB implementation by the authors and our Python adaptation. The original scripts, although not available online, are readily available on request. We present our adaptation of this model in the open source framework Python, which has been built based on the paper and the MATLAB scripts provided by the authors. In contrast to the original scripts, it is modular and is easily modifiable with external json descriptor files. We compare the results presented in the paper, with simulations of the MATLAB scripts, provided by the authors, and also with our adaptation in Python [1].

To validate our Python adaptation, we test the three algorithms on three tasks by simulating them using both the original scripts and our Python adaptation. For each algorithm-task combination, we produce ten simulations with arbitrary seeds initialising the random generator and one additional simulation using the default seed of MATLAB (equivalent to seed 5489 of the numpy random generator). Except for the default seed, the ten arbitrary seeds are different for the Python and MATLAB simulations, and for each algorithm-task combination. Both MATLAB and numpy use the Mersenne Twister pseudo-random number generator [6]. To evaluate the performance of the algorithm, the authors plot the "distance from target", i.e. the square root of the low pass filtered version of the mean squared error, over the progression of the simulation. In order to categorise the model performance as satisfactory or unsatisfactory, we further compute a **deviation metric** by calculating the mean "distance from target" over the testing phase. If this deviation metric is below the threshold of 0.5 (set by visual inspection), the model is said to have satisfactorily learnt and produced the target output.

2.1 Task 1

Here, we compare the simulations of the original scripts and our adaptation for Task 1, using FORCE, RMHL and SUPERTREX, with the results presented in the article. Task 1 is designed to test the performance of these three algorithms when generating a known target output. The objective of this task is to produce a time-series of 2-D coordinates required to traverse a target trajectory, in this case, the parameterized curve of a butterfly. The model is trained to generate an output which closely matches the target function.

The article claims that:

- under the FORCE framework, the target time-series is learned accurately and is maintained in a stable manner during the testing phase (Figure 1a).
- under the RMHL framework, the target time-series is generated accurately during the training phase, however is not maintained perfectly during the testing phase (Figure 1b).
- under the SUPERTREX framework, the target time-series is learned accurately and is also maintained in a stable manner during testing phase, albeit not as well as FORCE (Figure 1c).

We validate these observations with the MATLAB scripts provided by the authors as well as with our Python adaptation. To do so, we run the simulations with the default seed and repeat it ten times with different (arbitrarily chosen) seeds initialising the random number generator.
We observe that:

[1] In Figures 1- 4, the results presented in the paper have been reused in the column titled "original".

- under the FORCE framework, the target time-series is learned accurately and is maintained in a stable manner during the testing phase, as claimed. The mean deviation over eleven simulations, for both the original scripts (0.003 ± 0.002; n=11) and the Python adaptation (0.004 ± 0.003; n=11) is much lower than the threshold of 0.5 (Figure 1a, 2a).
- under the RMHL framework, the target time-series is generated accurately during the training phase, however is not maintained perfectly during the testing phase, as claimed. The mean deviation, for both the original scripts (0.168 ± 0.038; n=11) and the Python adaptation (0.182 ± 0.046; n=11), is higher than that with FORCE (Figure 1b, 2b).
- under the SUPERTREX framework, the target time-series is learned accurately and is also maintained in a stable manner during testing phase, albeit not as well as FORCE, as claimed. The mean deviation, for both the original scripts (0.006 ± 0.003; n=11) and the Python adaptation (0.006 ± 0.003; n=11), is much better than that for RMHL, but slightly worse than with FORCE (Figure 1c, 2c).

Both the original scripts and the Python adaptation are able to successfully closely reproduce the results presented in the paper for Task 1 (Figure 1,2; Table 1, 2).

2.2 Task 2

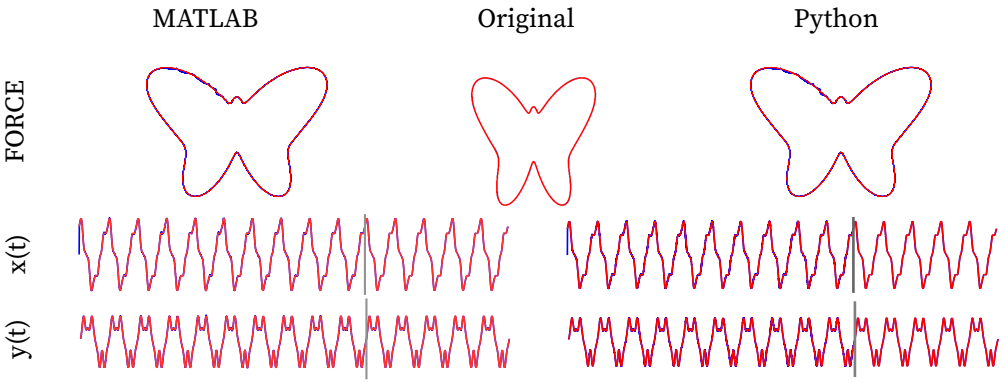
Here, we compare the simulations of the original scripts and our Python adaptation for Task 2, using FORCE, RMHL and SUPERTREX, with the results presented in the article. Task 2 is designed to test the performance of these three algorithms when generating an unknown target from an indirect error signal. Using the paradigm of a pivoted multi-segmented arm, the objective of this task is to produce a time-series by generating the angles between the arm segments. Motor output does not control the position of the end-effector of the arm, but instead controls the angles of the arm joints, which are non-linearly related to end-effector position.

The article claims that:

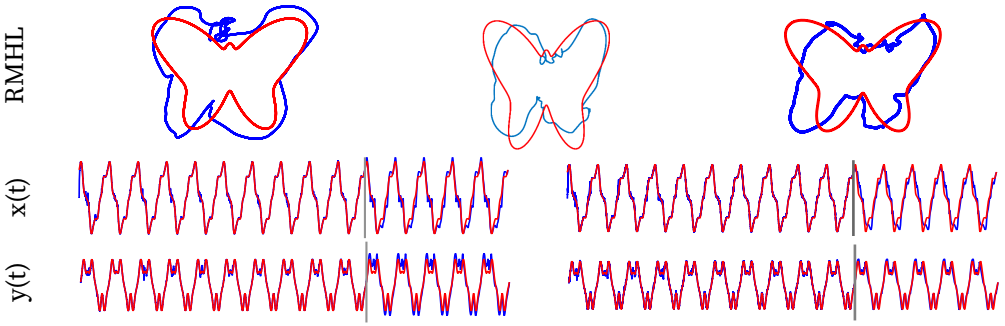
- the FORCE framework cannot be applied to this task, as FORCE requires the exact target to be provided as a supervisory error, which in this case would be the unknown target angles. Since, we do not have this information beforehand, and require the model to derive it, the FORCE framework is inapplicable to this task.
- under the RMHL framework, the target time-series is imitated well by the model during the training phase, however the weights do not converge, and hence, it is unable to maintain the time-series in a stable manner during the testing phase (Figure 3a).
- under the SUPERTREX framework, the target time-series is learned accurately and is also generated in a stable manner, with minor divergences, during testing phase, owing to the contribution of the pathway based on the FORCE algorithm (Figure 3b).

We verify these observations with the MATLAB scripts provided by the authors as well as with our Python adaptation. To do so, we run the simulations with the default seed of MATLAB and re-simulate it with ten arbitrary seeds initialising the random number generator. We do not modify any task conditions or model hyper-parameters.
We observe that:

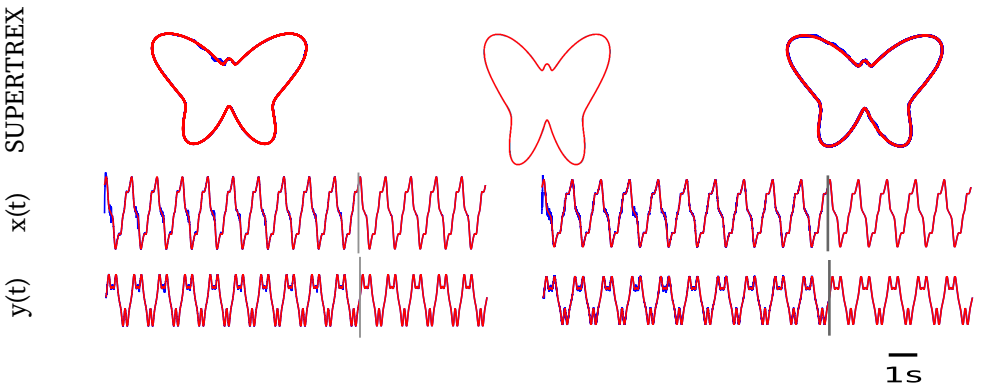
- indeed, the FORCE framework is inapplicable to this task.



(a) Results for Task 1 with the FORCE algorithm. The target time-series is learned accurately during the training phase and is maintained in a stable manner during the testing phase, in both implementations, as presented in [1].

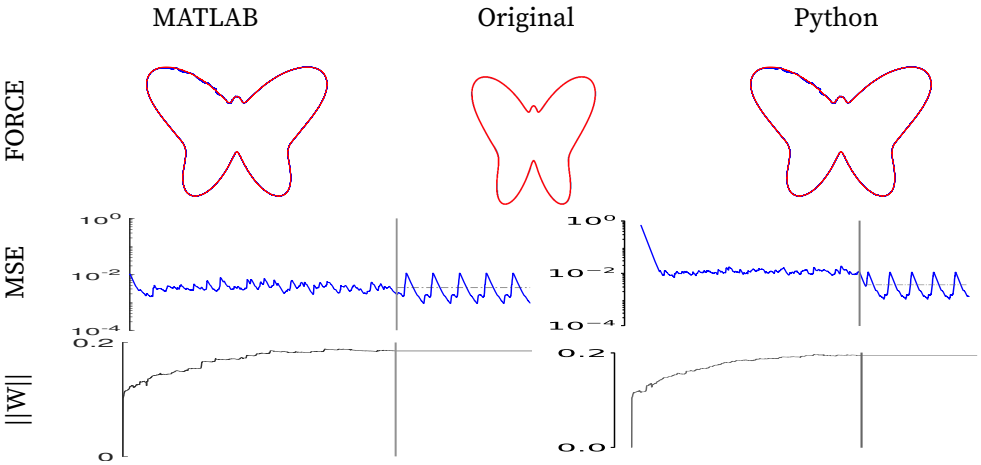


(b) Results for Task 1 with the RMHL algorithm. The target time-series is learned accurately during the training phase, though not maintained perfectly during the testing phase, in both implementations, as presented in [1].

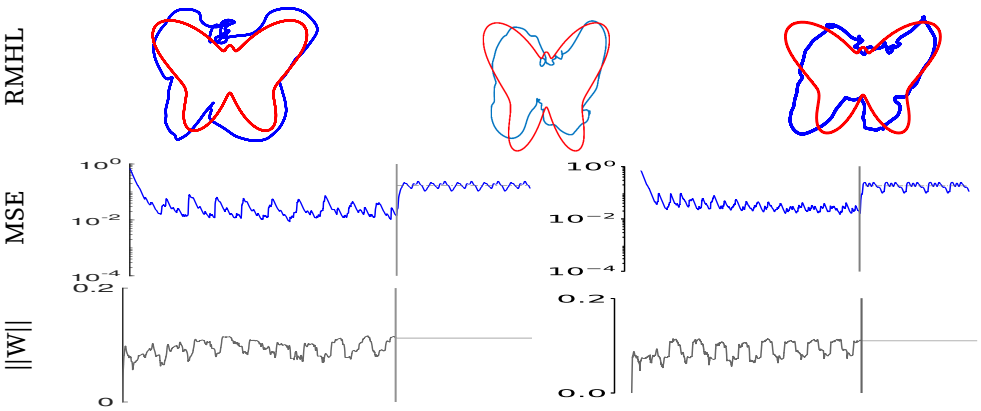


(c) Results for Task 1 with the SUPERTREX algorithm. The target time-series is learned accurately during the training phase, and is also maintained in a stable manner during testing phase, albeit not as well as FORCE, in both implementations, as presented in [1].

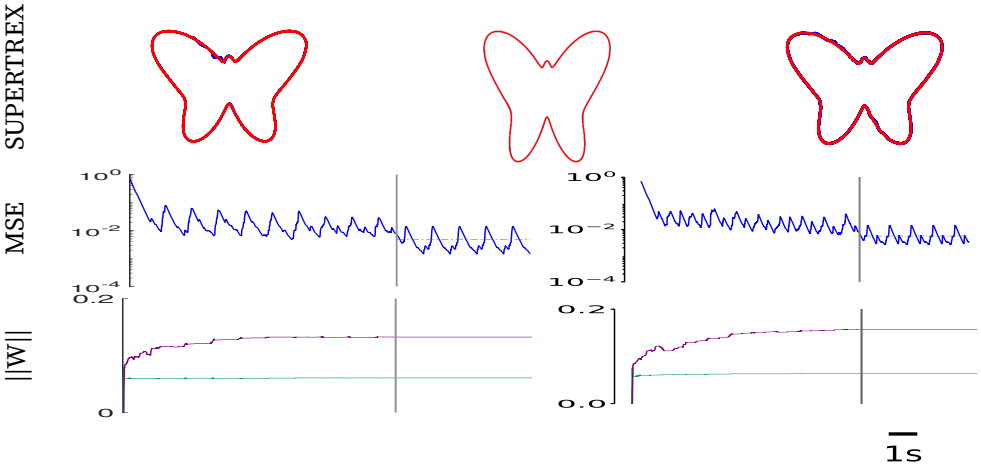
Figure 1. Comparison of the performances of the MATLAB scripts (left column) and the Python adaptation (right column) with the results presented in the original article (center column), for the three learning algorithms on Task 1 [1]. All simulations shown here use the MATLAB default (5489) as the seed for the random number generator. In each subfigure, the top row shows the target trajectory (red) with the trajectory generated by the model (blue) throughout the test phase. The second row shows the time-series (blue) generated by the model (x and y coordinates, in this case) along with the target time-series (red). The grey vertical line marks the separation of the training and testing phase.



(a) Results for Task 1 with the FORCE algorithm. The target time-series is learned accurately during the training phase and is maintained in a stable manner during the testing phase, as presented in [1].



(b) Results for Task 1 with the RMHL algorithm. The target time-series is learned accurately during the training phase, though not maintained perfectly during the testing phase, as presented in [1].



(c) Results for Task 1 with the SUPERTREX algorithm. The target time-series is learned accurately during the training phase, and is also maintained in a stable manner during testing phase, albeit not as well as FORCE, as presented in [1].

Figure 2. Comparison of the performances of the MATLAB scripts (left column) and the Python adaptation (right column) with the results presented in the original article (center column), for the three learning algorithms on Task 1 [1]. All simulations shown here use the MATLAB default (5489) as the seed for the random number generator. In each subfigure, the top row shows the target trajectory (red) with the trajectory generated by the model (blue) throughout the test phase. The second row shows the error metric (blue) over the simulation (x and y coordinates, in this case), using the log scale for the y axis. The bottom row shows the progression of the corresponding weight matrices (SUPERTREX: W_1 in purple; W_2, in green). The horizontal grey line, in the test phase, indicates the deviation metric.

- under the RMHL framework, the target time-series is imitated well by the model during the training phase, however the weights do not converge, and hence, it is unable to maintain the time-series in a stable manner during the testing phase. The mean deviation over eleven simulations, for both the original scripts (0.759 ± 0.284; n=11) and the Python adaptation (0.814 ± 0.288; n=11) is higher than the threshold of 0.5 (Figure 3a).
- under the SUPERTREX framework, the target time-series is learned accurately and is also generated in a stable manner, with minor divergences, during testing phase, owing to the contribution of the pathway based on the FORCE algorithm. The mean deviation over eleven simulations, for both the original scripts (0.011 ± 0.003; n=11) and the Python adaptation (0.012 ± 0.005; n=11) is below the threshold of 0.5 and much lower than that with RMHL (Figure 3b).

The MATLAB scripts provided by the authors and the Python adaptation are able to successfully closely reproduce the results presented for Task 2 in the paper, with the default seed as well as with the 10 arbitrary seeds (Figure 3; Table 1, 2).

2.3 Task 3

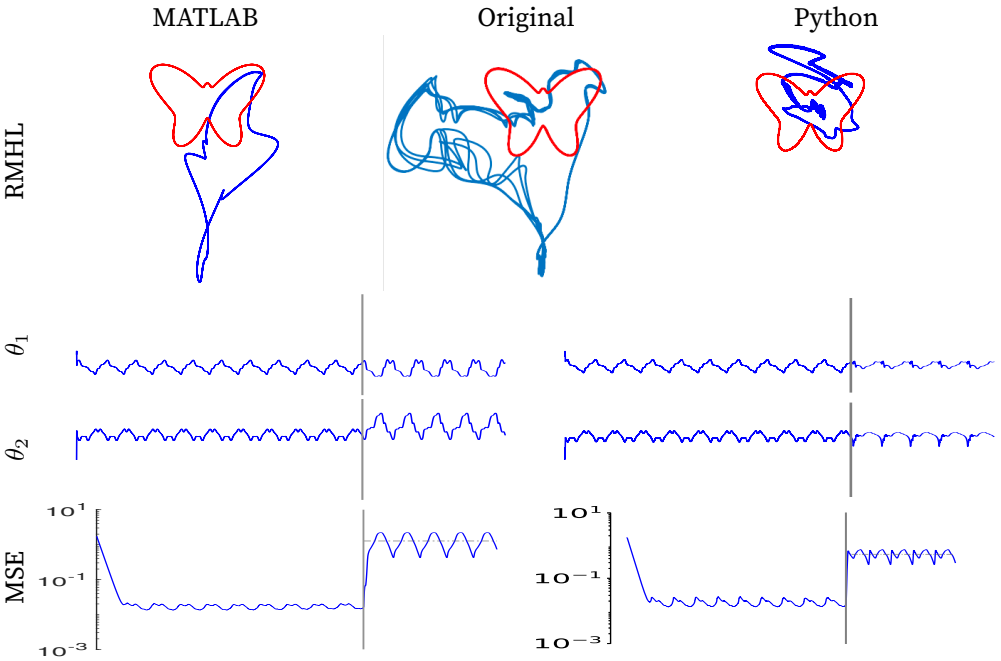
Here, we compare the performance of the MATLAB scripts and our Python adaptation on Task 3, for the three algorithms, with the results presented in the article. Task 3 is an extension of Task 2, designed to test the constraint optimisation ability of these three algorithms when generating an unknown target from an indirect error signal. Using the paradigm of a pivoted multi-segmented arm, the objective of this task is to produce a time-series by generating the angles between the arm segments, while also optimising the movement cost of each arm segment. Hence, the arm is required to traverse the butterfly, while carefully choosing the segment to rotate, in order to minimise the movement cost of its segments. Post the training phase, the readout weights are frozen and in the SUPERTREX model, the exploratory pathway is deactivated.

The article claims that:

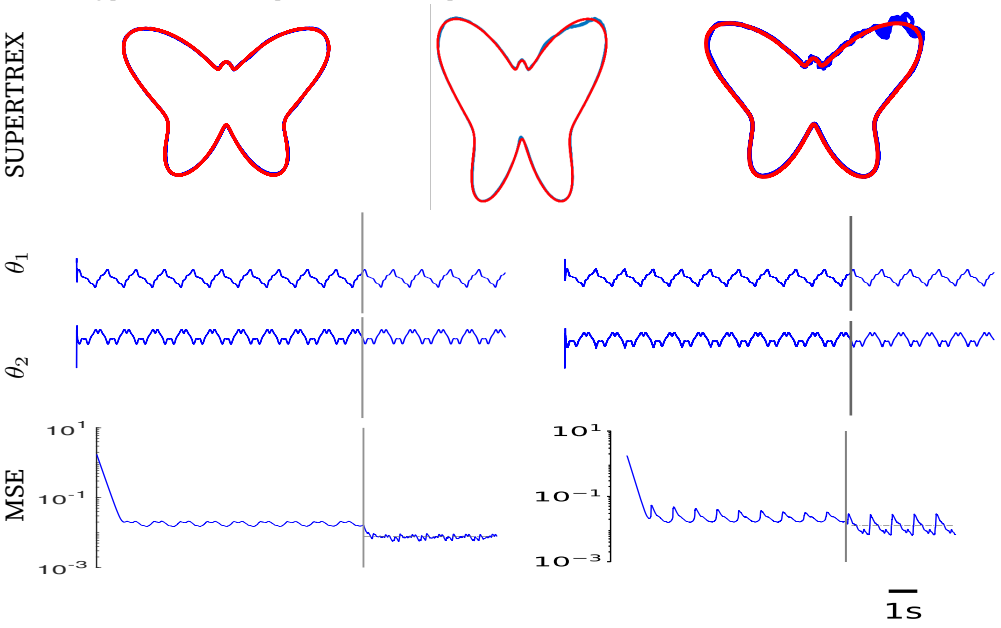
- FORCE can't be applied to this task, as explained for Task 2.
- under the RMHL framework, the target time-series is imitated well by the model during the training phase, however the weights do not converge, and hence, it poorly maintains the time-series during the testing phase (Figure 4a).
- under the SUPERTREX framework, the performance is much better than RMHL. The target time-series is learned accurately and is also generated with minor divergences, during testing phase (Figure 4b).

We verify these observations with the MATLAB scripts provided by the authors as well as with our Python adaptation. To do so, we run the simulations with the default seed of MATLAB and re-simulate it with ten arbitrary seeds initialising the random number generator.
We observe that:

- indeed, the FORCE framework is inapplicable to this task, as claimed.
- under the RMHL framework, the target time-series is imitated well by the model during the training phase, however the weights do not converge, and hence, it poorly maintains the time-series during the testing phase, as claimed. The mean deviation over eleven simulations, for both the original scripts (0.850 ± 0.313; n=11) and the Python adaptation (0.658 ± 0.216; n=11) is higher than the threshold of 0.5. All eleven simulations with different seeds did not generate the target output in a satisfactory manner (i.e. deviation > 0.5 for 11/11 seeds) (Figure 4a).



(a) Results for Task 2 with the RMHL algorithm. The target time-series is imitated well by the model during the training phase (not shown), however, it is unable to maintain the time-series in a stable manner during the testing phase, in both implementations, as presented in [1].



(b) Results for Task 2 with the SUPERTREX algorithm. The target time-series is learned accurately during the training phase, and is also maintained in a stable manner, during the testing phase, in both implementations, as presented in [1].

Figure 3. Comparison of the performances of original scripts (left column) and Python adaptation (right column) with the results presented in the original article (center column), for the RMHL and SUPERTREX, on Task 2 [1]. All simulations shown here use the MATLAB default (5489) as the seed for the random number generator. In each subfigure, the top row shows the target trajectory (red) with the trajectory generated by the algorithm (blue) throughout the test phase. The second row shows the time-series (blue) generated by the model (joint angles (θ_i), in this case). The bottom row shows the distance from target metric (blue) over the simulation (x and y coordinates, in this case), using the log scale for the y axis. The horizontal grey line, in the test phase, indicates the deviation metric. The grey vertical line marks the separation of the training and testing phase.

		MATLAB			Python adaptation			Python re-implementation		
Task	Model	Mean	Median	Std	Mean	Median	Std	Mean	Median	Std
#1	FORCE	0.003	0.002	0.002	0.004	0.003	0.003	0.003	0.002	0.003
	RMHL	0.168	0.165	0.038	0.182	0.182	0.046	0.201	0.203	0.053
	ST	0.006	0.004	0.003	0.006	0.005	0.003	0.004	0.003	0.003
#2	RMHL	0.759	0.740	0.284	0.814	0.799	0.288	0.697	0.681	0.263
	ST	0.011	0.010	0.003	0.012	0.011	0.005	0.010	0.009	0.004
#3	RMHL	0.850	0.794	0.313	0.658	0.647	0.216	0.849	0.794	0.360
	ST	0.881	0.845	0.224	0.837	0.827	0.241	0.140	0.116	0.071
#2'	RMHL	0.846	0.807	0.299	0.738	0.713	0.256	0.839	0.794	0.310
	ST	0.016	0.015	0.007	0.009	0.008	0.003	0.067	0.062	0.035

Table 1. Deviation metric showing the performance of the original MATLAB scripts, Python adaptation and Python re-implementation on different tasks. Each variant is simulated with the default seed (5489) and ten additional seeds. The mean, median and standard deviation of the deviation metric over these eleven simulations are tabulated here. Note that for task #2', the SUPERTREX statistics have been computed using only 2 simulations, for the original MATLAB scripts and Python adaptation. (ST: SUPERTREX; #2': 3 segment variant of Task 2)

		MATLAB		Python adaptation		Python re-implementation	
Task	Model	Satisfactory	Total	Satisfactory	Total	Satisfactory	Total
#1	FORCE	11	11	11	11	11	11
	RMHL	11	11	11	11	11	11
	ST	11	11	11	11	11	11
#2	RMHL	0	11	0	11	0	11
	ST	11	11	11	11	11	11
#3	RMHL	1	11	2	11	0	11
	ST	5	11	4	11	10	11
#2'	RMHL	0	11	1	11	0	11
	ST	2	2	2	2	11	11

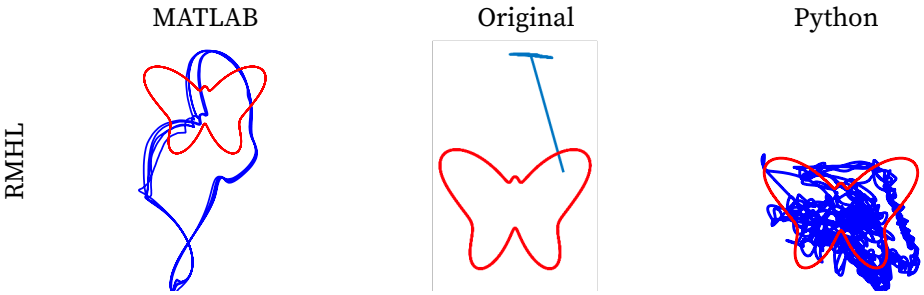
Table 2. The proportion of model simulations categorised as having satisfactory performance. Each variant is simulated with the default seed (5489) and ten additional seeds. Number of satisfactory simulations refers to the number of simulations that were below the threshold (0.5) for the deviation metric. The total number of simulations refer to the number of simulations which successfully reached completion, without the weights growing exponentially. (ST: SUPERTREX; 2': 3 segment variant of Task 2)

- under the SUPERTREX framework, the performance is not much better than RMHL, contrary to the article's claim. The target time-series is not generated in a satisfactory manner, during the testing phase, for more than 50% of the tested simulations (Original scripts: 6/11 and Python adaptation: 7/11). The mean deviation over eleven simulations, for both the original scripts (0.881 ± 0.224; n=11) and the Python adaptation (0.837 ± 0.241; n=11) is above the threshold of 0.5 and comparable with that of RMHL (Figure 4b).

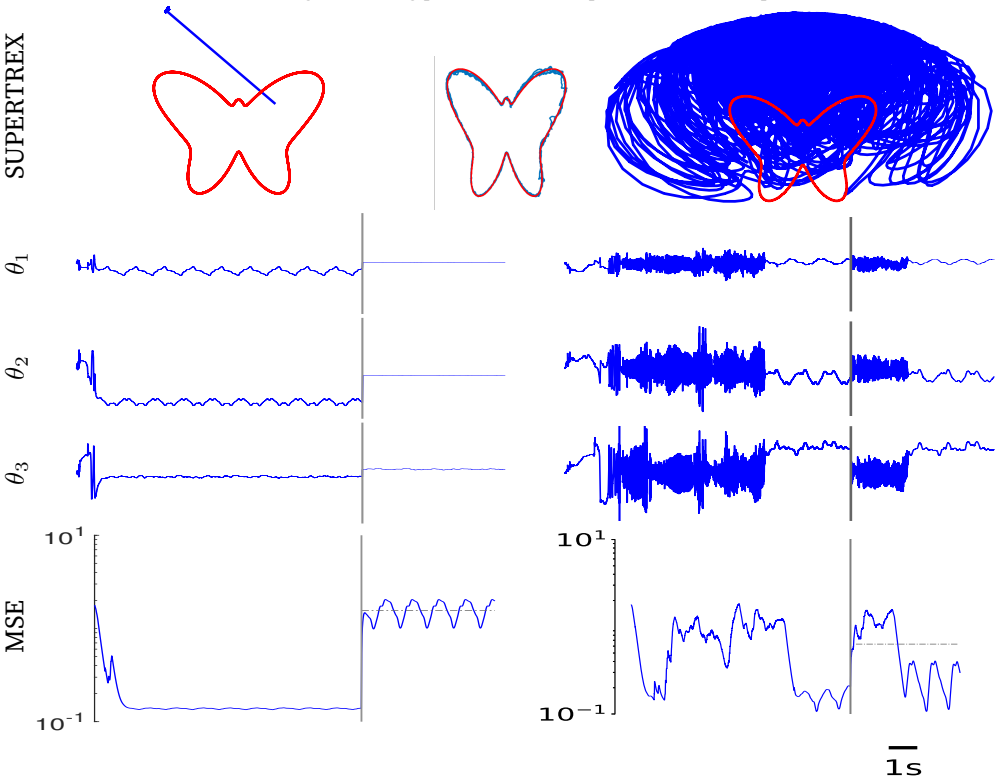
The original scripts and the Python adaptation are able to successfully reproduce the results presented in the paper with the default seed as well as with the 10 arbitrary seeds for the RMHL algorithm, but not for the SUPERTREX algorithm (Figure 3; Table 1, 2).

3 Modification

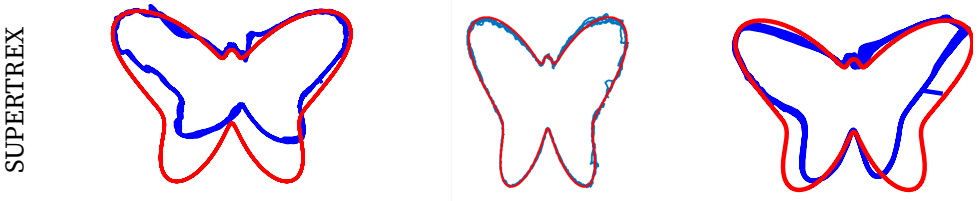
The Python adaptation is a close adaptation of the original MATLAB scripts provided by the authors. However, on simulating their performance on the three tasks, we observed that while, for the first two tasks, the models performed as described in Pyle and Rosenbaum [1], the performance of the SUPERTREX algorithm on Task 3 was not consistent, and was dependent on the seed used for the random number generator. On inspecting



(a) Results for Task 3 with the RMHL algorithm, using the default seed (5489) for the random number generator. The target trajectory is imitated well by the model during the training phase (not shown), however, it poorly maintains the time-series during the testing phase, in both implementations, as presented in [1].



(b) Results for Task 3 with the SUPERTREX algorithm using the default seed (5489) for the random number generator. The target time-series is learned accurately during the training phase, but is not maintained during the testing phase, in both implementations, in contrast to the results presented in [1].



(c) Results for Task 3 with the SUPERTREX algorithm using different implementations (MATLAB, left and Python adaptation, right) and different seeds (295728336, left and 5624282, right) for the random number generator. The target trajectory is learned accurately during the training phase, and is also maintained in a stable manner, with slight divergences (Deviation: 295728336: 0.215 ± 0.073; 5624282: 0.190 ± 0.054) , during the testing phase, in both implementations, similar to the results presented in [1].

Figure 4. Comparison of the performances of original scripts (left column) and Python adaptation (right column) with the results presented in the original article (center column), for RMHL and SUPERTREX, on Task 3 [1]. Each subfigure shows the target trajectory (red) with the trajectory generated by the algorithm (blue) throughout the test phase. In the second subfigure, the the middle rows show the time-series (blue) generated by the model (joint angles (θ_i), in this case). The bottom row shows the distance from target metric (blue) over the simulation (x and y coordinates, in this case), using the log scale for the y axis. The horizontal grey line, in the test phase, indicates the deviation metric. The grey vertical line marks the separation of the training and testing phase.

further, we notice that this was, in some cases, due to the uncontrolled exponential increase in the readout weights.

To look into the robustness of the implementations further, we test the performance of the RMHL and SUPERTREX algorithms on Task 2 with certain modifications to the task parameters, specifically, the number of arm segments and the length of the arm segments. It would be expected for the behaviour to be comparable with the performance on the original task performance, or undergo a gradual decline. We test Task 2 on the arm parameters, which were used in Task 3, i.e. by increasing the number of arm segments from two to three and changing the length of each arm segment. We observe that RMHL performance is comparable to the original Task 2, wherein the time series is generated during the training phase, but is not maintained beyond (Original scripts: 0.846 ± 0.299, Python adaptation: 0.738 ± 0.256; n=11). On the other hand, simulations of the SUPERTREX model, with 2 out of 11 seeds, were able to produce the target output satisfactorily (Original scripts: 0.016 ± 0.007, Python adaptation: 0.009 ± 0.003; n=2) (Figure 5). However, in simulations with 9 out of 11 seeds, the weights increase exponentially, rendering the simulation unable to progress in a meaningful manner (Table 1, 2).

In order to improve the model performance, make the model more scalable in terms of task parameters, and also more robust (as seen in Task 3, with respect to reproducibility with different random seeds), we introduce two minor alterations.

1. We introduce a compensation factor to the update of the readout weights in the exploratory pathway, inversely proportional to the number of segments. Specifically, when the number of segments is greater than two, we multiply the weight update by $0.1/n\_segs$ for Task 2 and by $0.5/n\_segs$ for Task 3.

2. The SUPERTREX model transfers the information from the exploratory pathway to the mastery pathway, only if the error is consistently below a certain threshold. In the original scripts, this threshold is set at 1.5e-3 for Task 1 and Task 2, while at 1.5e-2 for Task 3. We change the transfer threshold for Task 2 from 1.5e-3 to 1.5e-2.

These slight modifications address the shortcomings we encountered earlier with the performance of SUPERTREX in Task 2 and 3. Alteration #1, by including a compensation factor for the change in number of arm segments, prevents the weights from increasing exponentially, and lets the simulation proceed in a meaningful manner. Alteration #2, by increasing the error threshold governing the transfer of information to the mastery pathway, makes the model more tolerant of fluctuations, while continuing to explore and learn a good solution. Although this does not lead to a critical change for Task 1 (Original scripts: 0.006 ± 0.003, n=11; Modified Python re-implementation: 0.004 ± 0.003, n=11) and Task 2 (Original scripts: 0.011 ± 0.003, n=11; Modified Python re-implementation: 0.010 ± 0.004, n=11), this alteration improves the performance of SUPERTREX on Task 3 (Original scripts: 0.881 ± 0.224, n=11; Modified Python re-implementation: 0.140 ± 0.071, n=11). Simulations with 10 out of 11 seeds had satisfactory performance (Deviation < 0.5), compared to 6 out of 11 simulations for the original scripts. Further, it unlocks the potential for the model to be more scalable. We find that with these alterations, on merely increasing the number of time-steps per training cycle and with no further fine tuning of hyper-parameters, the model is able to proceed without an exponential increase in weights over a wider range of task parameters. For instance, on adding surplus segments with length 0.1 each, the model is able to perform in a satisfactory manner, for most cases, with up to 50 arm segments (Table 3, Figure 6). Better accuracy can be achieved by further fine tuning of the hyper-parameters.

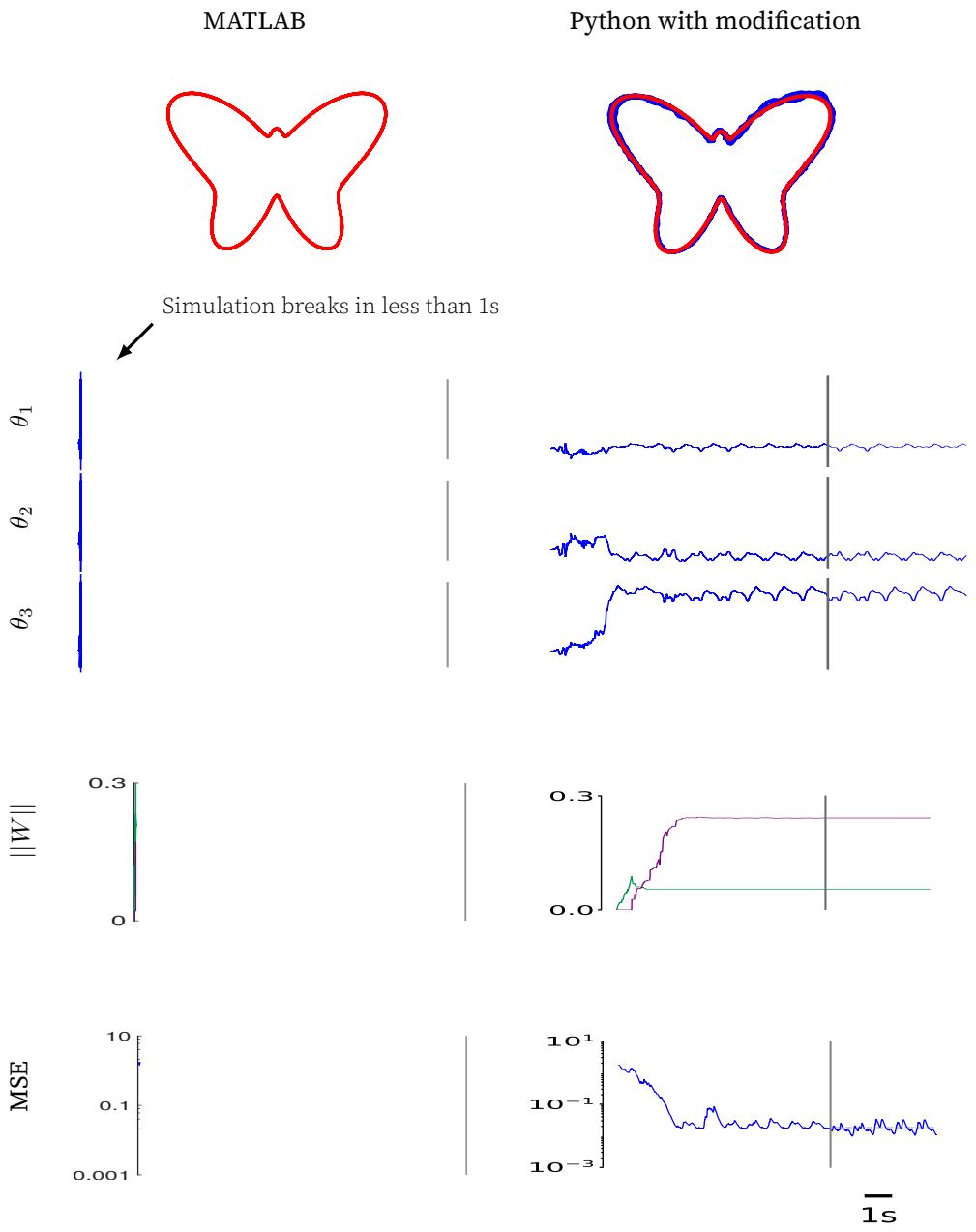


Figure 5. Robustness of the SUPERTREX model on a Task 2 variant. The performance of the original scripts (left column) and modified Python re-implementation (right column) is tested for the SUPERTREX learning algorithm on a variant of Task 2 with increased number of arm segments (lengths: 1.8, 1.2, 0.6). The top panel shows the target trajectory (red) with the trajectory generated by the algorithm (blue) throughout the test phase. The next three rows show the time-series (blue) generated by the model (joint angles (θ_i), in this case). The fourth row shows the progression of the norm of the weight matrix (W_1 in purple; W_2, in green). The bottom row shows the distance from target metric (blue) over the simulation, using the log scale for the y axis. The horizontal grey line, in the test phase, indicates the deviation metric. The grey vertical line marks the separation of the training and testing phase. Using the MATLAB scripts, the readout weights increase uncontrollably rendering the model unable to learn. The Python re-implementation, using a compensation factor to harness the weight update, is able to learn and converge to produce the target time-series.

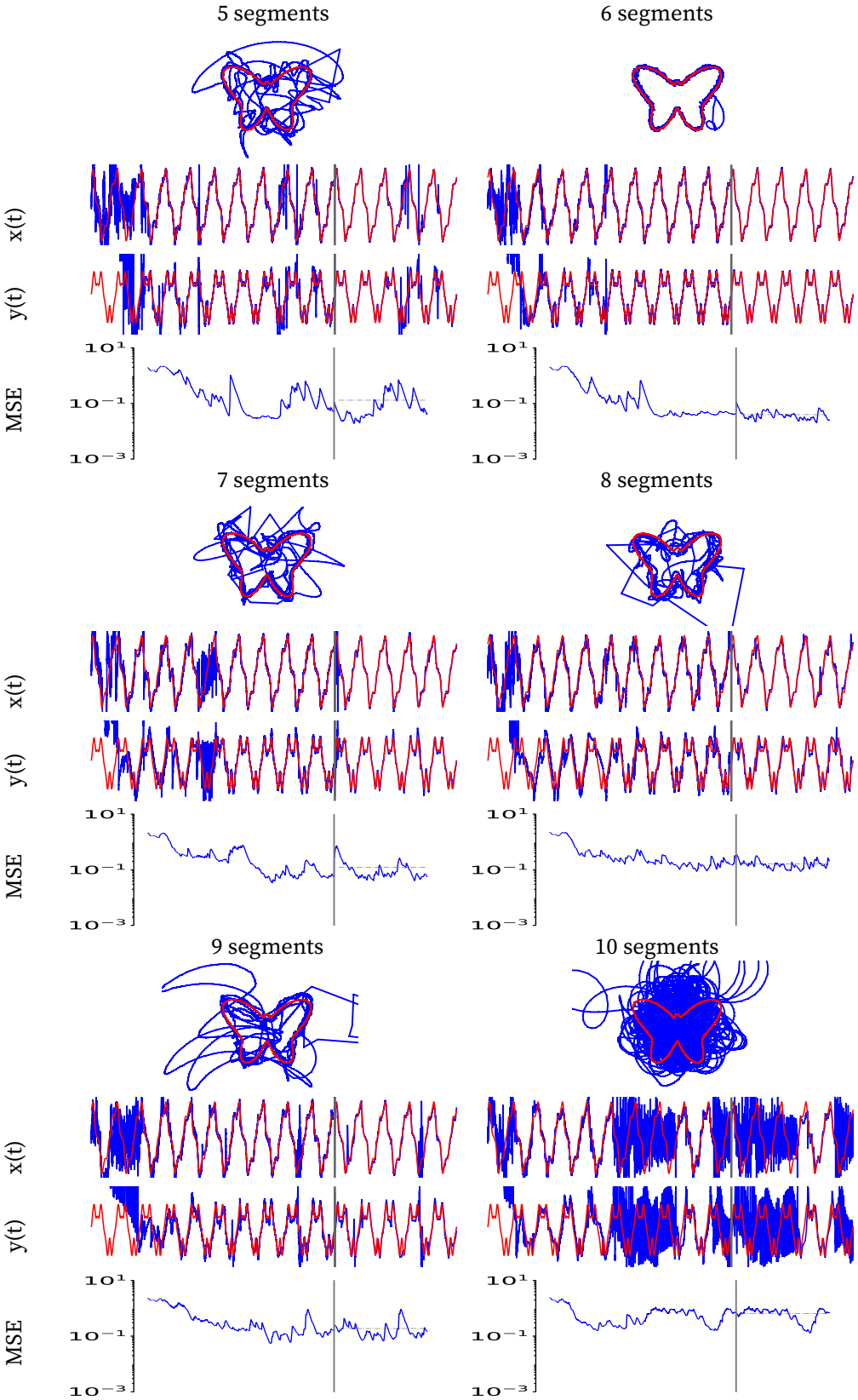


Figure 6. Scalability of the performance of the modified Python re-implementation using the SUPERTREX algorithm on Task 2. The lengths of the arm segments are 1.8, 1.2 and 0.6 for the first three segments (akin to Task 3) and 0.1 for each additional segment. Here, the simulations for Task 2 with 5 to 50 segments are shown, all using the default seed 5489 for the random number generator. (Continued on next page.)

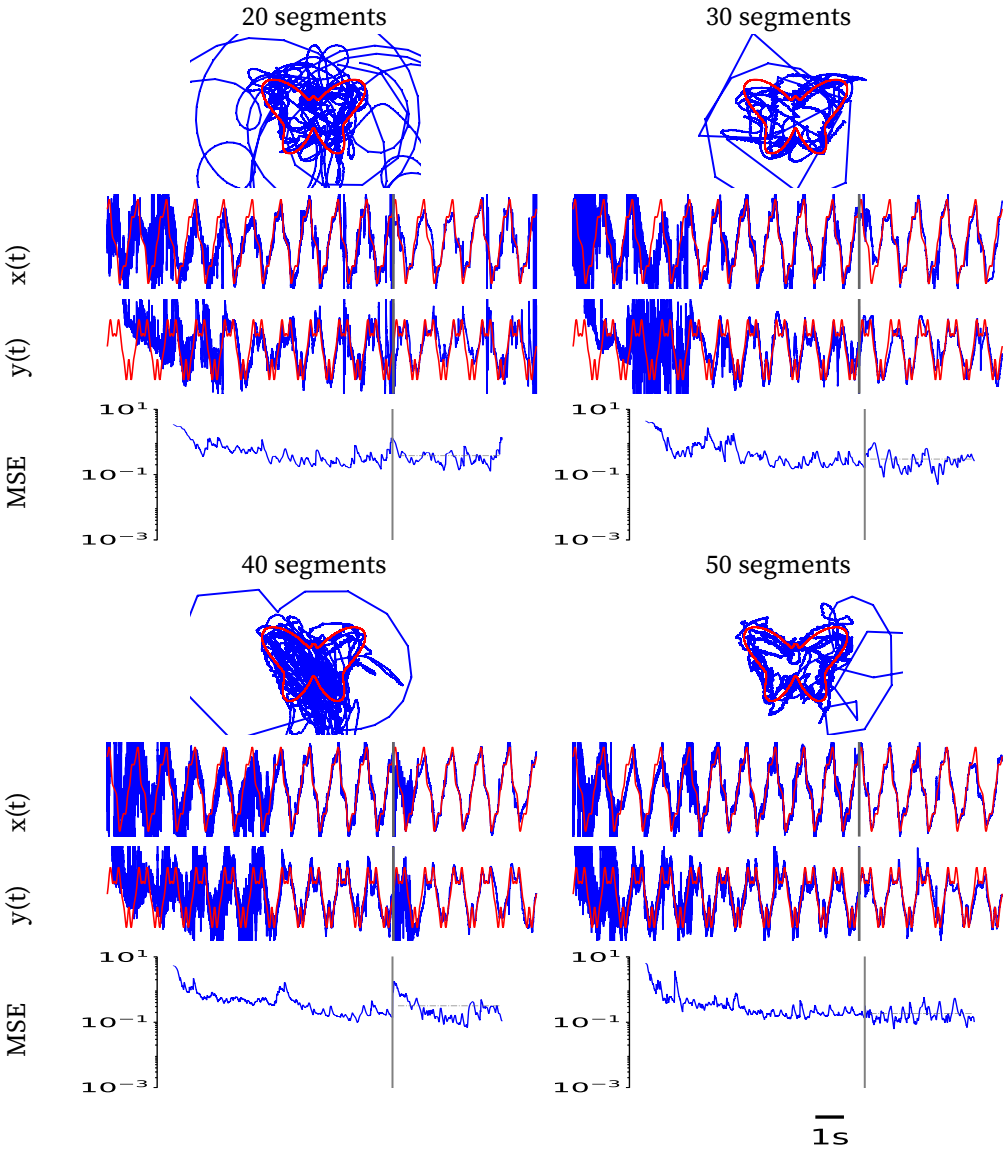


Figure 6. (Continued from previous page.) Scalability of the performance of the modified Python re-implementation using the SUPERTREX algorithm on Task 2. The lengths of the arm segments are 1.8, 1.2 and 0.6 for the first three segments (akin to Task 3) and 0.1 for each additional segment. Here, the simulations for Task 2 with 5 to 50 segments are shown, all using the default seed 5489 for the random number generator. In each subfigure, the top panel shows the produced trajectory, the middle panels show the evolution of the x and y coordinates of the end-effector of the arm (blue) throughout the training and test phase, along with the target coordinates (red). The grey vertical line marks the separation of the training and testing phase. The bottom panel shows the progression of the distance from target metric (blue) over the simulation, using the log scale for the y axis. The horizontal grey line, in the test phase, indicates the deviation metric.

Task 2 variant		Deviation metric		
No. of segments	Time steps	Mean	Median	Standard Deviation
3	10000	0.057	0.032	0.069
4	10000	0.242	0.221	0.136
5	10000	0.141	0.080	0.147
6	10000	0.181	0.160	0.133
7	10000	0.173	0.129	0.130
8	10000	0.253	0.230	0.137
9	10000	0.331	0.297	0.168
10	10000	0.417	0.409	0.151
15	10000	0.538	0.512	0.179
20	15000	0.366	0.324	0.188
30	20000	0.549	0.489	0.236
40	20000	0.372	0.313	0.228
50	30000	0.375	0.283	0.281

Table 3. Deviation metric showing the performance of the modified Python implementation on increasing number of segments for Task 2. Each variant is simulated with the default seed (5489) and ten additional seeds. The mean, median and standard deviation of the deviation metric over these eleven simulations are tabulated here.

4 Discussion

In this article, we discussed the SUPERTREX model presented by Pyle and Rosenbaum [1]. We compared the results presented in the paper, both with the results obtained using the original scripts, and with our modular and user-friendly Python adaptation. Furthermore, we were able to improve the robustness and scalability of the model with two minor alterations.

The Python adaptation strives to be a close adaptation of the original scripts in MATLAB and differs mainly in the method of initialisation of the reservoir connectivity matrix. This is due to the usage of the function `sprandn` in the original scripts, whose internal implementation is not freely available. Figure 7 shows that, on importing initialisation matrix from MATLAB, the exact same results can be obtained in the Python adaptation, as well. Most of the details for the implementation of the models are also described in the paper. Only two necessary details were missing, both concerning the update of the readout weights of the exploratory pathway in the RMHL and SUPERTREX models: One, the inclusion of a crucial learning rate of 0.0005, for Tasks 1-3, and two, an additional compensatory factor of 0.5, for Task 3. There is another discrepancy in the function $psi(x)$ for Task 3. The scripts provided by the authors use a factor of 0.005, whereas the article mentions this factor to be 0.025.

The three algorithms (FORCE, RMHL and SUPERTREX) have been tested on three tasks, presented in Pyle and Rosenbaum [1]. For Task 1 and 2, we verify that the three algorithms function as presented in the paper, and validate that our Python re-implementation produces comparable results. For Task 3, the SUPERTREX model's behaviour is also reproducible, although the performance is dependent on the seed used for the random number generator. Furthermore, we observed that this implementation is quite sensitive to changes in task parameters, such as the number of arms. This was due to the uninhibited increase in the readout weights. We propose the inclusion of a compensation factor for the number of arm segments, which inhibits the growth of the readout weights, and allows the simulation to proceed in a meaningful manner. This considerably improves the robustness and the scalability of the original model.

We conclude that the results presented in the paper are reproducible for two tasks, using the original MATLAB scripts provided by the authors, and also, replicable in Python for all tasks with comparable performance.

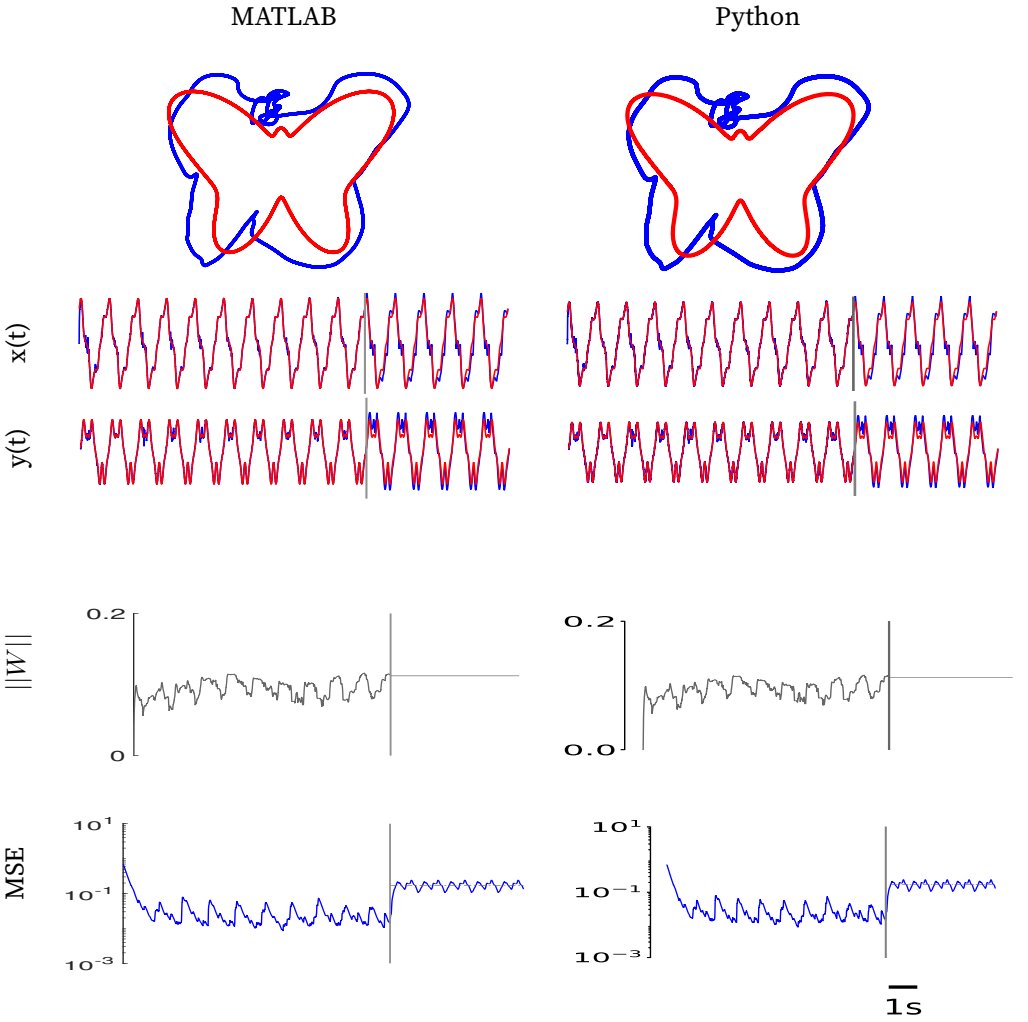


Figure 7. Similarity between the original scripts and the Python adaptation. The performance of the original scripts (left column) and the Python adaptation (right column) is tested for the RMHL learning algorithm on a Task 1. The reservoir connectivity matrix for the Python simulation was initialised using the MATLAB equivalent. Using this initialisation, the progression of the Python simulation is identical to that of the MATLAB simulation. The top panel shows the target trajectory (red) with the trajectory generated by the model (blue) throughout the test phase. The next two rows show the time-series (blue) generated by the model (x and y coordinates, in this case). The third row shows the progression of the norm of the weight matrix. The bottom row shows the distance from target metric (blue) over the simulation, using the log scale for the y axis. The horizontal grey line, in the test phase, indicates the deviation metric. The grey vertical line marks the separation of the training and testing phase.

References

1. R. Pyle and R. Rosenbaum. "A reservoir computing model of reward-modulated motor learning and automaticity." In: **Neural computation** 31.7 (2019), pp. 1430–1461.
2. D. Sussillo and L. F. Abbott. "Generating coherent patterns of activity from chaotic neural networks." In: **Neuron** 63.4 (2009), pp. 544–557.
3. G. M. Hoerzer, R. Legenstein, and W. Maass. "Emergence of complex computational structures from chaotic neural networks through reward-modulated Hebbian learning." In: **Cerebral cortex** 24.3 (2014), pp. 677–690.
4. M. S. Brainard and A. J. Doupe. "What songbirds teach us about learning." In: **Nature** 417.6886 (2002), pp. 351–358.
5. B. P. Ölveczky, T. M. Otchy, J. H. Goldberg, D. Aronov, and M. S. Fee. "Changes in the neural control of a complex motor sequence during learning." In: **Journal of neurophysiology** 106.1 (2011), pp. 386–397.
6. M. Matsumoto and T. Nishimura. "Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator." In: **ACM transactions on modeling and computer simulation (TOMACS)** 8.1 (1998), pp. 3–30.

List of Publications

1. R. Sankar, A. Leblois, and N. P. Rougier. "Dual pathway architecture underlying vocal learning in songbirds", 2022. DOI: 10.1101/2022.04.02.486814. URL: https://doi.org/10.1101/2022.04.02.486814. Awarded **Best Paper** at IEEE International Conference on Development and Learning 2022 (ICDL 2022).
2. R. Sankar, N. P. Rougier, and A. Leblois. "Computational benefits of structural plasticity, illustrated in songbirds". *Neuroscience and Biobehavioral Reviews* 132, 2022, pp. 1183–1196. DOI: 10.1016/j.neubiorev.2021.10.033. URL: https://doi.org/10.1016/j.neubiorev.2021.10.033.
3. R. Sankar, N. Thou, N. P. Rougier, and A. Leblois. "[Re] A Reservoir Computing Model of Reward-Modulated Motor Learning and Automaticity". *ReScience C* 7:11, 1 2021. DOI: 10.5281/ZENODO.5718075. URL: https://zenodo.org/record/5718075.

List of Figures

List of Tables

Bibliography

1. W. C. Abraham and J. M. Williams. "Properties and Mechanisms of LTP Maintenance". *The Neuroscientist* 9:6, 2003, pp. 463–474. DOI: 10.1177/1073858403259119. URL: https://doi.org/10.1177/1073858403259119.
2. J. Alliende, N. Giret, L. Pidoux, C. Del Negro, and A. Leblois. "Seasonal plasticity of song behavior relies on motor and syntactic variability induced by a basal ganglia–forebrain circuit". *Neuroscience* 359, 2017, pp. 49–68. ISSN: 18737544. DOI: 10.1016/j.neuroscience.2017.07.007. URL: http://dx.doi.org/10.1016/j.neuroscience.2017.07.007.
3. A. Alvarez-Buylla, M. Theelen, and F. Nottebohm. "Birth of projection neurons in the higher vocal center of the canary forebrain before, during, and after song learning." *Proceedings of the National Academy of Sciences of the United States of America* 85:22, 1988, pp. 8722–8726. ISSN: 00278424. DOI: 10.1073/pnas.85.22.8722.
4. A. Alvarez-Buylla and J. R. Kirn. "Birth, migration, incorporation, and death of vocal control neurons in adult songbirds". *Journal of Neurobiology* 33:5, 1997, pp. 585–601. ISSN: 00223034. DOI: 10.1002/(SICI)1097-4695(19971105)33:5<585::AID-NEU7>3.0.CO;2-0.
5. B. A. Alward, J. Balthazart, and G. F. Ball. "Differential effects of global versus local testosterone on singing behavior and its underlying neural substrate". *Proceedings of the National Academy of Sciences of the United States of America* 110:48, 2013, pp. 19573–19578. ISSN: 00278424. DOI: 10.1073/pnas.1311371110.
6. A. Amador, Y. S. Perl, G. B. Mindlin, and D. Margoliash. "Elemental gesture dynamics are encoded by song premotor cortical neurons". *Nature* 495:7439, 2013, pp. 59–64. DOI: 10.1038/nature11967.
7. N. Amin, A. Doupe, and F. E. Theunissen. "Development of selectivity for natural sounds in the songbird auditory forebrain". *Journal of Neurophysiology* 97:5, 2007, pp. 3517–3531. ISSN: 00223077. DOI: 10.1152/jn.01066.2006.
8. A. S. Andalman and M. S. Fee. "A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors". *Proceedings of the National Academy of Sciences of the United States of America* 106:30, 2009, pp. 12518–12523. ISSN: 00278424. DOI: 10.1073/pnas.0903214106.

9. A. S. Andalman and M. S. Fee. "A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors". *Proceedings of the National Academy of Sciences* 106:30, 2009, pp. 12518–12523. DOI: 10.1073/pnas.0903214106.

10. P. J. Angeline, G. M. Saunders, and J. B. Pollack. "An evolutionary algorithm that constructs recurrent neural networks". *IEEE Transactions on Neural Networks* 5:1, 1994, pp. 54–65. URL: https://ieeexplore.ieee.org/document/265960.

11. A. P. Arnold. "The effects of castration on song development in zebra finches (Poephila guttata)". *Journal of Experimental Zoology* 191:2, 1975, pp. 261–277. DOI: 10.1002/jez.1401910212. URL: https://doi.org/10.1002/jez.1401910212.

12. D. Aronov, A. S. Andalman, and M. S. Fee. "A specialized forebrain circuit for vocal babbling in the juvenile songbird". *Science* 320:5876, 2008, pp. 630–634. ISSN: 00368075. DOI: 10.1126/science.1155140.

13. F. G. Ashby, B. O. Turner, and J. C. Horvitz. "Cortical and basal ganglia contributions to habit learning and automaticity". *Trends in cognitive sciences* 14:5, 2010, pp. 208–215.

14. G. F. Ball and J. Balthazart. "Seasonal and hormonal modulation of neurotransmitter systems in the song control circuit". *Journal of Chemical Neuroanatomy* 39:2, 2010, pp. 82–95. ISSN: 08910618. DOI: 10.1016/j.jchemneu.2009.08.005.

15. S. A. Bamford, A. F. Murray, and D. J. Willshaw. "Synaptic rewiring for topographic mapping and receptive field development". *Neural Networks* 23:4, 2010, pp. 517–527. DOI: 10.1016/j.neunet.2010.01.005. URL: https://doi.org/10.1016/j.neunet.2010.01.005.

16. A. Belmeguenai, E. Hosy, F. Bengtsson, C. M. Pedroarena, C. Piochon, E. Teuling, Q. He, G. Ohtsuki, M. T. De Jeu, Y. Elgersma, et al. "Intrinsic plasticity complements long-term potentiation in parallel fiber input gain control in cerebellar Purkinje cells". *Journal of Neuroscience* 30:41, 2010, pp. 13630–13643.

17. F. Benureau. "Self Exploration of Sensorimotor Spaces in Robots." Theses. Université de Bordeaux, 2015. URL: https://tel.archives-ouvertes.fr/tel-01251324.

18. Y. Bernardinelli, I. Nikonenko, and D. Muller. "Structural plasticity: Mechanisms and contribution to developmental psychiatric disorders". *Frontiers in Neuroanatomy* 8:November, 2014, pp. 1–9. ISSN: 16625129. DOI: 10.3389/fnana.2014.00123.

19. K. P. Berry and E. Nedivi. "Experience-dependent structural plasticity in the visual system". *Annual review of vision science* 2, 2016, pp. 17–35.

20. G.-q. Bi and M.-m. Poo. "Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type". *Journal of neuroscience* 18:24, 1998, pp. 10464–10472.

21. D. Blalock, J. J. G. Ortiz, J. Frankle, and J. Guttag. *What is the State of Neural Network Pruning?* 2020. arXiv: 2003.03033 [cs.LG]. URL: arXiv:2003.03033.

22. T. V. Bliss and T. Lomo. "Long-lasting potentiation of synaptic transmission in the dentate area of the unanaesthetized rabbit following stimulation of the perforant path". *The Journal of Physiology* 232:2, 1973, pp. 331–356. ISSN: 14697793. DOI: 10.1113/jphysiol.1973.sp010274.

23. P. A. Bogdan, A. G. D. Rowley, O. Rhodes, and S. B. Furber. "Structural Plasticity on the SpiNNaker Many-Core Neuromorphic System". *Frontiers in Neuroscience* 12, 2018. DOI: 10.3389/fnins.2018.00434. URL: https://doi.org/10.3389/fnins.2018.00434.

24. J. R. Boivin, D. J. Piekarski, A. W. Thomas, and L. Wilbrecht. "Adolescent pruning and stabilization of dendritic spines on cortical layer 5 pyramidal neurons do not depend on gonadal hormones". *Developmental Cognitive Neuroscience* 30:January, 2018, pp. 100–107. ISSN: 18789307. DOI: 10.1016/j.dcn.2018.01.007. URL: https://doi.org/10.1016/j.dcn.2018.01.007.

25. B. Bontempi, C. Laurent-Demir, C. Destrade, and R. Jaffard. "Time-dependent reorganization of brain circuitry underlying long-term memory storage". *Nature* 400:6745, 1999, pp. 671–675. DOI: 10.1038/23270. URL: https://doi.org/10.1038/23270.

26. T. Boraud, A. Leblois, and N. P. Rougier. "A natural history of skills". *Progress in Neurobiology* 171, 2018, pp. 114–124. DOI: 10.1016/j.pneurobio.2018.08.003. URL: https://doi.org/10.1016/j.pneurobio.2018.08.003.

27. T. Boraud, A. Leblois, and N. P. Rougier. "A natural history of skills". *Progress in Neurobiology* 171, 2018, pp. 114–124. ISSN: 0301-0082. DOI: 10.1016/j.pneurobio.2018.08.003.

28. M. Bosch and Y. Hayashi. "Structural plasticity of dendritic spines". *Current opinion in neurobiology* 22:3, 2012, pp. 383–388.

29. S. W. Bottjer, S. L. Glaessner, and A. P. Arnold. "Ontogeny of brain nuclei controlling song learning and behavior in zebra finches". *Journal of Neuroscience* 5:6, 1985, pp. 1556–1562. DOI: https://doi.org/10.1523/JNEUROSCI.05-06-01556.1985.

30. S. W. Bottjer and F. Johnson. "Circuits, hormones, and learning: Vocal behavior in songbirds". *Journal of Neurobiology* 33:5, 1997, pp. 602–618. ISSN: 00223034. DOI: 10.1002/(SICI)1097-4695(19971105)33:5<602::AID-NEU8>3.0.CO;2-8.

31. S. W. Bottjer and F. Johnson. "Circuits, hormones, and learning: Vocal behavior in songbirds". *Journal of Neurobiology* 33:5, 1997, pp. 602–618. DOI: 10.1002/(sici)1097-4695(19971105)33:5<602::aid-neu8>3.0.co;2-8.

32. S. W. Bottjer, E. A. Miesner, and A. P. Arnold. "Changes in neuronal number, density and size account for increases in volume of song-control nuclei during song development in zebra finches". *Neuroscience Letters* 67:3, 1986, pp. 263–268. ISSN: 03043940. DOI: 10.1016/0304-3940(86)90319-8.

33. S. W. Bottjer, E. A. Miesner, and A. P. Arnold. "Forebrain Lesions Disrupt Development But Not Maintenance of Song in Passerine Birds". *Science* 224:4651, 1984, pp. 901–903. URL: https://science.sciencemag.org/content/224/4651/901.

34. L. Bottou et al. "Stochastic gradient learning in neural networks". *Proceedings of Neuro-Nimes* 91:8, 1991, p. 12.

35. P. H. Boucherie, M. C. Loretto, J. J. Massen, and T. Bugnyar. "What constitutes "social complexity" and "social intelligence" in birds? Lessons from ravens". *Behavioral Ecology and Sociobiology* 73:1, 2019. ISSN: 03405443. DOI: 10.1007/s00265-018-2607-2.

36. M. S. Brainard and A. J. Doupe. "Auditory feedback in learning and maintenance of vocal behaviour". *Nature Reviews Neuroscience* 1:1, 2000, pp. 31–40. DOI: 10.1038/35036205. URL: https://doi.org/10.1038/35036205.

37. M. S. Brainard and A. J. Doupe. "Translating Birdsong: Songbirds as a Model for Basic and Applied Medical Research". *Annual Review of Neuroscience* 36:1, 2013, pp. 489–517. ISSN: 0147-006X. DOI: 10.1146/annurev-neuro-060909-152826.

38. M. S. Brainard and A. J. Doupe. "What songbirds teach us about learning". *Nature* 417, 2002, pp. 351–417. ISSN: 0033524X. DOI: 10.4324/9781315709307-44.

39. E. A. Brenowitz. "Comparative approaches to the avian song system". *Journal of Neurobiology* · 4695:December 1997, 1997. DOI: 10.1002/(SICI)1097-4695(19971105)33.

40. E. A. Brenowitz. "Plasticity of the adult avian song control system". *Annals of the New York Academy of Sciences* 1016, 2004, pp. 560–585. ISSN: 00778923. DOI: 10.1196/annals.1298.006.

41. E. Bufill, J. Agustí, and R. Blesa. "Human neoteny revisited: The case of synaptic plasticity". *American Journal of Human Biology* 23:6, 2011, pp. 729–739. ISSN: 10420533. DOI: 10.1002/ajhb.21225.

42. D. V. Buonomano and M. M. Merzenich. "Cortical plasticity: From synapses to maps". *Annual Review of Neuroscience* 21, 1998, pp. 149–186. ISSN: 0147006X. DOI: 10.1146/annurev.neuro.21.1.149.

43. A. B. Butler, A. Reiner, and H. J. Karten. "Evolution of the amniote pallium and the origins of mammalian neocortex". *Annals of the New York Academy of Sciences* 1225:1, 2011, pp. 14–27. ISSN: 17496632. DOI: 10.1111/j.1749-6632.2011.06006.x.

44. M. Butz and A. van Ooyen. "A simple rule for dendritic spine and axonal bouton formation can account for cortical reorganization after focal retinal lesions". *PLoS computational biology* 9:10, 2013, e1003259.

45. M. Butz, I. D. Steenbuck, and A. van Ooyen. "Homeostatic structural plasticity increases the efficiency of small-world networks". *Frontiers in synaptic neuroscience* 6, 2014, p. 7.

46. M. Butz, F. Wörgötter, and A. van Ooyen. "Activity-dependent structural plasticity". *Brain Research Reviews* 60:2, 2009, pp. 287–305. DOI: 10.1016/j.brainresrev.2008.12.023. URL: https://doi.org/10.1016/j.brainresrev.2008.12.023.

47. G. Buzsáki. "Two-stage model of memory trace formation: A role for "noisy" brain states". *Neuroscience* 31:3, 1989, pp. 551–570. DOI: 10.1016/0306-4522(89)90423-5. URL: https://doi.org/10.1016/0306-4522(89)90423-5.

48. S. R. y. Cajal. *Degeneration & regeneration of the nervous system*. Oxford University Press, 1928. URL: https://books.google.fr/books/about/Degeneration_Regeneration_of_the_Nervous.html?id=ML9rAAAAMAAJ&redir_esc=y.

49. A. Calabrese and S. M. Woolley. "Coding principles of the canonical cortical microcircuit in the avian brain". *Proceedings of the National Academy of Sciences of the United States of America* 112:11, 2015, pp. 3517–3522. ISSN: 10916490. DOI: 10.1073/pnas.1408545112.

50. D. Caligiore, M. A. Arbib, R. C. Miall, and G. Baldassarre. "The super-learning hypothesis: Integrating learning processes across cortex, cerebellum and basal ganglia". *Neuroscience and Biobehavioral Reviews* 100, 2019, pp. 19–34. DOI: 10.1016/j.neubiorev.2019.02.008. URL: https://doi.org/10.1016/j.neubiorev.2019.02.008.

51. D. Caligiore, G. Pezzulo, et al. "Consensus Paper: Towards a Systems-Level View of Cerebellar Function: the Interplay Between Cerebellum, Basal Ganglia, and Cortex". *The Cerebellum* 16:1, 2016, pp. 203–229. DOI: 10.1007/s12311-016-0763-3. URL: https://doi.org/10.1007/s12311-016-0763-3.

52. E. Campanac and D. Debanne. "Spike timing-dependent plasticity: a learning rule for dendritic integration in rat CA1 pyramidal neurons". *The Journal of physiology* 586:3, 2008, pp. 779–793.

53. M. L. Caras, M. O'Brien, E. A. Brenowitz, and E. W. Rubel. "Estradiol selectively enhances auditory function in avian forebrain neurons". *Journal of Neuroscience* 32:49, 2012, pp. 17597–17611. ISSN: 02706474. DOI: 10.1523/JNEUROSCI.3938-12.2012.

54. S. Carey and E. Bartlett. "Acquiring a single new word.", 1978.

55. G. D. Carrillo and A. J. Doupe. "Is the Songbird Area X Striatal, Pallidal, or Both? An Anatomical Study". *Journal of Comparative Neurology* 473:3, 2004, pp. 415–437. ISSN: 00219967. DOI: 10.1002/cne.20099.

56. C. K. Catchpole and P. J. Slater. *Bird song: biological themes and variations*. Cambridge university press, 2003.

57. G. Chechik, I. Meilijson, and E. Ruppin. "Synaptic pruning in development: a computational account". *Neural computation* 10:7, 1998, pp. 1759–1777.

58. A. N. Chen and C. D. Meliza. "Experience-and sex-dependent intrinsic plasticity in the zebra finch auditory cortex during song memorization". *Journal of Neuroscience* 40:10, 2020, pp. 2047–2055.

59. H. Chen, L. Yang, Y. Xu, G.-y. Wu, J. Yao, J. Zhang, Z.-r. Zhu, Z.-a. Hu, J.-f. Sui, and B. Hu. "Prefrontal Control of Cerebellum-Dependent Associative Motor Learning". *The Cerebellum* 13:1, 2013, pp. 64–78. DOI: 10.1007/s12311-013-0517-4. URL: https://doi.org/10.1007/s12311-013-0517-4.

60. D. B. Chklovskii, B. Mel, and K. Svoboda. "Cortical rewiring and information storage". *Nature* 431:7010, 2004, pp. 782–788.

61. A. Citri and R. C. Malenka. "Synaptic plasticity: Multiple forms, functions, and mechanisms". *Neuropsychopharmacology* 33:1, 2008, pp. 18–41. ISSN: 0893133X. DOI: 10.1038/sj.npp.1301559.

62. G. Cornez, E. Jonckers, S. M. ter Haar, A. V. der Linden, C. A. Cornil, and J. Balthazart. "Timing of perineuronal net development in the zebra finch song control system correlates with developmental song learning". *Proceedings of the Royal Society B: Biological Sciences* 285:1883, 2018, p. 20180849. DOI: 10.1098/rspb.2018.0849. URL: https://doi.org/10.1098/rspb.2018.0849.

63. A. Daou and D. Margoliash. "Intrinsic neuronal properties represent song and error in zebra finch vocal learning". *Nature communications* 11:1, 2020, pp. 1–17.

64. A. Daou and D. Margoliash. "Intrinsic plasticity and birdsong learning". *Neurobiology of Learning and Memory* 180, 2021, p. 107407.

65. R. Darshan, W. E. Wood, S. Peters, A. Leblois, and D. Hansel. "A canonical neural mechanism for behavioral variability". *Nature Communications* 8:1, 2017. DOI: 10.1038/ncomms15415. URL: https://doi.org/10.1038/ncomms15415.

66. A. Datwani, T. Iwasato, S. Itohara, and R. S. Erzurumlu. "NMDA Receptor-Dependent Pattern Transfer from Afferents to Postsynaptic Cells and Dendritic Differentiation in the Barrel Cortex". *Molecular Cell Neuroscience* 21:3, 2002, pp. 477–492. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf.

67. G. De Groof, M. Verhoye, V. Van Meir, J. Balthazart, and A. Van Der Linden. "Seasonal rewiring of the songbird brain: An in vivo MRI study". *European Journal of Neuroscience* 28:12, 2008, pp. 2475–2485. ISSN: 0953816X. DOI: 10.1111/j.1460-9568.2008.06545.x.

68. M. De Roo, P. Klauser, and D. Muller. "LTP promotes a selective long-term stabilization and clustering of dendritic spines". *PLoS Biology* 6:9, 2008, pp. 1850–1860. ISSN: 15449173. DOI: 10.1371/journal.pbio.0060219.

69. D. Debanne. "Plasticity of neuronal excitability in vivo". *Journal of Physiology* 587:13, 2009, pp. 3057–3058. ISSN: 00223751. DOI: 10.1113/jphysiol.2009.175448.

70. C. Del Negro and J. M. Edeline. "Sex and season influence the proportion of thin spike cells in the canary HVc". *NeuroReport* 13:16, 2002, pp. 2005–2009. ISSN: 09594965. DOI: 10.1097/00001756-200211150-00003.

71. S. Derégnaucourt, P. P. Mitra, O. Fehér, C. Pytte, and O. Tchernichovski. "How sleep affects the developmental learning of bird song". *Nature* 433:7027, 2005, pp. 710–716. ISSN: 1476-4687. DOI: 10.1038/nature03275.

72. A. K. Dhawale, M. A. Smith, and B. P. Ölveczky. "The Role of Variability in Motor Learning". *Annual Review of Neuroscience* 40:1, 2017, pp. 479–498. DOI: 10.1146/annurev-neuro-072116-031548. URL: https://doi.org/10.1146/annurev-neuro-072116-031548.

73. L. Ding and D. J. Perkel. "Long-term potentiation in an avian basal ganglia nucleus essential for vocal learning". *Journal of Neuroscience* 24:2, 2004, pp. 488–494. DOI: 10.1523/jneurosci.4358-03.2004. URL: https://doi.org/10.1523/jneurosci.4358-03.2004.

74. L. Ding and D. J. Perkel. "Long-Term Potentiation in an Avian Basal Ganglia Nucleus Essential for Vocal Learning". *Journal of Neuroscience* 24:2, 2004, pp. 488–494. ISSN: 02706474. DOI: 10.1523/JNEUROSCI.4358-03.2004.

75. A. J. Doupe and P. K. Kuhl. "BIRDSONG AND HUMAN SPEECH: Common Themes and Mechanisms". *Annual Review of Neuroscience* 22:1, 1999, pp. 567–631. ISSN: 0147-006X. DOI: 10.1146/annurev.neuro.22.1.567.

76. A. J. Doupe, D. J. Perkel, A. Reiner, and E. A. Stern. "Birdbrains could teach basal ganglia research a new song". *Trends in Neurosciences* 28:7, 2005, pp. 353–363. ISSN: 01662236. DOI: 10.1016/j.tins.2005.05.005.

77. P. A. Downing, A. S. Griffin, and C. K. Cornwallis. "Group formation and the evolutionary pathway to complex sociality in birds". *Nature Ecology and Evolution* 4:3, 2020, pp. 479–486. ISSN: 2397334X. DOI: 10.1038/s41559-020-1113-x.

78. K. Doya. "Complementary roles of basal ganglia and cerebellum in learning and motor control". *Current Opinion in Neurobiology* 10:6, 2000, pp. 732–739. DOI: 10.1016/s0959-4388(00)00153-7. URL: https://doi.org/10.1016/s0959-4388(00)00153-7.

79. K. Doya and T. J. Sejnowski. "A Computational Model of Birdsong Learning by Auditory Experience and Auditory Feedback". In: *Central Auditory Processing and Neural Modeling*. Springer US, 1998, pp. 77–88. DOI: 10.1007/978-1-4615-5351-9_8. URL: https://doi.org/10.1007/978-1-4615-5351-9_8.

80. J. Doyon, V. Penhune, and L. G. Ungerleider. "Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning". *Neuropsychologia* 41:3, 2003, pp. 252–262. DOI: 10.1016/s0028-3932(02)00158-6. URL: https://doi.org/10.1016/s0028-3932(02)00158-6.

81. H. Eichenbaum. "Time cells in the hippocampus: a new dimension for mapping memories". *Nature Reviews Neuroscience* 15:11, 2014, pp. 732–744. DOI: 10.1038/nrn3827. URL: https://doi.org/10.1038/nrn3827.

82. T. Elliott and N. R. Shadbolt. "A Neurotrophic Model of the Development of the Retinogeniculocortical Pathway Induced by Spontaneous Retinal Waves". *The Journal of Neuroscience* 19:18, 1999, pp. 7951–7970. DOI: 10.1523/jneurosci.19-18-07951.1999. URL: https://doi.org/10.1523/jneurosci.19-18-07951.1999.

83. P. S. Eriksson, E. Perfilieva, T. Björk-Eriksson, A.-M. Alborn, C. Nordborg, D. A. Peterson, and F. H. Gage. "Neurogenesis in the adult human hippocampus". *Nature medicine* 4:11, 1998, pp. 1313–1317.

84. B. J. Everitt and M. E. Wolf. "Psychomotor Stimulant Addiction: A Neural Systems Perspective". *The Journal of Neuroscience* 22:9, 2002, pp. 3312–3320. DOI: 10.1523/jneurosci.22-09-03312.2002. URL: https://doi.org/10.1523/jneurosci.22-09-03312.2002.

85. M. A. Farries and A. L. Fairhall. "Reinforcement Learning With Modulated Spike Timing–Dependent Synaptic Plasticity". *Journal of Neurophysiology* 98:6, 2007, pp. 3648–3665. DOI: 10.1152/jn.00364.2007. URL: https://doi.org/10.1152/jn.00364.2007.

86. M. A. Farries and D. J. Perkel. "A Telencephalic Nucleus Essential for Song Learning Contains Neurons with Physiological Characteristics of Both Striatum and Globus Pallidus". *Journal of Neuroscience* 22:9, 2002, pp. 3776–3787. ISSN: 02706474. DOI: 10.1523/jneurosci.22-09-03776.2002.

87. M. Fee and J. Goldberg. "A hypothesis for basal ganglia-dependent reinforcement learning in the songbird". *Neuroscience* 198, 2011, pp. 152–170. DOI: 10.1016/j.neuroscience.2011.09.069. URL: https://doi.org/10.1016/j.neuroscience.2011.09.069.

88. M. S. Fee and A. Leonardo. "Miniature motorized microdrive and commutator system for chronic neural recording in small animals". *Journal of Neuroscience Methods* 112:2, 2001, pp. 83–94. DOI: 10.1016/s0165-0270(01)00426-5. URL: https://doi.org/10.1016/s0165-0270(01)00426-5.

89. M. S. FEE, A. A. KOZHEVNIKOV, and R. H. HAHNLOSER. "Neural Mechanisms of Vocal Sequence Generation in the Songbird". *Annals of the New York Academy of Sciences* 1016:1, 2004, pp. 153–170. DOI: 10.1196/annals.1298.022. URL: https://doi.org/10.1196/annals.1298.022.

90. M. S. Fee and C. Scharff. "The songbird as a model for the generation and learning of complex sequential behaviors". *ILAR Journal* 51:4, 2010, pp. 362–377. ISSN: 10842020. DOI: 10.1093/ilar.51.4.362.

91. I. R. Fiete, M. S. Fee, and H. S. Seung. "Model of Birdsong Learning Based on Gradient Estimation by Dynamic Perturbation of Neural Conductances". *Journal of Neurophysiology* 98:4, 2007, pp. 2038–2057. DOI: 10.1152/jn.01311.2006. URL: https://doi.org/10.1152/jn.01311.2006.

92. I. R. Fiete, W. Senn, C. Z. Wang, and R. H. Hahnloser. "Spike-Time-Dependent Plasticity and Heterosynaptic Competition Organize Networks to Produce Long Scale-Free Sequences of Neural Activity". *Neuron* 65:4, 2010, pp. 563–576. ISSN: 08966273. DOI: 10.1016/j.neuron.2010.02.003. URL: http://dx.doi.org/10.1016/j.neuron.2010.02.003.

93. M. P. Forrest, E. Parnell, and P. Penzes. "Dendritic structural plasticity and neuropsychiatric disease". *Nature Reviews Neuroscience* 19:4, 2018, pp. 215–234. DOI: 10.1016/j.physbeh.2017.03.040. URL: file:///C:/Users/Carla%20Carolina/Desktop/Artigos%20para%20acrescentar%20na%20qualifica%7B%5Cc%7Bc%7D%7D%7B%5C~%7Ba%7D%7Do/The%20impact%20of%20birth%20weight%20on%20cardiovascular%20disease%20risk%20in%20the.pdf.

94. B. Fritzke. "A Growing Neural Gas Network Learns Topologies". In: *Advances in Neural Information Processing Systems 7*. Ed. by G. Tesauro, D. S. Touretzky, and T. K. Leen. MIT Press, 1995, pp. 625–632. URL: http://papers.nips.cc/paper/893-a-growing-neural-gas-network-learns-topologies.pdf.

95. V. Gadagkar and J. H. Goldberg. "Dopamine Neurons Encode Performance Error in Singing Birds". *Science* 176:1, 2016, pp. 139–148. DOI: 10.1016/j.physbeh.2017.03.040.

96. V. Gadagkar, P. A. Puzerey, R. Chen, E. Baird-Daniel, A. R. Farhang, and J. H. Goldberg. "Dopamine neurons encode performance error in singing birds". *Science* 354:6317, 2016, pp. 1278–1282. DOI: 10.1126/science.aah6837.

97. J. Garst-Orozco, B. Babadi, and B. P. Ölveczky. "A neural circuit mechanism for regulating vocal variability during song learning in zebra finches". *eLife* 3, 2014. DOI: 10.7554/elife.03697.

98. J. Garst-Orozco, B. Babadi, and B. P. Ölveczky. "A neural circuit mechanism for regulating vocal variability during song learning in zebra finches". *eLife* 3, 2014, e03697. ISSN: 2050084X. DOI: 10.7554/eLife.03697.

99. M. J. Geden, S. E. Romero, and M. Deshmukh. "Apoptosis versus Axon Pruning: Molecular Intersection of Two Distinct Pathways for Axon Degeneration", 2019, pp. 3–8. DOI: 10.1016/j.neures.2018.11.007.Apoptosis.

100. W. Gerstner, R. Kempter, J. L. Van Hemmen, and H. Wagner. "A neuronal learning rule for sub-millisecond temporal coding". *Nature* 383:6595, 1996, pp. 76–78.

101. C. D. Gilbert. "Adult cortical dynamics". *Physiological Reviews* 78:2, 1998, pp. 467–485. ISSN: 00319333. DOI: 10.1152/physrev.1998.78.2.467.

102. N. Giret, J. Kornfeld, S. Ganguli, and R. H. R. Hahnloser. "Evidence for a causal inverse model in an avian cortico-basal ganglia circuit". *Proceedings of the National Academy of Sciences* 111:16, 2014, pp. 6063–6068. DOI: 10.1073/pnas.1317087111. URL: https://doi.org/10.1073/pnas.1317087111.

103. C. M. Glaze and T. W. Troyer. "Development of temporal structure in zebra finch song". *Journal of Neurophysiology* 109:4, 2013, pp. 1025–1035. DOI: 10.1152/jn.00578.2012. URL: https://doi.org/10.1152/jn.00578.2012.

104. "Global Optimization". *Science* 285:5432, 1999, 1325n–1325. DOI: 10.1126/science.285.5432.1325n. URL: https://doi.org/10.1126/science.285.5432.1325n.

105. G. Goh. "Why Momentum Really Works". *Distill*, 2017. DOI: 10.23915/distill.00006. URL: http://distill.pub/2017/momentum.

106. S. A. Goldman and F. Nottebohm. "Neuronal production, migration, and differentiation in a vocal control nucleus of the adult female canary brain". *Proceedings of the National Academy of Sciences of the United States of America* 80:8 I, 1983, pp. 2390–2394. ISSN: 00278424. DOI: 10.1073/pnas.80.8.2390.

107. M. Gori and A. Tesi. "On the problem of local minima in backpropagation". *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14:1, 1992, pp. 76–86.

108. J. Gottlieb, M. Lopes, and P.-Y. Oudeyer. "Motivated Cognition: Neural and Computational Mechanisms of Curiosity, Attention, and Intrinsic Motivation". In: *Advances in Motivation and Achievement*. Emerald Group Publishing Limited, 2016, pp. 149–172. DOI: 10.1108/s0749-742320160000019017. URL: https://doi.org/10.1108/s0749-742320160000019017.

109. E. Gould. "How widespread is adult neurogenesis in mammals?" *Nature Reviews Neuroscience* 8:6, 2007, pp. 481–488.

110. O. Güntürkün. "The avian 'prefrontal cortex' and cognition". *Current Opinion in Neurobiology* 15:6, 2005, pp. 686–693. ISSN: 09594388. DOI: 10.1016/j.conb.2005.10.003.

111. O. Güntürkün. "The convergent evolution of neural substrates for cognition". *Psychological Research* 76:2, 2012, pp. 212–219. ISSN: 03400727. DOI: 10.1007/s00426-011-0377-9.

112. O. Güntürkün and T. Bugnyar. "Cognition without Cortex". *Trends in Cognitive Sciences* 20:4, 2016, pp. 291–303. ISSN: 1879307X. DOI: 10.1016/j.tics.2016.02.001. URL: http://dx.doi.org/10.1016/j.tics.2016.02.001.

113. S. Haesler, K. Wada, A. Nshdejan, E. E. Morrisey, T. Lints, E. D. Jarvis, and C. Scharff. "FoxP2 Expression in Avian Vocal Learners and Non-Learners". *Journal of Neuroscience* 24:13, 2004, pp. 3164–3175. ISSN: 02706474. DOI: 10.1523/JNEUROSCI.4369-03.2004.

114. R. H. Hahnloser and A. Kotowicz. "Auditory representations and memory in birdsong learning". *Current Opinion in Neurobiology* 20:3, 2010, pp. 332–339. ISSN: 09594388. DOI: 10.1016/j.conb.2010.02.011. URL: http://dx.doi.org/10.1016/j.conb.2010.02.011.

115. R. H. Hahnloser, A. A. Kozhevnikov, and M. S. Fee. "An ultra-sparse code underlies the generation of neural sequences in a songbird". *Nature* 419:65, 2002, p. 69. ISSN: 00280836. DOI: 10.1038/nature01221. URL: https://doi.org/10.1038/nature01221.

116. A. Hanuschkin, S. Ganguli, and R. H. R. Hahnloser. "A Hebbian learning rule gives rise to mirror neurons and links them to control theoretic inverse models". *Frontiers in Neural Circuits* 7, 2013. DOI: 10.3389/fncir.2013.00106. URL: https://doi.org/10.3389/fncir.2013.00106.

117. M. Hartley, N. Taylor, and J. Taylor. "Understanding spike-time-dependent plasticity: A biologically motivated computational model". *Neurocomputing* 69:16-18, 2006, pp. 2005–2016.

118. S. Hélie, S. W. Ell, and F. G. Ashby. "Learning robust cortico-cortical associations with the basal ganglia: An integrative review". *Cortex* 64, 2015, pp. 123–135. ISSN: 19738102. DOI: 10.1016/j.cortex.2014.10.011.

119. T. K. Hensch. "Critical period plasticity in local cortical circuits". *Nature Reviews Neuroscience* 6:11, 2005, pp. 877–888. ISSN: 1471003X. DOI: 10.1038/nrn1787.

120. C. Herold, P. Schlömer, I. Mafoppa-Fomat, J. Mehlhorn, K. Amunts, and M. Axer. "The hippocampus of birds in a view of evolutionary connectomics". *Cortex* 118, 2019, pp. 165–187. ISSN: 19738102. DOI: 10.1016/j.cortex.2018.09.025.

121. K. Herrmann and A. P. Arnold. "The development of afferent projections to the robust archistriatal nucleus in male zebra finches: A quantitative electron microscopic study". *Journal of Neuroscience* 11:7, 1991, pp. 2063–2074. ISSN: 02706474. DOI: 10.1523/jneurosci.11-07-02063.1991.

122. K. Herrmann and H.-J. .-. Bischof. "Delayed development of song control nuclei in the zebra finch is related to behavioral development". *Journal of Comparative Neurology* 245:2, 1986, pp. 167–175. ISSN: 10969861. DOI: 10.1002/cne.902450204.

123. K. M. Hill and T. J. DeVoogd. "Altered daylength affects dendritic structure in a song-related brain region in red-winged blackbirds". *Behavioral and Neural Biology* 56:3, 1991, pp. 240–250. ISSN: 01631047. DOI: 10.1016/0163-1047(91)90379-5.

124. E. Hisey, M. G. Kearney, and R. Mooney. "A common neural circuit mechanism for internally guided and externally reinforced forms of motor learning". *Nature Neuroscience* 21:4, 2018, pp. 589–597. ISSN: 15461726. DOI: 10.1038/s41593-018-0092-6. URL: http://dx.doi.org/10.1038/s41593-018-0092-6.

125. A. J. Holtmaat, J. T. Trachtenberg, L. Wilbrecht, G. M. Shepherd, X. Zhang, G. W. Knott, and K. Svoboda. "Transient and Persistent Dendritic Spines in the Neocortex In Vivo". *Neuron* 45:2, 2005, pp. 279–291. DOI: 10.1016/j.neuron.2005.01.003.

126. D. H. Hubel and T. N. Wiesel. "Brain mechanisms of vision." *Scientific American* 241:3, 1979, pp. 150–162. ISSN: 00368733. DOI: 10.1038/scientificamerican0979-150.

127. P. R. Huttenlocher et al. "Synaptic density in human frontal cortex-developmental changes and effects of aging". *Brain Res* 163:2, 1979, pp. 195–205.

128. P. R. Huttenlocher, C. de Courten, L. J. Garey, and H. Van der Loos. "Synaptogenesis in human visual cortex—evidence for synapse elimination during normal development". *Neuroscience letters* 33:3, 1982, pp. 247–252.

129. K. Immelmann. "Song development in the zebra finch and other estrildid finches". *Bird Vocalizations*, 1969, pp. 61–77. URL: https://cir.nii.ac.jp/crid/1571135649961747968.

130. G. M. Innocenti and D. J. Price. "Exuberance in the development of cortical networks". *Nature Reviews Neuroscience* 6:12, 2005, pp. 955–965. ISSN: 1471003X. DOI: 10.1038/nrn1790.

131. M. Ito. "Error detection and representation in the olivo-cerebellar system". *Frontiers in Neural Circuits* 7, 2013. DOI: 10.3389/fncir.2013.00001. URL: https://doi.org/10.3389/fncir.2013.00001.

132. M. Ito. "Mechanisms of motor learning in the cerebellum11Published on the World Wide Web on 24 November 2000." *Brain Research* 886:1-2, 2000, pp. 237–245. DOI: 10.1016/s0006-8993(00)03142-5. URL: https://doi.org/10.1016/s0006-8993(00)03142-5.

133. R. B. Ivry, S. Keele, and H. Diener. "Dissociation of the lateral and medial cerebellum in movement timing and movement execution". *Experimental brain research* 73:1, 1988, pp. 167–180.

134. S. Iyengar, S. S. Viswanathan, and S. W. Bottjer. "Development of topography within song control circuitry of zebra finches during the sensitive period for song learning". *Journal of Neuroscience* 19:14, 1999, pp. 6037–6057. ISSN: 02706474. DOI: 10.1523/jneurosci.19-14-06037.1999.

135. L. F. Jacobs. "The economy of winter: Phenotypic plasticity in behavior and brain structure". *Biological Bulletin* 191:1, 1996, pp. 92–100. ISSN: 00063185. DOI: 10.2307/1543068.

136. F. Johnson and S. W. Bottjer. "Growth and regression of thalamic efferents in the song-control system of male zebra finches". *Journal of Comparative Neurology* 326:3, 1992, pp. 442–450. ISSN: 10969861. DOI: 10.1002/cne.903260309.

137. M. I. Jordan and D. E. Rumelhart. "Forward Models: Supervised Learning with a Distal Teacher". *Cogn. Sci.* 16:3, 1992, pp. 307–354. DOI: 10.1207/s15516709cog1603_1. URL: https://doi.org/10.1207/s15516709cog1603%5C_1.

138. M. H. Kao, A. J. Doupe, and M. S. Brainard. "Contributions of an avian basal ganglia–forebrain circuit to real-time modulation of song". *Nature* 433:7026, 2005, pp. 638–643. DOI: 10.1038/nature03127. URL: https://doi.org/10.1038/nature03127.

139. D. Kappel, R. Legenstein, S. Habenschuss, M. Hsieh, and W. Maass. "A Dynamic Connectome Supports the Emergence of Stable Computational Function of Neural Circuits through Reward-Based Learning". *eneuro* 5:2, 2018, ENEURO.0301–17.2018. DOI: 10.1523/eneuro.0301-17.2018. URL: https://doi.org/10.1523/eneuro.0301-17.2018.

140. H. J. Karten. "Vertebrate brains and evolutionary connectomics: On the origins of the mammalian 'neocortex'". *Philosophical Transactions of the Royal Society B: Biological Sciences* 370:1684, 2015. ISSN: 14712970. DOI: 10.1098/rstb.2015.0060.

141. R. Kawai, T. Markman, R. Poddar, R. Ko, A. L. Fantana, A. K. Dhawale, A. R. Kampff, and B. P. Ölveczky. "Motor Cortex Is Required for Learning but Not for Executing a Motor Skill". *Neuron* 86:3, 2015, pp. 800–812. DOI: 10.1016/j.neuron.2015.03.024. URL: https://doi.org/10.1016/j.neuron.2015.03.024.

142. F. W. Kerr. "Structural and functional evidence of plasticity in the central nervous system". *Experimental Neurology* 48:3 PART 2, 1975, pp. 16–31. ISSN: 10902430. DOI: 10.1016/0014-4886(75)90169-7.

143. Y. Kim, S. Kwon, R. Rajan, C. Mori, and S. Kojima. "Intrinsic motivation for singing in songbirds is enhanced by temporary singing suppression and regulated by dopamine". *Scientific Reports* 11:1, 2021. DOI: 10.1038/s41598-021-99456-w. URL: https://doi.org/10.1038/s41598-021-99456-w.

144. J. Kirn, A. Alvarez-Buylla, and F. Nottebohm. "Production and survival of projection neurons in a forebrain vocal center of adult male canaries". *The Journal of Neuroscience* 11:6, 1991, pp. 1756–1762. DOI: 10.1523/jneurosci.11-06-01756.1991. URL: https://doi.org/10.1523/jneurosci.11-06-01756.1991.

145. A. Knoblauch. "Impact of structural plasticity on memory formation and decline". In: *The Rewiring Brain*. Elsevier, 2017, pp. 361–386.

146. A. Knoblauch, E. Körner, U. Körner, and F. T. Sommer. "Structural synaptic plasticity has high memory capacity and can explain graded amnesia, catastrophic forgetting, and the spacing effect". *PloS one* 9:5, 2014, e96485.

147. A. Knoblauch, G. Palm, and F. T. Sommer. "Memory capacities for synaptic and structural plasticity". *Neural Computation* 22:2, 2010, pp. 289–341.

148. A. Knoblauch and F. T. Sommer. "Structural plasticity, effectual connectivity, and memory in cortex". *Frontiers in neuroanatomy* 10, 2016, p. 63.

149. B. Knowlton and T. Moody. "Procedural Learning in Humans". In: *Learning and Memory: A Comprehensive Reference*. Elsevier, 2008, pp. 321–340. DOI: 10.1016/b978-012370509-9.00123-6. URL: https://doi.org/10.1016/b978-012370509-9.00123-6.

150. T. Kohonen. "Self-organized formation of topologically correct feature maps". *Biological Cybernetics* 43:1, 1982, pp. 59–69. DOI: 10.1007/bf00337288. URL: https://doi.org/10.1007/bf00337288.

151. S. Kojima and A. J. Doupe. "Social performance reveals unexpected vocal competency in young songbirds". *Proceedings of the National Academy of Sciences* 108:4, 2011, pp. 1687–1692. DOI: 10.1073/pnas.1010502108. URL: https://doi.org/10.1073/pnas.1010502108.

152. B. Kolb and R. Gibb. *Brain Plasticity and Behaviour in the Developing Brain*. 2011. URL: https://pubmed.ncbi.nlm.nih.gov/22114608/.

153. S. Kollmorgen, R. H. R. Hahnloser, and V. Mante. "Nearest neighbours reveal fast and slow components of motor learning". *Nature* 577:7791, 2020, pp. 526–530. DOI: 10.1038/s41586-019-1892-x.

154. M. Konishi. "Birdsong: From Behavior to Neuron". *Annual Review of Neuroscience* 8:1, 1985, pp. 125–170. ISSN: 0147006X. DOI: 10.1146/annurev.neuro.8.1.125.

155. K. P. Körding and D. M. Wolpert. "Bayesian decision theory in sensorimotor control". *Trends in Cognitive Sciences* 10:7, 2006, pp. 319–326. DOI: 10.1016/j.tics.2006.05.003. URL: https://doi.org/10.1016/j.tics.2006.05.003.

156. S. Korsia and S. W. Bottjer. "Chronic testosterone treatment impairs vocal learning in male zebra finches during a restricted period of development". *Journal of Neuroscience* 11:8, 1991, pp. 2362–2371. ISSN: 02706474. DOI: 10.1523/jneurosci.11-08-02362.1991.

157. J. W. Krakauer, A. M. Hadjiosif, J. Xu, A. L. Wong, and A. M. Haith. *Motor Learning*. 2019. DOI: 10.1002/cphy.c170043. URL: https://doi.org/10.1002/cphy.c170043.

158. M. Kubota and I. Taniguchi. "Electrophysiological Characteristics of Classes of Neuron in the HVc of the Zebra Finch". *Journal of Neurophysiology* 80:2, 1998, pp. 914–923. DOI: 10.1152/jn.1998.80.2.914. URL: https://doi.org/10.1152/jn.1998.80.2.914.

159. P. K. Kuhl. "Brain Mechanisms in Early Language Acquisition". *Neuron* 67:5, 2010, pp. 713–727. DOI: 10.1016/j.neuron.2010.08.038. URL: https://doi.org/10.1016/j.neuron.2010.08.038.

160. M. Kuperstein. "Neural Model of Adaptive Hand-Eye Coordination for Single Postures". *Science* 239:4845, 1988, pp. 1308–1311. DOI: 10.1126/science.3344437. URL: https://doi.org/10.1126/science.3344437.

161. L. D. LaDage, T. C. Roth, and V. V. Pravosudov. "Hippocampal neurogenesis is associated with migratory behaviour in adult but not juvenile sparrows (Zonotrichia leucophrys ssp.)" *Proceedings of the Royal Society B: Biological Sciences* 278:1702, 2011, pp. 138–143. ISSN: 14712970. DOI: 10.1098/rspb.2010.0861.

162. R. Laje, T. J. Gardner, and G. B. Mindlin. "Neuromuscular control of vocalizations in birdsong: A model". *Physical Review E* 65:5, 2002. DOI: 10.1103/physreve.65.051921. URL: https://doi.org/10.1103/physreve.65.051921.

163. B. M. Lake, R. Salakhutdinov, J. Gross, and J. B. Tenenbaum. "One shot learning of simple visual concepts". In: *Proceedings of the 33th Annual Meeting of the Cognitive Science Society, CogSci 2011, Boston, Massachusetts, USA, July 20-23, 2011*. Ed. by L. A. Carlson, C. Hölscher, and T. F. Shipley. cognitivesciencesociety.org, 2011. URL: https://mindmodeling.org/cogsci2011/papers/0601/index.html.

164. J. V. Le Bé and H. Markram. "Spontaneous and evoked synaptic rewiring in the neonatal neocortex". *Proceedings of the National Academy of Sciences of the United States of America* 103:35, 2006, pp. 13214–13219. ISSN: 00278424. DOI: 10.1073/pnas.0604691103.

165. A. Leblois, Á. L. Bodor, A. L. Person, and D. J. Perkel. "Millisecond timescale disinhibition mediates fast information transmission through an avian basal ganglia loop". *Journal of Neuroscience* 29:49, 2009, pp. 15420–15433. ISSN: 02706474. DOI: 10.1523/JNEUROSCI.3060-09.2009.

166. A. Leblois, B. J. Wendel, and D. J. Perkel. "Striatal dopamine modulates basal ganglia output and regulates social context-dependent behavioral variability through D1 receptors". *Journal of Neuroscience* 30:16, 2010, pp. 5730–5743. ISSN: 02706474. DOI: 10.1523/JNEUROSCI.5974-09.2010.

167. Y. LeCun, J. S. Denker, and S. A. Solla. "Optimal Brain Damage". In: *Advances in Neural Information Processing Systems 2*. Ed. by D. S. Touretzky. Morgan-Kaufmann, 1990, pp. 598–605. URL: http://papers.nips.cc/paper/250-optimal-brain-damage.pdf.

168. B. Leuner and E. Gould. "Structural plasticity and hippocampal function". *Annual review of psychology* 61, 2010, pp. 111–140.

169. D. A. Lewis and S. R. Sesack. "Dopamine systems in the primate brain". *Handbook of Chemical Neuroanatomy* 13, 1997, pp. 263–375. ISSN: 09248196. DOI: 10.1016/S0924-8196(97)80008-5.

170. T. L. Lewis Jr, J. Courchet, and F. Polleux. "Cellular and molecular mechanisms underlying axon formation, growth, and branching". *Journal of Cell Biology* 202:6, 2013, pp. 837–848.

171. T. Lightheart, S. Grainger, and T.-F. Lu. "Spike-timing-dependent construction". *Neural computation* 25:10, 2013, pp. 2611–2645.

172. R. B. Livingston. "Brain Mechanisms in Conditioning and Learning". 4:3, 1966, pp. 235–347. URL: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19670018681.pdf.

173. P. M. Lledo, M. Alonso, and M. S. Grubb. "Adult neurogenesis and functional plasticity in neuronal circuits". *Nature Reviews Neuroscience* 7:3, 2006, pp. 179–193. ISSN: 1471003X. DOI: 10.1038/nrn1867.

174. M. A. Long and M. S. Fee. "Using temperature to analyse temporal dynamics in the songbird motor pathway". *Nature* 456:7219, 2008, pp. 189–194. ISSN: 14764687. DOI: 10.1038/nature07448. URL: https://doi.org/10.1038/nature07448.

175. M. Luo, L. Ding, and D. J. Perkel. "An avian basal ganglia pathway essential for vocal learning forms a closed topographic loop". *Journal of Neuroscience* 21:17, 2001, pp. 6836–6845. ISSN: 02706474. DOI: 10.1523/jneurosci.21-17-06836.2001.

176. M. Luo and D. J. Perkel. "A GABAergic, strongly inhibitory projection to a thalamic nucleus in the zebra finch song system". *Journal of Neuroscience* 19:15, 1999, pp. 6700–6711. ISSN: 02706474. DOI: 10.1523/jneurosci.19-15-06700.1999.

177. K. D. MacDermot et al. "Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits". *American Journal of Human Genetics* 76:6, 2005, pp. 1074–1080. ISSN: 00029297. DOI: 10.1086/430841.

178. R. C. Malenka and M. F. Bear. "LTP and LTD". *Neuron* 44:1, 2004, pp. 5–21. DOI: 10.1016/j.neuron.2004.09.012. URL: https://doi.org/10.1016/j.neuron.2004.09.012.

179. M. M. Mariette and K. L. Buchanan. "Prenatal acoustic communication programs offspring for higher posthatching temperatures in a songbird". *Behavioral Ecology* 353:6301, 2016, pp. 812–814. DOI: 10.1002/jez.2022.Pre-.

180. H. Markram, J. Lübke, M. Frotscher, and B. Sakmann. "Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs". *Science* 275:5297, 1997, pp. 213–215.

181. P. Marler. "A comparative approach to vocal learning: Song development in white-crowned sparrows." *Journal of Comparative and Physiological Psychology* 71:2, Pt.2, 1970, pp. 1–25. DOI: 10.1037/h0029144. URL: https://doi.org/10.1037/h0029144.

182. P. Marler and S. Peters. "The Role of Song Phonology and Syntax in Vocal Learning Preferences in the Song Sparrow, Melospiza melodia". *Ethology* 77:2, 2010, pp. 125–149. DOI: 10.1111/j.1439-0310.1988.tb00198.x. URL: https://doi.org/10.1111/j.1439-0310.1988.tb00198.x.

183. D. Marr. "Simple memory: a theory for archicortex". *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* 262:841, 1971, pp. 23–81. DOI: 10.1098/rstb.1971.0078. URL: https://doi.org/10.1098/rstb.1971.0078.

184. P. D. Martin and M. L. Shapiro. "Disparate effects of long-term potentiation on evoked potentials and single CA1 neurons in the hippocampus of anesthetized rats". *Hippocampus* 10:3, 2000, pp. 207–212. DOI: 10.1002/1098-1063(2000)10:3<207::aid-hipo1>3.0.co;2-s. URL: https://doi.org/10.1002/1098-1063(2000)10:3%3C207::aid-hipo1%3E3.0.co;2-s.

185. T. Maviel. "Sites of Neocortical Reorganization Critical for Remote Spatial Memory". *Science* 305:5680, 2004, pp. 96–99. DOI: 10.1126/science.1098180. URL: https://doi.org/10.1126/science.1098180.

186. L. Medina and A. Reiner. "Do birds possess homologues of mammalian primary visual, somatosensory and motor cortices?" *Evolution of Nervous Systems* 23, 2000, pp. 1–12. DOI: 10.1016/B0-12-370878-8/00132-4.

187. W. H. Mehaffey and A. J. Doupe. "Naturalistic stimulation drives opposing heterosynaptic plasticity at two inputs to songbird cortex". *Nature Neuroscience* 18:9, 2015, pp. 1272–1280. DOI: 10.1038/nn.4078. URL: https://doi.org/10.1038/nn.4078.

188. J. Meitzen, D. J. Perkel, and E. A. Brenowitz. "Seasonal changes in intrinsic electrophysiological activity of song control neurons in wild song sparrows". *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology* 193:6, 2007, pp. 677–683. ISSN: 03407594. DOI: 10.1007/s00359-007-0222-1.

189. M. Migaud, M. Batailler, D. Pillon, I. Franceschini, and B. Malpaux. "Seasonal changes in cell proliferation in the adult sheep brain and pars tuberalis". *Journal of Biological Rhythms* 26:6, 2011, pp. 486–496. ISSN: 07487304. DOI: 10.1177/0748730411420062.

190. B. Milner, L. R. Squire, and E. R. Kandel. "Cognitive Neuroscience and the Study of Memory". *Neuron* 20:3, 1998, pp. 445–468. DOI: 10.1016/s0896-6273(00)80987-3. URL: https://doi.org/10.1016/s0896-6273(00)80987-3.

191. J. F. Montiel, N. A. Vasistha, F. Garcia-Moreno, and Z. Molnár. "From sauropsids to mammals and back: New approaches to comparative cortical development". *Journal of Comparative Neurology* 524:3, 2016, pp. 630–645. ISSN: 10969861. DOI: 10.1002/cne.23871.

192. R. Mooney and M. Rao. "Waiting periods versus early innervation: the development of axonal connections in the zebra finch song system". *The Journal of Neuroscience* 14:11, 1994, pp. 6532–6543. DOI: 10.1523/jneurosci.14-11-06532.1994.

193. R. Mooney. "The HVC Microcircuit: The Synaptic Basis for Interactions between Song Motor and Vocal Plasticity Pathways". *Journal of Neuroscience* 25:8, 2005, pp. 1952–1964. DOI: 10.1523/jneurosci.3726-04.2005. URL: https://doi.org/10.1523/jneurosci.3726-04.2005.

194. R. Mooney. "Neural mechanisms for learned birdsong". *Learning and Memory* 16:11, 2009, pp. 655–669. ISSN: 10720502. DOI: 10.1101/lm.1065209.

195. R. Mooney and M. Konishi. "Two distinct inputs to an avian song nucleus activate different glutamate receptor subtypes on individual neurons". *Proceedings of the National Academy of Sciences of the United States of America* 88:10, 1991, pp. 4075–4079. ISSN: 00278424. DOI: 10.1073/pnas.88.10.4075. URL: https://www.pnas.org/content/88/10/4075.short.

196. R. Mooney and M. Rao. "Waiting periods versus early innervation: The development of axonal connections in the zebra finch song system". *Journal of Neuroscience* 14:11 I, 1994, pp. 6532–6543. ISSN: 02706474. DOI: 10.1523/jneurosci.14-11-06532.1994.

197. J. M. Murray and G. S. Escola. "Remembrance of things practiced with fast and slow learning in cortical and subcortical pathways". *Nature Communications* 11:1, 2020. DOI: 10.1038/s41467-020-19788-5. URL: https://doi.org/10.1038/s41467-020-19788-5.

198. T. Nakamori, F. Maekawa, K. Sato, K. Tanaka, and H. Ohki-Hamazaki. "Neural basis of imprinting behavior in chicks". *Development, Growth & Differentiation* 55:1, 2013, pp. 198–206. DOI: 10.1111/dgd.12028. URL: https://doi.org/10.1111/dgd.12028.

199. D. A. Nelson and P. Marler. "Selection-based learning in bird song development." *Proceedings of the National Academy of Sciences* 91:22, 1994, pp. 10498–10501. DOI: 10.1073/pnas.91.22.10498. URL: https://doi.org/10.1073/pnas.91.22.10498.

200. R. I. Nicolson and A. J. Fawcett. "Procedural learning difficulties: reuniting the developmental disorders?" *Trends in Neurosciences* 30:4, 2007, pp. 135–141. DOI: 10.1016/j.tins.2007.02.003. URL: https://doi.org/10.1016/j.tins.2007.02.003.

201. K. W. Nordeen and E. J. Nordeen. "Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches". *Behavioral and Neural Biology* 57:1, 1992, pp. 58–66. DOI: 10.1016/0163-1047(92)90757-u. URL: https://doi.org/10.1016/0163-1047(92)90757-u.

202. K. W. Nordeen and E. J. Nordeen. "Projection neurons within a vocal motor pathway are born during song learning in zebra finches". *Nature* 334:6178, 1988, pp. 149–151. ISSN: 00280836. DOI: 10.1038/334149a0.

203. F. Nottebohm. "A Brain for All Seasons: Cyclical Anatomical Changes in Song Control Nuclei of the Canary Brain". *Science* 214:4527, 1981, pp. 1368–1370. URL: https://science.sciencemag.org/content/214/4527/1368.

204. F. Nottebohm, M. E. Nottebohm, and L. Crane. "Developmental and seasonal changes in canary song and their relation to changes in the anatomy of song-control nuclei". *Behavioral and Neural Biology* 46:3, 1986, pp. 445–471. ISSN: 01631047. DOI: 10.1016/S0163-1047(86)90485-1.

205. F. Nottebohm, M. E. Nottebohm, L. A. Crane, and J. C. Wingfield. "Seasonal changes in gonadal hormone levels of adult male canaries and their relation to song". *Behavioral and Neural Biology* 47:2, 1987, pp. 197–211. ISSN: 01631047. DOI: 10.1016/S0163-1047(87)90327-X.

206. F. Nottebohm, T. M. Stokes, and C. M. Leonard. "Central control of song in the canary, Serinus canarius". *The Journal of Comparative Neurology* 165:4, 1976, pp. 457–486. DOI: 10.1002/cne.901650405. URL: https://doi.org/10.1002/cne.901650405.

207. W. C. Oh, T. C. Hill, and K. Zito. "Synapse-specific and size-dependent mechanisms of spine structural plasticity accompanying synaptic weakening". *Proceedings of the National Academy of Sciences of the United States of America* 110:4, 2013. ISSN: 00278424. DOI: 10.1073/pnas.1214705110.

208. E. Oja. "Simplified neuron model as a principal component analyzer". *Journal of Mathematical Biology* 15:3, 1982, pp. 267–273. DOI: 10.1007/bf00275687. URL: https://doi.org/10.1007/bf00275687.

209. B. P. Ölveczky, A. S. Andalman, and M. S. Fee. "Vocal experimentation in the juvenile songbird requires a basal ganglia circuit". *PLoS Biology* 3:5, 2005, pp. 0902–0909. ISSN: 15457885. DOI: 10.1371/journal.pbio.0030153.

210. B. P. Ölveczky, T. M. Otchy, J. H. Goldberg, D. Aronov, and M. S. Fee. "Changes in the neural control of a complex motor sequence during learning". *Journal of Neurophysiology* 106:1, 2011, pp. 386–397. ISSN: 00223077. DOI: 10.1152/jn.00018.2011.

211. A. van Ooyen and E. Butz-Ostendorf Markus. *The rewiring brain: a computational approach to structural plasticity in the adult brain*. Academic Press, 2017. URL: https://www.sciencedirect.com/book/9780128037843/the-rewiring-brain.

212. S. Pagliarini, X. Hinaut, and A. Leblois. "A bio-inspired model towards vocal gesture learning in songbird". In: *2018 Joint IEEE 8th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*. IEEE, 2018. DOI: 10.1109/devlrn.2018.8761009. URL: https://doi.org/10.1109/devlrn.2018.8761009.

213. M. F. Paredes, S. F. Sorrells, J. M. Garcia-Verdugo, and A. Alvarez-Buylla. "Brain size and limits to adult neurogenesis". *Journal of Comparative Neurology* 524:3, 2016, pp. 646–664. ISSN: 10969861. DOI: 10.1002/cne.23896.

214. A. Pascual-Leone and F. Torres. "Plasticity of the sensorimotor cortex representation of the reading finger in braille readers". *Brain* 116:1, 1993, pp. 39–52. ISSN: 14602156. DOI: 10.1093/brain/116.1.39.

215. A. Pasupathy and E. K. Miller. "Different time courses of learning-related activity in the prefrontal cortex and striatum". *Nature* 433:7028, 2005, pp. 873–876.

216. A. Pasupathy and E. K. Miller. "Different time courses of learning-related activity in the prefrontal cortex and striatum". *Nature* 433:7028, 2005, pp. 873–876. DOI: 10.1038/nature03287. URL: https://doi.org/10.1038/nature03287.

217. M. H. Patton, J. A. Blundon, and S. S. Zakharenko. "Rejuvenation of plasticity in the brain: opening the critical period". *Current Opinion in Neurobiology* 54, 2019, pp. 83–89. ISSN: 18736882. DOI: 10.1016/j.conb.2018.09.003. URL: https://doi.org/10.1016/j.conb.2018.09.003.

218. I. M. Pepperberg. "Cognitive and communicative abilities of Grey parrots". *Applied Animal Behaviour Science* 100:1-2, 2006, pp. 77–86. DOI: 10.1016/j.applanim.2006.04.005. URL: https://doi.org/10.1016/j.applanim.2006.04.005.

219. P. Poirazi and B. W. Mel. "Impact of active dendrites and structural plasticity on the memory capacity of neural tissue". *Neuron* 29:3, 2001, pp. 779–796.

220. C. Portera-Cailliau, R. M. Weimer, V. De Paola, P. Caroni, and K. Svoboda. "Diverse modes of axon elaboration in the developing neocortex". *PLoS Biology* 3:8, 2005. ISSN: 15449173. DOI: 10.1371/journal.pbio.0030272.

221. J. F. Prather, S. Peters, S. Nowicki, and R. Mooney. "Persistent Representation of Juvenile Experience in the Adult Songbird Brain". *Journal of Neuroscience* 30:31, 2010, pp. 10586–10598. DOI: 10.1523/jneurosci.6042-09.2010. URL: https://doi.org/10.1523/jneurosci.6042-09.2010.

222. H. Prior, A. Schwarz, and O. Gunturkun. "Mirror-Induced Behavior in the Magpie (Pica pica): Evidence of Self-Recognition". *Plos One* 6:8, 2008, pp. 1642–1650. ISSN: 1553-7404. DOI: 10.1371/journal.Citation. URL: papers2://publication/uuid/108B0ADD-6B1F-4262-8CA2-F23F6E665F89.

223. R. Pyle and R. Rosenbaum. "A Reservoir Computing Model of Reward-Modulated Motor Learning and Automaticity". *Neural Computation* 31:7, 2019, pp. 1430–1461. DOI: 10.1162/neco_a_01198. URL: https://doi.org/10.1162/neco_a_01198.

224. C. L. Pytte. "Adult neurogenesis in the songbird: Region-specific contributions of new neurons to behavioral plasticity and stability". *Brain, Behavior and Evolution* 87:3, 2016, pp. 191–204. ISSN: 14219743. DOI: 10.1159/000447048.

225. G. Raisman. "Neuronal plasticity in the septal nuclei of the adult rat". *Brain Research* 14:1, 1969, pp. 25–48. ISSN: 00068993. DOI: 10.1016/0006-8993(69)90029-8.

226. R. Reed. "Pruning algorithms-a survey". *IEEE Transactions on Neural Networks* 4:5, 1993, pp. 740–747. DOI: 10.1109/72.248452. URL: https://doi.org/10.1109/72.248452.

227. A. Reiner, L. Medina, and C. L. Veenman. "Structural and functional evolution of the basal ganglia in vertebrates". *Brain Research Reviews* 28:3, 1998, pp. 235–285. ISSN: 01650173. DOI: 10.1016/S0165-0173(98)00016-2.

228. A. Reiner, D. J. Perkel, et al. "Revised Nomenclature for Avian Telencephalon and Some Related Brainstem Nuclei". *Journal of Comparative Neurology* 473:3, 2004, pp. 377–414. ISSN: 15378276. DOI: 10.1038/jid.2014.371. arXiv: NIHMS150003. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2688697/pdf/nihms-110443.pdf.

229. T. F. Roberts, K. A. Tschida, M. E. Klein, and R. Mooney. "Rapid spine stabilization and synaptic enhancement at the onset of behavioural learning". *Nature* 463:7283, 2010, pp. 948–952. DOI: 10.1038/nature08759.

230. T. F. Roberts, K. A. Tschida, M. E. Klein, and R. Mooney. "Rapid spine stabilization and synaptic enhancement at the onset of behavioural learning". *Physiology & behavior* 176:1, 2016, pp. 139–148. DOI: 10.1016/j.physbeh.2017.03.040.

231. M. T. Ross, D. Flores, R. Bertram, F. Johnson, and R. L. Hyson. "Neuronal intrinsic physiology changes during development of a learned behavior". *eneuro* 4:5, 2017.

232. S. Ruder. *An overview of gradient descent optimization algorithms*. 2016. DOI: 10.48550/ARXIV.1609.04747. URL: https://arxiv.org/abs/1609.04747.

233. L. T. Rutledge. "The effects of denervation and stimulation upon synaptic ultrastructure". *Journal of Comparative Neurology* 178:1, 1978, pp. 117–128. ISSN: 10969861. DOI: 10.1002/cne.901780107.

234. D. Saar, Y. Grossman, and E. Barkai. "Reduced after-hyperpolarization in rat piriform cortex pyramidal neurons is associated with increased learning capability during operant conditioning". *European Journal of Neuroscience* 10:4, 1998, pp. 1518–1523.

235. P. Sah and J. M. Bekkers. "Apical dendritic location of slow afterhyperpolarization current in hippocampal pyramidal neurons: implications for the integration of long-term potentiation". *Journal of Neuroscience* 16:15, 1996, pp. 4537–4542.

236. K. A. Sailor, M. T. Valley, M. T. Wiechert, H. Riecke, G. J. Sun, W. Adams, J. C. Dennis, S. Sharafi, G.-l. Ming, H. Song, et al. "Persistent structural plasticity optimizes sensory information processing in the olfactory bulb". *Neuron* 91:2, 2016, pp. 384–396.

237. J. T. Sakata, C. M. Hampton, and M. S. Brainard. "Social Modulation of Sequence and Syllable Variability in Adult Birdsong". *Journal of Neurophysiology* 99:4, 2008, pp. 1700–1711. DOI: 10.1152/jn.01296.2007. URL: https://doi.org/10.1152/jn.01296.2007.

238. R. Sankar, A. Leblois, and N. P. Rougier. "Dual pathway architecture underlying vocal learning in songbirds", 2022. DOI: 10.1101/2022.04.02.486814. URL: https://doi.org/10.1101/2022.04.02.486814.

239. R. Sankar, N. P. Rougier, and A. Leblois. "Computational benefits of structural plasticity, illustrated in songbirds". *Neuroscience and Biobehavioral Reviews* 132, 2022, pp. 1183–1196. DOI: 10.1016/j.neubiorev.2021.10.033. URL: https://doi.org/10.1016/j.neubiorev.2021.10.033.

240. R. Sankar, N. P. Rougier, and A. Leblois. *rsankar9/Review-Arm-Exploration-Model: Arm Exploration Miniature Model*. Version v1.0.0. 2020. DOI: 10.5281/zenodo.4063714. URL: https://doi.org/10.5281/zenodo.4063714.

241. R. Sankar, N. P. Rougier, and A. Leblois. *rsankar9/Review-momentum-illustration*. Version v1.0.0. 2021. DOI: 10.5281/zenodo.4063714. URL: https://doi.org/10.5281/zenodo.4063714.

242. N. Schaefer et al. "The malleable brain: plasticity of neural circuits and behavior – a review from students to students". *Journal of Neurochemistry* 142:6, 2017, pp. 790–811. ISSN: 14714159. DOI: 10.1111/jnc.14107.

243. C. Scharff and F. Nottebohm. "A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: Implications for vocal learning". *Journal of Neuroscience* 11:9, 1991, pp. 2896–2913. ISSN: 02706474. DOI: 10.1523/jneurosci.11-09-02896.1991.

244. M. F. Schmidt, J. Mclean, and F. Goller. "Breathing and vocal control: The respiratory system as both a driver and a target of telencephalic vocal motor circuits in songbirds". *Experimental Physiology* 97:4, 2012, pp. 455–461. ISSN: 1469445X. DOI: 10.1113/expphysiol.2011.058669.

245. W. Schultz, P. Dayan, and P. R. Montague. "A neural substrate of prediction and reward". *Science* 275:5306, 1997, pp. 1593–1599. ISSN: 00368075. DOI: 10.1126/science.275.5306.1593.

246. M. Schwartze and S. A. Kotz. "Contributions of cerebellar event-based temporal processing and preparatory function to speech perception". *Brain and Language* 161, 2016, pp. 28–32. DOI: 10.1016/j.bandl.2015.08.005. URL: https://doi.org/10.1016/j.bandl.2015.08.005.

247. S. S. Shank and D. Margoliash. "Sleep and sensorimotor integration during early vocal learning in a songbird". *Nature* 458:7234, 2009, pp. 73–77.

248. D. F. Sherry and J. S. Hoshooley. "Seasonal hippocampal plasticity in food-storing birds". *Philosophical Transactions of the Royal Society B: Biological Sciences* 365:1542, 2010, pp. 933–943. ISSN: 14712970. DOI: 10.1098/rstb.2009.0220.

249. D. F. Sherry and S. A. MacDougall-Shackleton. "Seasonal change in the avian hippocampus". *Frontiers in Neuroendocrinology* 37, 2015, pp. 158–167. DOI: 10.1016/j.yfrne.2014.11.008. URL: https://doi.org/10.1016/j.yfrne.2014.11.008.

250. H. Z. Shouval, M. F. Bear, and L. N. Cooper. "A unified model of NMDA receptor-dependent bidirectional synaptic plasticity". *Proceedings of the National Academy of Sciences* 99:16, 2002, pp. 10831–10836.

251. T. Sigurdsson, V. Doyère, C. K. Cain, and J. E. LeDoux. "Long-term potentiation in the amygdala: A cellular mechanism of fear learning and memory". *Neuropharmacology* 52:1, 2007, pp. 215–227. DOI: 10.1016/j.neuropharm.2006.06.022. URL: https://doi.org/10.1016/j.neuropharm.2006.06.022.

252. J. Sirosh and R. Miikkulainen. "Self-organization and functional role of lateral connections and multisize receptive fields in the primary visual cortex". *Neural Processing Letters* 3:1, 1996, pp. 39–48. DOI: 10.1007/bf00417788. URL: https://doi.org/10.1007/bf00417788.

253. M. Sizemore and D. J. Perkel. "Premotor synaptic plasticity limited to the critical period for song learning". *Proceedings of the National Academy of Sciences* 108:42, 2011, pp. 17492–17497. DOI: 10.1073/pnas.1104255108.

254. M. Skocik and A. Kozhevnikov. "Real-time system for studies of the effects of acoustic feedback on animal vocalizations". *Frontiers in Neural Circuits* 6, 2013. DOI: 10.3389/fncir.2012.00111. URL: https://doi.org/10.3389/fncir.2012.00111.

255. S. J. Sober, M. J. Wohlgemuth, and M. S. Brainard. "Central Contributions to Acoustic Variation in Birdsong". *Journal of Neuroscience* 28:41, 2008, pp. 10370–10379. DOI: 10.1523/jneurosci.2448-08.2008.

256. S. J. Sober, S. Sponberg, I. Nemenman, and L. H. Ting. "Millisecond Spike Timing Codes for Motor Control". *Trends in Neurosciences* 41:10, 2018, pp. 644–648. DOI: 10.1016/j.tins.2018.08.010. URL: https://doi.org/10.1016/j.tins.2018.08.010.

257. F. Sohrabji, E. J. Nordeen, and K. W. Nordeen. "Characterization of neurons born and incorporated into a vocal control nucleus during avian song learning". *Brain Research* 620:2, 1993, pp. 335–338. ISSN: 00068993. DOI: 10.1016/0006-8993(93)90176-N.

258. R. Spiess, R. George, M. Cook, and P. U. Diehl. "Structural plasticity denoises responses and improves learning speed". *Frontiers in computational neuroscience* 10, 2016, p. 93.

259. K. H. Srivastava, C. P. H. Elemans, and S. J. Sober. "Multifunctional and Context-Dependent Control of Vocal Acoustics by Individual Muscles". *Journal of Neuroscience* 35:42, 2015, pp. 14183–14194. DOI: 10.1523/jneurosci.3610-14.2015.

260. A. Stepanyants, P. R. Hof, and D. B. Chklovskii. "Geometry and structural plasticity of synaptic connectivity". *Neuron* 34:2, 2002, pp. 275–288.

261. S. Steven and C. Blakemore. "Cortical Plasticity in the Adult Human Brain." *The cognitive neurosciences (3rd ed.).*, 2004, pp. 1243–1254. URL: https://psycnet.apa.org/record/2005-01373-083.

262. G. F. Striedter. "Bird brains and tool use: Beyond instrumental conditioning". *Brain, Behavior and Evolution* 82:1, 2013, pp. 55–67. ISSN: 00068977. DOI: 10.1159/000352003.

263. R. S. Sutton, A. G. Barto, et al. *Introduction to reinforcement learning*. Vol. 135. MIT press Cambridge, 1998.

264. R. S. Sutton, A. G. Barto, and R. J. Williams. "Reinforcement learning is direct adaptive optimal control". *IEEE control systems magazine* 12:2, 1992, pp. 19–22.

265. O. Tchernichovski, P. P. Mitra, T. Lints, and F. Nottebohm. "Dynamics of the Vocal Imitation Process: How a Zebra Finch Learns Its Song". *Science* 291:5513, 2001, pp. 2564–2569. DOI: 10.1126/science.1058522. URL: https://doi.org/10.1126/science.1058522.

266. K. Teffer and K. Semendeferi. "Human prefrontal cortex: evolution, development, and pathology". In: *Progress in brain research*. Vol. 195. Elsevier, 2012, pp. 191–218. URL: https://doi.org/10.1016/B978-0-444-53860-4.00009-X.

267. T. Teşileanu, B. Ölveczky, and V. Balasubramanian. "Rules and mechanisms for efficient two-stage learning in neural circuits". *Elife* 6, 2017, e20944.

268. M. Tessier-Lavigne and C. S. Goodman. "The Molecular Biology of Axon Guidance". *Science* 274:5290, 1996, pp. 1123–1133. ISSN: 0036-8075. DOI: 10.1126/science.274.5290.1123. eprint: https://science.sciencemag.org/content/274/5290/1123.full.pdf. URL: https://science.sciencemag.org/content/274/5290/1123.

269. G. Teuchert-Noodt, K.-H. Breuker, and R. Dawirs. "Neuronal lysosome accumulation in degrading synapses of sensory-motor and limbic subsystems in the duck Anas platyrhynchos: indication of rearrangements during avian brain development?" *Developmental neuroscience* 13:3, 1991, pp. 151–163. URL: https://doi.org/10.1159/000112153.

270. J. A. Thompson and F. Johnson. "HVC Microlesions Do Not Destabilize the Vocal Patterns of Adult Male Zebra Finches with Prior Ablation of LMAN". *Developmental Neurobiology* 67: 2007, pp. 205–218. ISSN: 1932-846X. DOI: 10.1002/dneu.

271. N.-W. Tien and D. Kerschensteiner. "Homeostatic plasticity in neural development". *Neural Development* 13:1, 2018. DOI: 10.1186/s13064-018-0105-x. URL: https://doi.org/10.1186/s13064-018-0105-x.

272. J. T. Trachtenberg, B. E. Chen, G. W. Knott, G. Feng, J. R. Sanes, E. Welker, and K. Svoboda. "Long-term in vivo imaging of experience-dependent synaptic plasticity in adult cortex". *Nature* 420:6917, 2002, pp. 788–794. ISSN: 00280836. DOI: 10.1038/nature01273.

273. A. D. Tramontin and E. A. Brenowitz. "Seasonal plasticity in the adult brain". *Trends in Neurosciences* 23:6, 2000, pp. 251–258. ISSN: 01662236. DOI: 10.1016/S0166-2236(00)01558-7.

274. T. W. Troyer and A. J. Doupe. "An Associational Model of Birdsong Sensorimotor Learning I. Efference Copy and the Learning of Song Syllables". *Journal of Neurophysiology* 84:3, 2000, pp. 1204–1223. DOI: 10.1152/jn.2000.84.3.1204. URL: https://doi.org/10.1152/jn.2000.84.3.1204.

275. C. Tsallis and D. A. Stariolo. "Generalized simulated annealing". *Physica A: Statistical Mechanics and its Applications* 233:1-2, 1996, pp. 395–406. DOI: 10.1016/s0378-4371(96)00271-3.

276. K. A. Tschida and R. Mooney. "Deafening drives cell-type-specific changes to dendritic spines in a sensorimotor nucleus important to learned vocalizations". *Neuron* 73:5, 2012, pp. 1028–1039.

277. E. C. Tumer and M. S. Brainard. "Performance variability enables adaptive plasticity of 'crystallized' adult birdsong". *Nature* 450:7173, 2007, pp. 1240–1244. ISSN: 14764687. DOI: 10.1038/nature06390.

278. T. Van Kerkoerle, S. A. Marik, S. M. Z. A. Borgloh, and C. D. Gilbert. "Axonal plasticity associated with perceptual learning in adult macaque primary visual cortex". *Proceedings of the National Academy of Sciences of the United States of America* 115:41, 2018, pp. 10464–10469. ISSN: 10916490. DOI: 10.1073/pnas.1812932115.

279. L. Vasung, E. A. Turk, S. L. Ferradal, J. Sutin, J. N. Stout, B. Ahtam, P.-Y. Lin, and P. E. Grant. "Exploring early human brain development with structural and physiological neuroimaging". *NeuroImage* 187, 2019, pp. 226–254. DOI: 10.1016/j.neuroimage.2018.07.041. URL: https://doi.org/10.1016/j.neuroimage.2018.07.041.

280. G. E. Vates and F. Nottebohm. "Feedback circuitry within a song-learning pathway". *Neurobiology* 92:May, 1995, pp. 5139–5143. URL: https://doi.org/10.1073/pnas.92.11.5139.

281. D. S. Vicario. "Organization of the zebra finch song control system: Functional organization of outputs from nucleusrobustus archistriatalis". *The Journal of Comparative Neurology* 309:4, 1991, pp. 486–494. DOI: 10.1002/cne.903090405. URL: https://doi.org/10.1002/cne.903090405.

282. N. Vieillard, B. Scherrer, O. Pietquin, and M. Geist. "Momentum in reinforcement learning". In: *International Conference on Artificial Intelligence and Statistics*. PMLR. 2020, pp. 2529–2538.

283. E. Vu, M. Mazurek, and Y. Kuo. "Identification of a forebrain motor programming network for the learned song of zebra finches". *The Journal of Neuroscience* 14:11, 1994, pp. 6924–6934. DOI: 10.1523/jneurosci.14-11-06924.1994. URL: https://doi.org/10.1523/jneurosci.14-11-06924.1994.

284. K. Wada et al. "A molecular neuroethological approach for identifying and characterizing a cascade of behaviorally regulated genes". *Proceedings of the National Academy of Sciences of the United States of America* 103:41, 2006, pp. 15212–15217. ISSN: 00278424. DOI: 10.1073/pnas.0607098103.

285. Y. Wang, A. Brzozowska-Prechtl, and H. J. Karten. "Laminar and columnar auditory cortex in avian brain". *Proceedings of the National Academy of Sciences of the United States of America* 107:28, 2010, pp. 12676–12681. ISSN: 00278424. DOI: 10.1073/pnas.1006645107.

286. T. L. Warren, E. C. Tumer, J. D. Charlesworth, and M. S. Brainard. "Mechanisms and time course of vocal learning and consolidation in the adult songbird". *Journal of Neurophysiology* 106:4, 2011, pp. 1806–1821. ISSN: 00223077. DOI: 10.1152/jn.00311.2011.

287. C. S. Whaling, D. A. Nelson, and P. Marler. "Testosterone-induced shortening of the storage phase of song development in birds interferes with vocal learning". *Developmental Psychobiology* 28:7, 1995, pp. 367–376. ISSN: 10982302. DOI: 10.1002/dev.420280703.

288. S. A. White, F. S. Livingston, and R. Mooney. "Androgens Modulate NMDA Receptor–Mediated EPSCs in the Zebra Finch Song System". *Journal of Neurophysiology* 82:5, 1999, pp. 2221–2234. DOI: 10.1152/jn.1999.82.5.2221. URL: https://doi.org/10.1152/jn.1999.82.5.2221.

289. J. R. Wickens, J. N. Reynolds, and B. I. Hyland. "Neural mechanisms of reward-related motor learning". *Current Opinion in Neurobiology* 13:6, 2003, pp. 685–690. DOI: 10.1016/j.conb.2003.10.013. URL: https://doi.org/10.1016/j.conb.2003.10.013.

290. R. J. Williams. "Simple statistical gradient-following algorithms for connectionist reinforcement learning". *Machine Learning* 8:3-4, 1992, pp. 229–256. DOI: 10.1007/bf00992696.

291. D. J. Willshaw and C. V. D. Malsburg. "A marker induction mechanism for the establishment of ordered neural mappings: its application to the retinotectal problem". *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* 287:1021, 1979, pp. 203–243. DOI: 10.1098/rstb.1979.0056. URL: https://doi.org/10.1098/rstb.1979.0056.

292. D. M. Wolpert and J. Flanagan. "Motor prediction". *Current Biology* 11:18, 2001, R729–R732. DOI: 10.1016/s0960-9822(01)00432-8. URL: https://doi.org/10.1016/s0960-9822(01)00432-8.

293. D. M. Wolpert, J. Diedrichsen, and J. R. Flanagan. "Principles of sensorimotor learning". *Nature Reviews Neuroscience* 12:12, 2011, pp. 739–751. DOI: 10.1038/nrn3112. URL: https://doi.org/10.1038/nrn3112.

294. R. H. Wurtz, V. F. Castellucci, and J. M. Nusrala. "Synaptic Plasticity: The Effect of the Action Potential in the Postsynaptic Neuron". *Experimental Neurology* 18:3, 1967, pp. 350–368. URL: https://doi.org/10.1016/0014-4886(67)90054-4.

295. L. Xiao, G. Chattree, F. G. Oscos, M. Cao, M. J. Wanat, and T. F. Roberts. "A Basal Ganglia Circuit Sufficient to Guide Birdsong Learning". *Neuron* 98:1, 2018, 208–221.e5. ISSN: 10974199. DOI: 10.1016/j.neuron.2018.02.020. URL: https://doi.org/10.1016/j.neuron.2018.02.020.

296. T. Xu, X. Yu, A. J. Perlik, W. F. Tobin, J. A. Zweig, K. Tennant, T. Jones, and Y. Zuo. "Rapid formation and selective stabilization of synapses for enduring motor memories". *Nature* 462:7275, 2009, pp. 915–919. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2844762/.

297. Y. Yazaki-Sugiyama. "Neuronal mechanisms regulating the critical period of sensory experience-dependent song learning". *Neuroscience Research* 140, 2019, pp. 53–58. ISSN: 18728111. DOI: 10.1016/j.neures.2018.11.002. URL: https://doi.org/10.1016/j.neures.2018.11.002.

298. I. B. Yildiz and S. J. Kiebel. "A Hierarchical Neuronal Model for Generation and Online Recognition of Birdsongs". *PLoS Computational Biology* 7:12, 2011. Ed. by O. Sporns, e1002303. DOI: 10.1371/journal.pcbi.1002303. URL: https://doi.org/10.1371/journal.pcbi.1002303.

299. A. C. Yu and D. Margoliash. "Temporal Hierarchical Control of Singing in Birds". *Science* 273:5283, 1996, pp. 1871–1875. DOI: 10.1126/science.273.5283.1871. URL: https://doi.org/10.1126/science.273.5283.1871.

300. A. T. Zai, S. Cavé-Lopez, M. Rolland, N. Giret, and R. H. R. Hahnloser. "Sensory substitution reveals a manipulation bias". *Nature Communications* 11:1, 2020. DOI: 10.1038/s41467-020-19686-w. URL: https://doi.org/10.1038/s41467-020-19686-w.