CONTEXTUAL MODULATION OF NATURALISTIC BEHAVIOR AND SENSORY PROCESSING

by

DAVID GRAHAM WYRICK

A DISSERTATION

Presented to the Department of Biology
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

September 2022

DISSERTATION APPROVAL PAGE

Student: David Graham Wyrick

Title: Contextual Modulation of Naturalistic Behavior and Sensory Processing

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Biology by:

Cris Niell	Chair
Luca Mazzucato	Advisor
James Murry	Core Member
Kip Keller	Core Member
Matt Smear	Institutional Representative

and

Krista Chronister	Vice Provost of Graduate Studies

Original approval signatures are on file with the University of Oregon Division of Graduate Studies.

Degree awarded September 2022

DISSERTATION ABSTRACT

David Graham Wyrick

Doctor of Philosophy

Department of Biology

September 2022

Title: Contextual Modulation of Naturalistic Behavior and Sensory Processing

Forming representations of the sensory world and navigating within it require an animal to actively interact and sample key aspects of the environment. In the olfactory system, many organisms search for food by actively sensing the odor environment, employing search programs to successfully navigate across turbulent odor plumes to the source. In the visual system, we effortlessly perceive the world of light, shadow, color, and movement as a stable, predictable, and unified percept. We use this perception to extract information about the outside world we live in, to move our eyes across a line of text, and critically, to move through our environment. Here, we present three studies investigating how animal behavior and sensory processing interact to produce the complex dynamics we observe in behavioral and neural data.

In Chapter II, we characterized the dynamics of olfactory search behavior in freely moving mice using latent state space modelling. We identified behavioral motifs that constitute the overall search strategy of the mouse. By segmenting the behavior into these identifiable and reoccurring motifs, we determined that mice actively sample the environment in a sniff-synchronized, two-state strategy to gain information about concentration gradient cues.

In Chapter III, we linked theoretical predictions from a spiking neural network model to observations from electrophysiological recordings in mouse visual cortex and to the behavioral state of the animal. We connect mechanisms in our model to neural activity by modeling the locomotion-induced perturbations in cortex as an increase in variance of the input currents to excitatory neurons, which decreases the gain of single neurons and an acceleration of stimulus-processing speed.

In Chapter IV, we quantified the effects of temporal context and expectation on the sensory processing of natural images in mouse visual cortex. We found that all areas we recorded from predominately encode for the temporal context in which the images were presented in. In other words, it matters how and when a stimulus is presented to an animal. Overall, the conjunctive encoding of representations of natural scenes and temporal context was modulated by the expectation about sequential events.

This dissertation consists of previously published co-authored material.

CURRICULUM VITAE

NAME OF AUTHOR: David Graham Wyrick

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, Oregon, USA
Washington State University, Pullman, WA, USA

DEGREES AWARDED:

Doctor of Philosophy in Biology, 2022, University of Oregon
Bachelor of Science in Physics, 2013, Washington State University

AREAS OF SPECIAL INTEREST:

Neural dynamics, sensory coding, serotonin, psychedelics, causality, complexity

PROFESSIONAL EXPERIENCE:

Graduate Research Assistant, University of Oregon
Electro-Optics Scientist, Boeing Company

PUBLICATIONS:

David G Wyrick* , Hannah Choi, Nicholas Cain , Rylan Larsen, Jerome Lecoq, Marina Garrett, Luca Mazzucato. Differential encoding of temporal context and expectation across the visual hierarchy. BioRxiv. 2022.

Philip RL Parker, Eliott TT Abe, Natalie T Beatie, Emmalyn SP Leonard, Dylan M Martins, Shelby L Sharp, David G Wyrick, Luca Mazzucato, Cristopher M Niell. Distance estimation from monocular cues in an ethological visuomotor task. BioRxiv. 2021.

David G Wyrick, Luca Mazzucato. State-dependent regulation of cortical processing speed via gain modulation. Journal of Neuroscience 41 (18). 2021.

Teresa M Findley*, David G Wyrick*, Jennifer L Cramer, Morgan A Brown, Blake Holcomb, Robin Attey, Dorian Yeh, Eric Monasevitch, Nelly Nouboussi, Isabelle Cullen, Jeremea O Songco, Jared F King, Yashar Ahmadian, Matthew C Smear. Sniff-synchronized, gradient-guided olfactory search by freely moving mice. eLife 10:e58523. 2021.

A-N Chené*, D Wyrick*, J Borissova, M Kuhn, A Hervé, S Ramírez Alegría, C Bonatto, J-C Bouret, R Kurtev. Improving distances to Galactic Wolf-Rayet stars. Wolf-Rayet Stars: Proceedings of an International Workshop held in Potsdam, Germany, 1.–5. June 2015

Sheng-Ting Hung, Shiva K Ramini, David G Wyrick, Koen Clays, Mark G Kuzyk. The role of the polymer host on reversible photodegradation in disperse orange 11 dye. Proc. SPIE 8474, Optical Processes in Organic Materials and Nanostructures, 84741A. 2012.

ACKNOWLEDGEMENTS

I would like to thank my mentor, Luca Mazzucato, for his support and guidance throughout the past half decade. His unwavering enthusiasm and energy provided a perfect counterbalance to my deep-set skepticism of scientific progress (as it pertains to disentangling the complexity of brains). I am forever grateful for the engaging discussions I was able to have with him. He has helped me become a more precise scientist, a more structured writer, and yes, a more optimistic person. Thank you to all of my academic mentors - Matt Smear, Cris Niell, Yashar Ahmadian, James Murray, Santiago Jaramillo, and Kip Keller - for serving on my committee and helping to guide my research.

Next, I would like to acknowledge my experimental colleagues, whose data we computational neuroscientists so desperately want and need. Much of the work contained within this dissertation and outside it would not have been possible without you. Thank you to Reese Findley and Matt Smear, for collaborating with me on the olfactory search project and helping me start my graduate career with a success. Thank you to Phil Parker, Angie Michaiel, and Cris Niell, for sharing their precious data on psychedelics with me. Even though I didn't necessarily find what I was looking for, I learned an immense amount and am inspired to keep investigating. Thank you to Evan Vickers, who can collect more data than any one person can analyze. I would also like to acknowledge all of the scientists and organizations that share data to the field at large. The data portion of Chapter III and IV was collected and shared by scientists at the Allen Institute for Brain Science. Thank you for such a monumental survey of data. Open science is better science. And thank you to all of the collaborators I have been fortunate enough to work with and learn from -

Hannah Choi, Marina Garrett, Amin Nejatbakhsh, Francesco Fumarola, and Stefano Recanatesi.

I would be remiss not to acknowledge the members - past and present - of the neurotheory group. It is not an exaggeration to say that Elliott Abe helped convince me to finish graduate school. Not because he had a compelling argument to stay, but because he consistently reminded me that he didn't want to be the only computational person in our class. One is the loneliest number they say. All jokes aside, I am grateful for his words of encouragement during those early days. Throughout our studies, he has always been there to lend a helping hand, or to have an intellectual discussion with. His work ethic and enthusiasm for neuroscience inspired me to dig deeper. To Nicu Istrate, thank you for being my friend / housemate during covid times. To the old guard - Gabe Barello, Caleb Holt, Takafumi Arakaki - thank you for your mentorship, friendship, and computer wizardry. I should not have been given sudo privileges, but you gave them to me anyway. To the present members of the joint labs - Lia Papadopolous, Audra McNamee, Christian Schmidt, Ben Lemberger, Matthew Trappett, James Murray, Danny Burnham - I will miss you all. I am grateful to have been part of this scientific community. Maybe Elliott was on to something after all. And finally, I would like to express my deepest gratitude to my friends and family, who supported me through this ordeal we call graduate school. I could not have done it without you.

This work is dedicated to my mom, for trusting me to find my own way; to my dad, for supporting me despite our disagreements; to my sister, for showing me what true grit looks like; to my brother, for inspiring me to look inward; to Kevin^2, for the conversations; and to Leigh, for loving me when I needed it most.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

LIST OF TABLES

Table Page

CHAPTER I

INTRODUCTION

Trying to understand perception by understanding neurons is like trying to understand a bird's flight by studying only feathers. It just cannot be done.

- David Marr, Vision: A Computational Approach, 1982

We are products of our context, in each moment.

- J. Jones

Let's talk about numbers. It is estimated that the observable Universe contains 100 billion galaxies, spanning a diameter of 8.8 x 10^{26} meters. In this immensity, our own galaxy hosts 300 billion stars, 1 of which we call home. Orbiting at a favorable distance of 1.5 x 10^{8} m and with a significant portion of heavy elements, life began and ended for 99% of all species to appear on earth. Fast forward through 3.7 billion years of evolution to the last cosmic second. The human brain contains roughly 86 billion neurons, with up to 500 trillion synaptic connections. In only the last few hundred years, humans have used their 3 pound gooey masses together in the scientific pursuit to understand the Universe. Modern computing (a CPU can perform at 300 billion floating point operations per second) and neural recording technologies (a 2-photon microscope can record from 10,000s of neurons across the cortex simultaneously) have uniquely positioned the scientific community to *investigate the investigator*. Or at least a distant relative of the investigator, like the mouse, in this dissertation. So, how do we even begin to tease apart the complexity of the brain and the behavior it manifests?

Fundamental to the science of biology is observation. An evolutionary biologist observes the natural world - the different types of animals and plants that constitute a local environment - to say something about the underlying processes that shaped

the origin and dynamics of the species of interest. A molecular biologist observes the inner workings of cells, mini universes unto themselves, to form hypotheses about the molecular mechanisms driving some biological function. A neuroscientist observes the activity of neurons - while attempting to control the sensory input the animal *observes* - to understand how representations of the sensory world are encoded and manipulated by the brain to generate behavior. Historically, this was conducted while an animal was under anesthesia or while an awake animal was head restricted to limit the sensory inputs to those determined by the experimenter and allow imaging and electrophysiological technologies to be utilized. But this disrupts the loop for which the brain evolved to perform. That is, to interpret the sensory world whilst simultaneously interacting and changing it. Sensory representations guide the behavior of an animal in it's environment. In turn, animal behavior - movement - shapes the sensory input the animal observes.

Connecting the function of neural circuits to the emergent behavior of animals is a paramount goal in the field. Once an after thought in the analysis of neural data, characterizing the phenomenology of naturalistic animal behavior is of vital importance for understanding how populations of neurons process the sensory world. Indeed separating the analysis of neural data and behavior, through techniques like head-fixation and trial averaging, is no longer proving adequate when it comes to explaining neural activity in sensory areas. Recent studies have shown that movement related signals - locomotion (Niell and Stryker, 2010), arousal (Stringer et al., 2019b), respiration (Tort et al., 2018), and uninstructed "task-irrelevant" movements like facial twitches (Musall et al., 2019) - pervade areas across the brain, including sensory areas. Why may this be the case? Put simply, sensing the environment is not a passive process. It requires an animal to actively move through the world to sample sensory

scenes. Furthermore, movement - through eye saccades, whisking, self-motion, etc - directly changes the spatiotemporal statistics of sensory stimuli that an animal is exposed to (Wachowiak, 2011; Michaiel et al., 2020). Thus, it is not surprising to observe such ubiquitous representations of behavior, especially in sensory areas.

In the olfactory system, active sensation through sniffing is required to expose olfactory bulb neurons to odorants in the environment. Sniffing entrains neural activity of these sensory neurons and even areas further up the hierarchy like the primary olfactory cortex and pre-frontal cortex (Shusterman et al., 2011; Tort et al., 2018). Ethologically, active sensation through sniffing enables animals to navigate turbulent odor environments to sources of food or mates. An experimental paradigm which allowed mice to perform this behavior unrestrained, in a freely moving context would provide further insights into the behavioral strategies animals use and how these relate to the underlying neural activity. In chapter II, we tackle the first task in this pair and investigate behaving animals in an ethologically relevant olfactory search task. By developing a novel method for quantifying animal behavior, we were able to monitor the animal's movements moment by moment and connect them to the external contexts (i.e. environmental variables like odor concentration gradients, or task vs no-task) in which the animal was operating in. With this quantification, we were able to characterize the strategies animals use while actively sensing an olfactory scene. We learned that mice actively move their snout in a sniff-synchronized way that only appeared during goal-directed, odor-guided behavior. Crucially, this was only observable in a freely-moving context, where the animal was able to initiate trials of olfactory search.

Like olfaction, vision is not a passive process. It entails active scanning of the eyes across the visual scene (Yarbus, 1967). The traditional feedforward,

hierarchical view of visual processing consists of the brain constructing, in each moment, a representation of the world from bottom-up sensory evidence (Van Essen, 1979). This view, while still a mainstay of our understanding, has been modified to encompass context-dependent modulations of visual responses by top-down and lateral connections between cortices and by neuromodulatory input (Khan and Hofer). What emerges is a view of visual processing that is strongly influenced by internal models of the world (Fiser et al., 2016), expectations (Poort et al., 2015), the state of the animal (Musall et al., 2019), and the statistics of the natural world (Field, 1987). In other words, the context in which an animal processes visual information. The most salient example is that of self-motion, where it has been shown that projections from anterior cingulate and secondary motor cortex into primary visual cortex (V1) send motor-related signals that can change to reflect new sensory feedback when visuomotor coupling is inverted (Keller et al., 2012). These top-down signals into V1 can be understood in terms of a prediction-based and context-dependent processing of the sensory world (Keller and Mrsic-Flogel, 2018), where behavior and sensory processing form a closed loop process. In chapter III and IV, we investigate internally driven (through the state of the animal) and externally driven (through experimental conditions) contextual modulations of visual processing in an effort to understand this interplay between behavior and sensory processing.

As a whole, this dissertation is focused on how context - both externally and internally driven - shapes naturalistic behavior and sensory processing. In the broadest sense, context can be defined as the behavioral state(s) of the animal while recording from neural populations. Is the animal running or at rest? Is the animal whisking or spontaneously moving its face? Context can be the specific characteristics of the environment an animal is in or the particular stimuli an animal is exposed to.

Are there odor gradients present or not? Are the visual stimuli presented expected or unexpected? A deeper understanding of the context in which animals engage in behavior and the context in which we record neuron populations will allow us to bridge the gap between neural circuits and behavior. We conclude this chapter with the abstracts for Chapters II - IV.

1.1 Chapter II: Quantifying active sensation

For many organisms, searching for relevant targets such as food or mates entails active, strategic sampling of the environment. Finding odorous targets may be the most ancient search problem that motile organisms evolved to solve. While chemosensory navigation has been well characterized in micro-organisms and invertebrates, spatial olfaction in vertebrates is poorly understood. We have established an olfactory search assay in which freely-moving mice navigate noisy concentration gradients of airborne odor. Mice solve this task using concentration gradient cues and do not require stereo olfaction for performance. During task performance, respiration and nose movement are synchronized with tens of milliseconds precision. This synchrony is present during trials and largely absent during inter-trial intervals, suggesting that sniff-synchronized nose movement is a strategic behavioral state rather than simply a constant accompaniment to fast breathing. To reveal the spatiotemporal structure of these active sensing movements, we used machine learning methods to parse motion trajectories into elementary movement motifs. Motifs fall into two clusters, which correspond to investigation and approach states. Investigation motifs lock precisely to sniffing, such that the individual motifs preferentially occur at specific phases of the sniff cycle. The allocentric structure of investigation and approach indicate an advantage to sampling

both sides of the sharpest part of the odor gradient, consistent with a serial sniff strategy for gradient sensing. This work clarifies sensorimotor strategies for mouse olfactory search and guides ongoing work into the underlying neural mechanisms.

1.2 Chapter III: Moving through the visual world

To thrive in dynamic environments, animals must be capable of rapidly and flexibly adapting behavioral responses to a changing context and internal state. Examples of behavioral flexibility include faster stimulus responses when attentive and slower responses when distracted. Contextual or state-dependent modulations may occur early in the cortical hierarchy and may be implemented via top-down projections from cortico-cortical or neuromodulatory pathways. However, the computational mechanisms mediating the effects of such projections are not known. Here, we introduce a theoretical framework to classify the effects of cell-type specific top-down perturbations on the information processing speed of cortical circuits. Our theory demonstrates that perturbation effects on stimulus processing can be predicted by intrinsic gain modulation, which controls the timescale of the circuit dynamics. Our theory leads to counter-intuitive effects such as improved performance with increased input variance. We tested the model predictions using large-scale electrophysiological recordings from the visual hierarchy in freely running mice, where we found that a decrease in single-cell intrinsic gain during locomotion led to an acceleration of visual processing. Our results establish a novel theory of cell-type specific perturbations, applicable to top-down modulation as well as optogenetic and pharmacological manipulations. Our theory links context, connectivity, dynamics, and information processing via gain modulation.

1.3 Chapter IV: Temporal context and expectations modulate the visual experience

The classic view that neural populations in the visual cortex preferentially encode responses to visual stimuli has been strongly challenged by recent experimental studies. A large fraction of variance in visual responses in rodents can be attributed to behavioral state and movements, trial-history, or salience. Here, we present a comprehensive experimental and theoretical study showing that the cortical visual hierarchy differentially encodes the temporal context and expectation of naturalistic visual stimuli. We measured layer-specific neural responses to expected and unexpected sequences of natural scenes across three visual areas using 2p imaging in behaving mice: the primary visual cortex (V1), the posterior medial higher order visual area (PM), and retrosplenial cortex (RSP). We found that all three areas predominantly encode for the temporal context in which the images were presented. Information about image identity in neural population activity only emerged when images were presented within a recurring sequence, and decreased along the visual hierarchy. We found that the conjunctive encoding of temporal context and image identity was modulated by the emergence of expectation about sequential events. We found enhanced oddball responses, signaling expectation violation, in V1 and PM. We found evidence for predictive coding in RSP, where the oddball response recapitulated the identity of the missing image. We further found evidence for representational drift in visual responses in all areas on the timescale of minutes. Despite this drift, the population responses in V1 and PM, but not in RSP, maintained stable encoding of visual information. Our results establish temporal context and expectation as substantial encoding dimensions in the visual hierarchy and suggest that differential responses along the visual hierarchy instantiate a predictive coding mechanism.

CHAPTER II

SNIFF-SYNCHRONIZED, GRADIENT-GUIDED OLFACTORY SEARCH BY FREELY MOVING MICE

2.1 Author contributions

Originally published as **Teresa M Findley***, **David G Wyrick***, Jennifer L Cramer, Morgan A Brown, Blake Holcomb, Robin Attey, Dorian Yeh, Eric Monasevitch, Nelly Nouboussi, Isabelle Cullen, Jeremea O Songco, Jared F King, Yashar Ahmadian, Matthew C Smear. Sniff-synchronized, gradient guided olfactory search by freely moving mice. eLife, 2021. 10:e58523. *Authors contributed equally. TF and MS conceived the study. MB, TF, and MS designed, built, and maintained the behavioral rig. TF, JC, BH, RA, DY, EM, NN, IC, JS, and JK trained mice. JC and TF performed naris occlusion experiments. DW and TF established and implemented Deeplabcut tracking, DW and YA designed and performed the AR-HMM analysis. TF, DW, BH, YA, and MS analyzed data and wrote the manuscript.

2.2 Introduction

Sensory observations are often made in concert with movements (Ahissar and Assa, 2016; Gibson, 1966). During active search behavior, animals make sampling movements in order to extract relevant sensory information from the environment (Gibson, 1962; Schroeder et al., 2010). Sampling behavior is flexible, and can be customized for the problem the animal is trying to solve (Kleinfeld et al., 2006; Yarbus, 1967). In the brain, sensory and motor systems interact extensively (Andersen and Mountcastle, 1983; Duhamel et al., 1992; McGinley et al., 2015b; Musall et al., 2019; Niell and Stryker, 2010; Poulet and Hedwig, 2006; Sommer and Wurtz, 2008; Stringer et al., 2019b), which reflects the importance of interpreting self-induced stimulus

dynamics (Sperry, 1950; von Holst and Mittelstaedt, 1950; Webb, 2004). Here, we show how mice sample the environment while navigating a noisy odor gradient.

Navigating by chemical cues may be one of the most ancient problems motile organisms evolved to solve, and it remains crucial in the lives of almost all modern species. Unicellular organisms and some invertebrates navigate chemical gradients by chemotaxis (Bargmann, 2006; Berg, 2000; Lockery, 2011). In essence, their movement programs can be described as having two states: they move straight when the concentration is increasing and reorient their movements when the concentration is decreasing. Whereas chemical gradients are stable and informative at the spatial scale of these organisms, for many larger or flying organisms, odor gradient cues do not provide useful positional information (Crimaldi et al., 2002; Murlis et al., 1992). At this larger spatial scale, turbulent airflow moves odor molecules in dynamic spatiotemporal patterns, disrupting concentration gradients and nullifying classical chemotaxis strategies. Instead, olfactory cues often gate movements that depend on other sensory modalities. Here too, these organisms' behavioral structure can be described as transitions between two states: detection of odor promotes upwind movement while the absence of odor promotes crosswind casting movement (Kennedy and Marsh, 1974; van Breugel and Dickinson, 2014; Vickers and Baker, 1994). In this behavioral program, known as odor-gated anemotaxis, odor cues gate behavioral responses to positional information provided by another modality. In both chemotaxis and odor-gated anemotaxis, search tasks can be described with a two-state search model.

In comparison to invertebrates, our understanding of olfactory search behavior in vertebrates is more rudimentary, even in commonly-studied rodent models. In these animals, access to the olfactory environment is gated by respiration, which is in

turn responsive to incoming olfactory stimulation (Kepecs et al., 2006; Wachowiak, 2011). Novel odors evoke rapid sniffing, during which respiration synchronizes with whisker, nose, and head movements on a cycle-by-cycle basis (Kurnikova et al., 2017; Moore et al., 2013; Ranade et al., 2013). Thus, during active mammalian olfaction, sensory and motor systems interact in a closed loop via the environment, as is true for other sensory modalities such as vision or somatosensation (Ahissar and Assa, 2016; Gibson, 1966). The cyclical sampling movements coordinated by respiration further synchronize with activity in widespread brain regions (Karalis and Sirota, 2018; Kay, 2005; Macrides et al., 1982; Vanderwolf, 1992; Yanovsky et al., 2014; Zelano et al., 2016) similarly to correlates of locomotor, pupillary, and facial movements observed throughout the brain (McGinley et al., 2015a; Musall et al., 2019; Niell and Stryker, 2010; Stringer et al., 2019b). Respiratory central pattern generators may coordinate sampling movements to synchronize sensory dynamics across modalities with internal brain rhythms (Kleinfeld et al., 2014).

Previous work has shown that rodents follow odor trails, where the concentration gradient is steep and stable, with rapid sniffing accompanied by side-to-side head movements(Jones and Urban, 2018; Khan et al., 2012). In these conditions, serial sniffing and stereo olfactory cues guide movements of the nose. Likewise, moles used concentration comparisons across space and time to locate a food source in a sealed experimental chamber in which a lack of airflow allowed for even diffusion of a chemical gradient (Catania, 2013). In this study, when input to the nares was reversed, moles navigated towards odor sources at a distance, but demonstrated significant deficits at identifying odor location when near the source. Behavioral modeling in mice further supports that inter-naris concentration comparison plays a more important role in search near the source (Liu et al., 2020). Thus, both serial sniffing and stereo

cues can guide olfactory search behavior. The sensory computations and movement strategies employed during navigation of an airborne odor plume are less clear. In previous experiments where rodents searched in airborne odor plumes, mice developed a memory-based strategy of serially sampling each possible reward location for the presence of odor, turning search tasks into detection tasks (Bhattacharyya and Singh Bhalla, 2015; Gire et al., 2017). Thus, it remains unclear whether mammals can follow noisy concentration gradients under turbulent conditions.

To better understand the sensory computations and sampling strategies for olfactory search, we designed a two-choice behavioral assay where mice use olfactory cues to locate an odor source while we monitor sniffing and movements of the head, nose, and body. We found that mice use a concentration gradient-guided search strategy to navigate olfactory environments that contain turbulent flow. We found that these navigational behaviors are robust to perturbations including introduction of a novel odorant, varying the concentration gradient, and naris occlusion. Given the fundamental importance of sniffing to olfactory function, we hypothesized that mice would selectively sample the environment such that nose movement would be tightly coupled to respiration. Consistent with this hypothesis, we found that mice synchronize rhythmic three-dimensional head movements with the sniff cycle during search. These sniff-synchronized movement rhythms are prominent during trials, and largely absent during the inter-trial interval, suggesting that sniff synchronous movement is a pro-active strategy rather than a reactive reflex. To find structure in this search strategy, we used unsupervised computational methods to parse movement trajectories into discrete motifs. These movement motifs are organized into two distinguishable behavioral states corresponding to investigation and approach, reminiscent of the two-state olfactory search programs described in smaller organisms.

Temporally, investigation motifs lock to the sniff cycle with precision at a tens of milliseconds scale. Spatially, patterns of investigation and approach usage indicate a strategic advantage for investigating across the steepest part of the odor gradient. Our findings reveal the microstructure of olfactory search behavior in mice, identifying sensory computations and movement strategies that are shared across a broad range of species.

2.3 Results

2.3.1 Olfactory search in noisy gradients of airborne odor

We developed a two-alternative choice task in which freely-moving mice report odor source location for water rewards (Methods and Fig. 2.1A). To capture the search behavior, we measured respiration using nasal thermistors (McAfee et al., 2016) and video-tracked the animal's body, head, and nose position in real time at 80 frames/s (Fig. 2.1B,E and Fig. 2.11). The mouse initiates a trial by inserting its nose in a port (Fig. 2.1C; "Initiation"), which activates odor release from two ports at the opposite end of the arena. The mouse reports the location of higher odor concentration by walking toward it (Fig. 2.1C; "Search"). In previous studies, rodents performing olfactory search tasks developed memory-guided foraging strategies. In essence, animals run directly to potential odor sources and sample each in turn, thus converting the search tasks to detection tasks (Bhattacharyya and Singh Bhalla, 2015; Gire et al., 2017). To prevent mice from adopting sample-and-detect strategies, our task forces mice to commit to a decision at a distance from the actual source. Using real-time video-tracking (Lopes et al., 2015), we enforced a virtual "decision line", such that the trial outcome is determined by the mouse's location when it crosses this decision line (Fig. 2.1C; "Outcome"). For stimuli, we deliver odor from two separate flow-dilution olfactometers, giving independent control over odor concentration on

the two sides. To test olfactory search over a range of difficulties, we presented four odor patterns, defined by the ratio of odor concentration released from the two sides (100:0, 80:20, 60:40, 0:0).

We measured the spatiotemporal distribution of odor using a photoionoization detector (PID) in a 5x7 grid of sampling locations (Fig. 2.1D and Fig. 2.12). Pinene was used for the majority of experiments, because it is a neutral-valence odorant that is sensitively detected by the PID. As designed, varying the concentration ratios produced across-trial averaged gradients of different magnitudes. Airflow in the arena

---

Figure 2.1 (next page). Behavioral assay for freely-moving olfactory search. **A)** Diagram of experimental chamber where mice are tracked by an overhead camera while performing olfactory search. **B)** Top. Nose and head position are tracked using red paint at the top of the head. Sniffing is monitored via an intranasally implanted thermistor.Bottom. Example of sniffing overlaid on a trace of nose position across a single trial. **C)** Diagram of trial structure. Initiation. Mice initiate a trial via an initiation poke (grey oval). Search. Odor is then released from both odor ports (grey rectangles) at different concentrations. Outcome. Mice that cross the decision line (red) on the side delivering the higher concentration as tracked by the overhead camera receive a reward at the corresponding water port (blue ovals). **D)** Color maps of average odor concentration across 15 two-second trials captured by a 7x5 grid of sequential photoionization detector recordings. Rows represent side of stimulus presentation (left or right). Odor concentration beyond the decision line were not measured. **E)** Comparison of sniff recordings taken with an intranasally implanted thermistor and intranasally implanted pressure cannula. These are implanted on the same mouse in different nostrils. Top. Example trace of simultaneous pressure cannula (blue) and thermistor (red) recordings with inhalation points (as detected in all future analyses) overlaid on the traces in their respective colors. Bottom Left. Histogram of peak latencies (pressure inhalation onset – thermistor inhalation onset). 14/301 inhalations (4.7%) were excluded as incorrect sniff detections. These were determined as incorrect, because they fell more than 2 standard deviations outside the mean in peak latency (mean = 1.61585ms, SD = $\pm$14.93223ms). Bottom Right. Peak latencies, defined as the difference between pressure inhalation onset and thermistor inhalation onset, plotted against instantaneous sniff frequency.

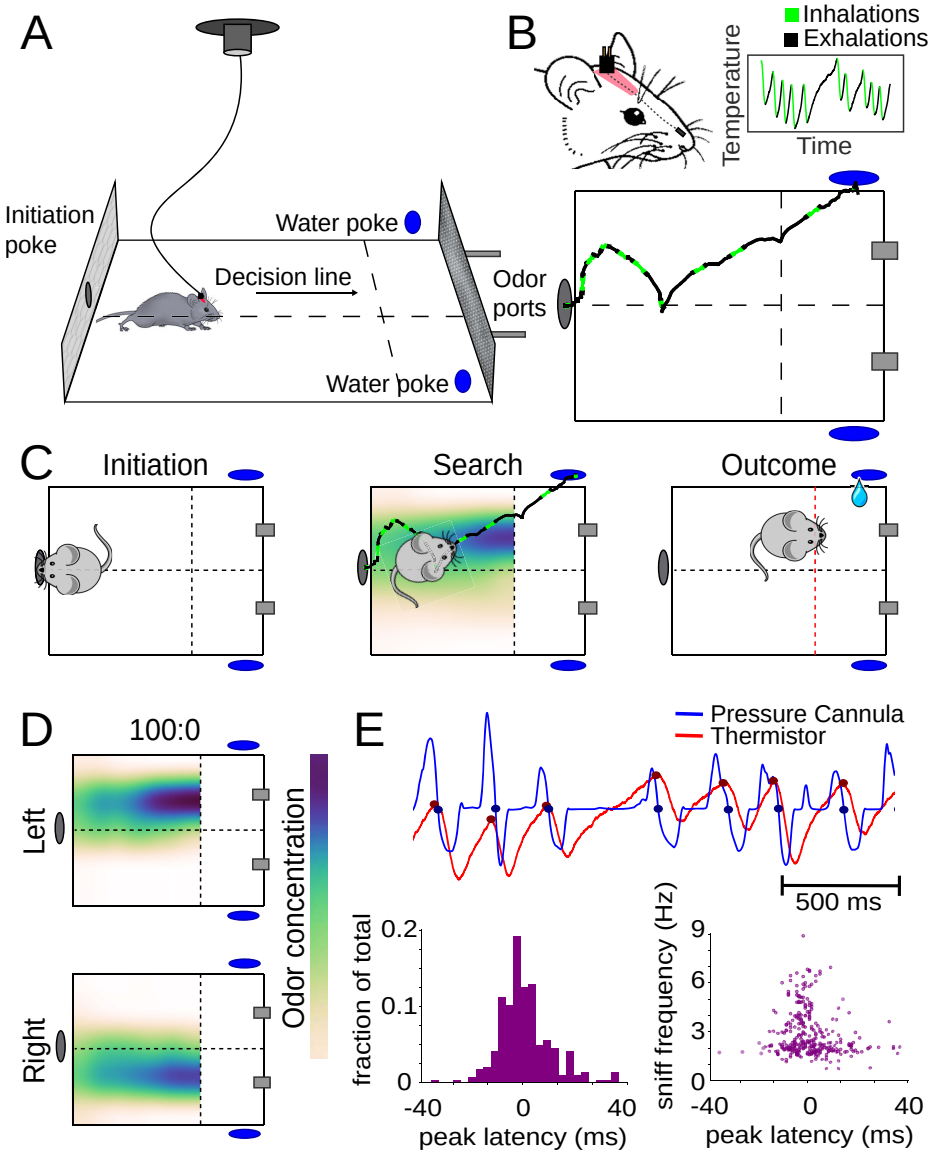

Figure 2.1.

is turbulent, imposing temporal fluctuations on the odor gradient (Video 2). Thus, our assay tests an animal's ability to navigate noisy odor gradients.

Mice learn the olfactory search task rapidly and robustly. We trained mice in the following sequence (Fig. 2.2A): First, naïve, water-restricted mice obtained water rewards from all ports in an alternating sequence ("water sampling"). In the next phase of training, we added odor stimulation such that odor delivery alternated in the

same sequence as reward, so that the mice would learn to associate odor with reward ports ("odor association"). Following these initial training steps, mice were introduced to the olfactory search paradigm. Odor was pseudo-randomly released from either the left or right odor source ("100:0"), signaling water availability at the corresponding reward port. Almost all mice performed above chance in the first session (Fig. 2.2B; binomial test, $p < 0.05$ for 24 out of 25 mice, $75 \pm 9.2\%$ correct, mean $\pm$ s.d.). Within 4 sessions, most animals exceeded 80% performance (19 out of 26). Following 100:0, mice were introduced to the 80:20 condition with mean performance across mice in the first session reaching 60% (Fig. 2.13D). Most subjects improved to exceed 70% performance over the next 7 sessions (17 out of 24). The mice that did not were excluded from subsequent experiments.

Next, we tested whether mice trained to search pinene plumes would generalize their search behavior to a novel odorant. We chose vanillin as the novel odorant, because, unlike pinene, vanillin does not activate the trigeminal fibers of the nose (Cometto-Muniz and Abraham, 2010; Doty et al., 1978; Hummel et al., 2009). Thus we could test whether trigeminal chemosensation is necessary for performance in our task. We found no differences in performance between vanillin and pinene sessions for these mice (Fig. 2.14A; Wilcoxon rank-sum test, $p = 0.827$, $n = 3$). These data suggest that this search behavior generalizes across odors and does not rely on the trigeminal system.

2.3.2 Mice can use gradient cues in turbulent flow

We reasoned that mice would solve this task using odor gradient cues. To vary odor gradients between trials, we trained mice in sessions with interleaved concentration ratios (100:0, 80:20, 60:40) across the trials of a session. In addition to these concentration ratios, odor omission probe trials (0:0) were randomly interleaved

into all experimental sessions. During these trials, airflow was identical to 80:20 trials, but air was directed through an empty vial rather than a vial containing odorant solution. These odor omission trials served a two-fold purpose: they acted as controls to ensure behavior was indeed odor-guided and they allowed us to observe how absence of odor impacts search behavior. On these probe trials, mice performed at chance (binomial test, $p = 0.9989$), with longer trial durations (Wilcoxon rank-sum test, $p < 0.05$) and more tortuous trajectories (Wilcoxon rank-sum test, $p < 0.05$) than on non-probe trials (Fig. 2.2C; $n = 19$, all data from 80:20 condition with probe trials). Performance drops with the concentration ratio (ΔC), consistent with our reasoning that mice would use odor gradient cues in this task (Fig. 2.2D; pairwise

---

Figure 2.2 (next page). Mice use concentration gradient cues in turbulent flow to perform search. **A)** Initial training steps. Water Sampling. In this task, mice alternate in sequence between the initiation, left, and right nose pokes to receive water rewards. Odor Association. Next, mice run the alternation sequence as above with without water rewards released from the initiation poke, making its only utility to initiate a trial. Further, odor is released on the same side of water availability to create an association between odor and reward. Odor Search. Here, mice initiate trials by poking the initiation poke. Odor is then randomly released from the left or right odor port. Correct localization (see Fig. 2.1C, decision line) results in a water reward and incorrect is deterred by an increased inter-trial interval (ITI). **B)** Performance curve across sessions for the Odor Search (100:0) training step ($n = 26$). **c-f)** Session statistics for four different experiments. Each colored line is the average of an individual mouse across all sessions, black points are means across mice, and whiskers are ± 1 standard deviation across mice. Top. Percent of correct trials. Middle. Average trial duration. Bottom. Average path tortuosity (total path length of nose trajectory/shortest possible path length). **C)** Odor omission. The 80:20 concentration ratio (Fig. 1) and odor omission (0:0) conditions randomly interleaved across a session. Data shown includes all sessions for each mouse ($n = 19$). **D)** Variable ΔC, Constant $|C|$. Three concentration ratio conditions (100:0, 80:20, 60:40) randomly interleaved across a session. Data shown includes all sessions for each mouse ($n = 15$). **E)** Constant ΔC, Variable $|C|$. Concentration ratio conditions 90:30 and 30:10 randomly interleaved across a session ($n = 5$). Data shown for first session only. **F)** Data shown includes all naris occlusion sessions, even if the mouse did not perform under every experimental condition.

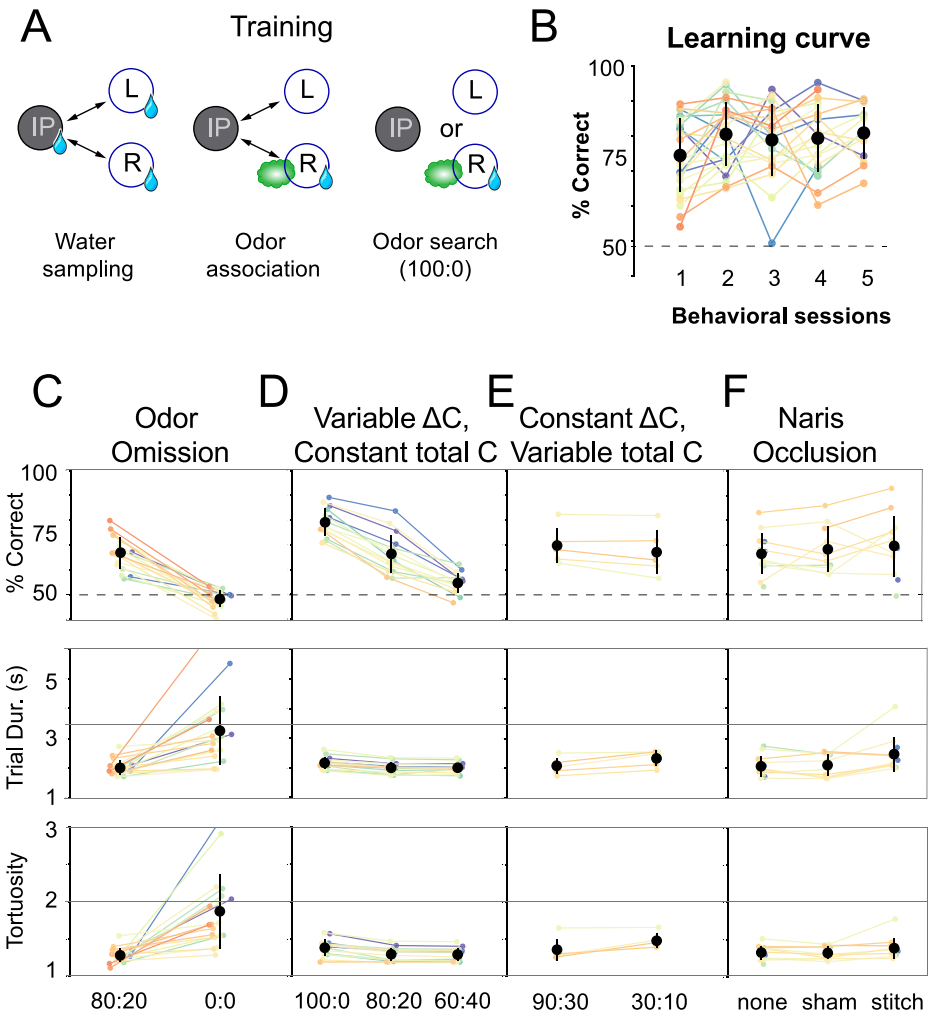

Figure 2.2.

Wilcoxon rank-sum tests, $p < 0.05$, $n = 15$). Varying the concentration ratio from 80:20 to 60:40 did not affect trial duration or path tortuosity, defined as actual path length divided by direct path length (Fig. 2.2D; pairwise Wilcoxon rank-sum tests, $p > 0.05$). However, trial duration and path tortuosity were slightly, but statistically significantly longer in the 100:0 condition (pairwise Wilcoxon rank-sum tests, $p < 0.05$).

Given that these results were obtained using a single absolute concentration ($|C|$) across ratios, mice could be solving our task with two distinct categories of sensory

computation. One possibility is that information about source location is extracted from the odor gradient. An alternative strategy would be to make an odor intensity judgement that gates a response to positional information from non-olfactory cues, such as wind direction, visual landmarks, or self-motion. This computation would be reminiscent of the odor-gated visual and mechanosensory behaviors observed in insects (Álvarez-Salvado et al., 2018; Kennedy and Marsh, 1974; van Breugel and Dickinson, 2014). To distinguish between these possible strategies, we tested mice in sessions interleaving the air dilution ratios 90:30 and 30:10. 30 is the correct answer in one condition and incorrect in the other, so that mice cannot use an intensity judgement strategy to perform well in both ratio conditions. In both conditions, mice performed equally well in the first session of training (Fig. 2.2E; Wilcoxon rank-sum test, $p = 0.465$, $n = 5$). This equal performance is true within the first 20 trials of the session (Fig. 2.14; Wilcoxon rank-sum test, $p = 0.296$). These results indicate that odor gradients guide olfactory search under these conditions.

We next asked how the mice are sensing the concentration gradient. Many mammals can use stereo-olfaction: comparing odor concentration samples between the nares (Catania, 2013; Parthasarathy and Bhalla, 2013; Porter et al., 2007; Rabell et al., 2017; Rajan et al., 2006). To test the role of stereo comparisons in our olfactory search task, we performed naris occlusion experiments. Mice were tested in three conditions on alternating days: naris occlusion, sham occlusion, and no procedure. We found that naris occlusion did not significantly impact performance or path tortuosity (pairwise Wilcoxon rank-sum tests, $p > 0.05$). When compared with the no-stitch condition, the naris stitch condition resulted in a slight, but statistically significant, increase in trial duration (pairwise Wilcoxon rank-sum test, $p < 0.05$).

This is not true when the stitch condition is compared with the sham condition (pairwise Wilcoxon rank-sum test, $p > 0.05$) indicating this may be a result of undergoing a surgical procedure. These overall results indicate that stereo comparison is not necessary in this task (Fig. 2.2F; n = 13), and that temporal comparisons across sniffs (Catania, 2013; Parabucki et al., 2019) play a larger role under our task conditions.

2.3.3 Sniff rate and occupancy are consistent across trials and gradient conditions

To investigate active sampling over the time course of trials, we tracked the animals' sniffing, position, and posture during behavioral sessions. Next, the mice emitted a rapid burst of sniffs, then sniffed more slowly as they approached the target (Fig. 2.3A). In this active behavioral state, inhalation and sniff durations were shorter during trials than during inter-trial intervals ($p \ll 0.01$ for all mice; Kolmogorov-Smirnov test; Fig. 2.3B,C), and strikingly shorter than those observed in head-fixed rodents (Bolding and Franks, 2017; Shusterman et al., 2011; Wesson et al.,

---

Figure 2.3 (next page). Distributions of sniffs and nose positions during search task. **A)** Above. Sniff raster plot for three sessions. Each black point is an inhalation, each row is a trial aligned to trial initiation (dashed line). Rows are sorted by trial length. Blue region represents trial initiation to trial end. Below. Mean instantaneous sniff rate across all trials for all mice aligned to time from trial initiation. Thin lines are individual mice, the thick line is the mean across mice, and shaded region is ± 1 standard deviation. **B)** Histogram of inhalation duration time across all mice (n = 11). Thick lines and shaded regions are mean and ± 1 standard deviation, thin lines are individual mice. Green: within-trial sniffs, Pink: inter-trial interval sniffs. **C)** Histogram of sniff duration time across all mice (n = 11). **D)** The nose traces of each trial across a single session, colored by chosen side. **E)** Location of all inhalations across a single session, colored by chosen side.**F)** Two-dimensional histogram of occupancy (fraction of frames spent in each 0.5 cm^2 bin). Colormap represents grand mean across mice (n = 19). **G)** Grand mean sniff rate colormap across mice (n = 11).

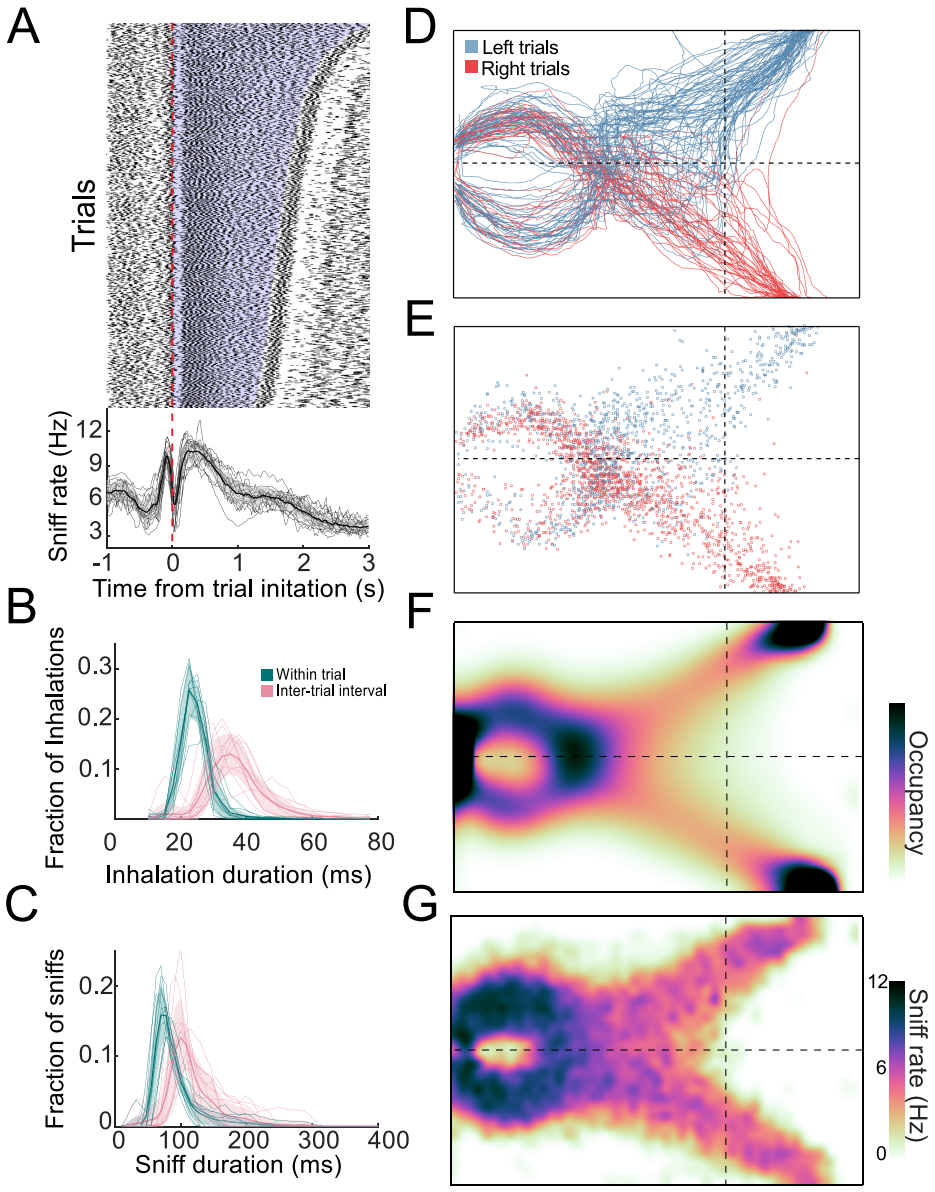

Figure 2.3.

2009). After the decision, there is a second rapid burst of sniffing followed by a long exhalation or pause during reward anticipation and retrieval (Fig. 2.3A). The overall sniff pattern was consistent across trials with an inhalation just before trial initiation followed by a long exhalation or pause at the beginning of the trial (Fig. 2.3A). During this sniffing behavior, the mice moved their nose through tortuous trajectories that were not stereotyped from trial to trial (Fig. 2.3D,E). Although individual mice

showed position biases (Fig. 2.15), these biases were not systematic across mice, so that the across-mouse mean occupancy distribution was evenly distributed across the two sides of the arena (Fig. 2.3F; n = 19). Consistent with this sniffing and movement pattern, the sniff rate was highest near the initiation port, and slower on the approach to target (Fig. 2.3G). These measures of active sampling were not statistically distinguishable across gradient or naris occlusion conditions, but changed significantly on odor omission probe trials, with more fast sniffing and head turns overall.

2.3.4 Mice synchronize three-dimensional kinematic rhythms with sniffing during olfactory search

To test the hypothesis that nose movement locks to respiration during olfactory search, we aligned movement dynamics with the sniff signal. Using Deeplabcut (Mathis et al., 2018; Mathis and Mathis, 2020), we tracked the position of three points: tip of snout, back of head, and center of mass (Fig. 2.4A). From the dynamics of these three points, we extracted the kinematic parameters nose speed, head yaw velocity, and Z-velocity (Fig. 2.4B-D). Synchrony between movement oscillations and sniffing is apparent on a sniff-by-sniff basis (Fig. 2.5), consistent across mice, and selectively executed during olfactory search. On average, nose speed accelerates during exhalation, peaks at inhalation onset, and decelerates during inhalation (Fig. 2.5A.i).

Head yaw velocity, which we define as toward or away (Fig. 2.4; centripetal or centrifugal) from the body-head axis, reaches peak centrifugal velocity at inhalation, decelerates and moves centripetally over the course of inhalation (Fig. 2.5A.ii). Although our videos are in two dimensions, we can approximate movement in depth by analyzing the distance between the tip of the snout and the back of the head

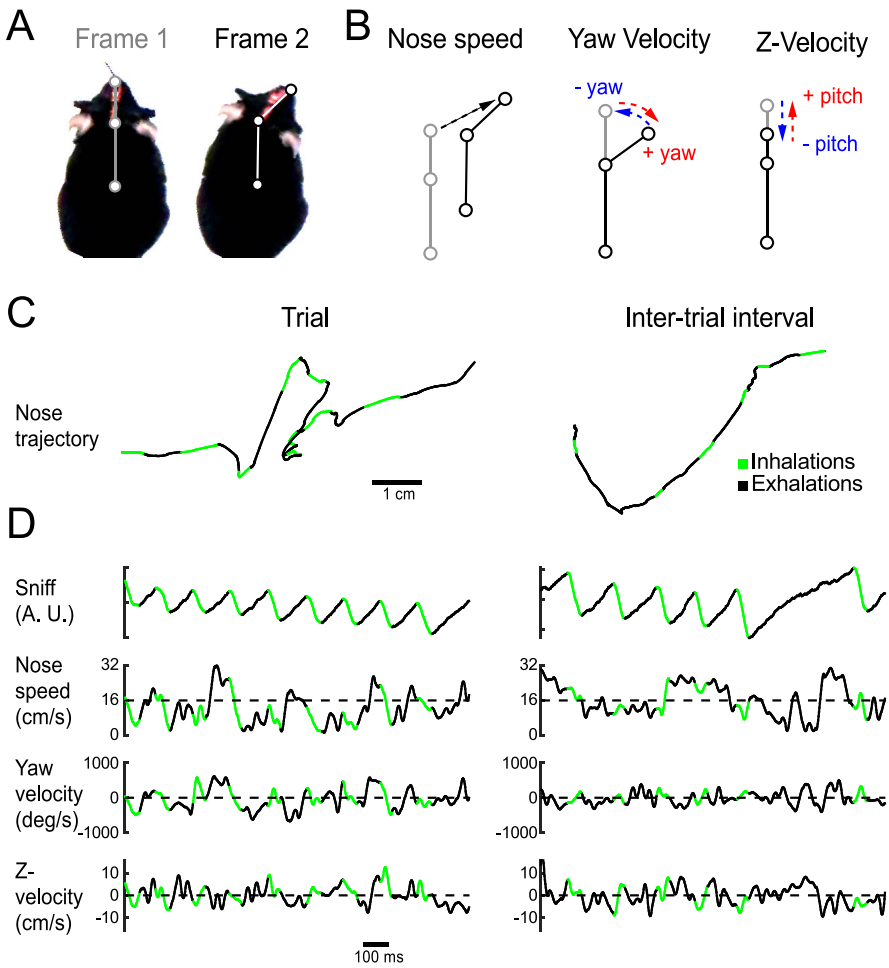

Figure 2.4. Quantifying kinematic parameters during olfactory search. **A)** Schematic of kinematic parameters. Left. Two example frames from one mouse, with the three tracked points marked: tip of snout, back of head, and center of mass. **B)** Quantified kinematic parameters: "nose speed": displacement of the tip of the snout per frame (12.5ms inter-frame interval). "Yaw velocity": change in angle between the line segment connecting snout and head and the line segment connecting head and center of mass. Centrifugal movement is positive, centripetal movement is negative. "Z-velocity": change in distance between tip of snout and back of head. Note that this measure confounds pitch angle and Z-axis translational movements. **C)** Segments of example trajectories. Left. The trajectory of the nose during one second of trial time. Green: path during inhalations. Black: path during the rest of the sniff. Right. Same for an inter-trial interval trajectory. **D)** Traces of sniff and kinematic parameters during the time windows shown in C. Color scheme as in C

(Fig. 2.4B). This measure confounds pitch angular motion and vertical translational motion, so we conservatively refer to this parameter as "z-velocity". Because mice point their head downward during task performance, shortening of the distance between the tip of the snout and the back of the head indicates downward movement,

while increases in the distance correspond to upward movements. The z-velocity reaches peak upward velocity at inhalation onset, decelerates and goes downward during inhalation, and rises again at exhalation (Fig. 2.5A.iii). These modulations were absent from trial-shuffled data (Fig. 2.16; permutation test, $p < 0.001$). Cross-correlation and spectral coherence analysis further demonstrate the synchrony between nose movement and sniffing (Fig. 2.5B,C). These results demonstrate that kinematic rhythms lock to sniffing with tens of millisecond precision, consistent with a previous report demonstrating that rats make similar movements during novel odor-evoked investigative behavior (Kurnikova et al., 2017). Our findings show that precise cycle-by-cycle synchronization can also be a feature of goal-directed odor-guided behavior. Mice selectively deploy this pattern of sniff-synchronized three-dimensional nose movement. For nose speed, yaw velocity, and z-velocity, sniff synchrony is significantly reduced during the inter-trial interval when the mouse is returning from the reward port to initiate the next trial, even when the mouse is sniffing rapidly. Modulations in nose speed were slightly different than trial-shuffled data, showing that sniff-synchronized movement is not totally absent during the inter-trial interval, whereas modulations in yaw velocity and z-velocity were indistinguishable from trial-shuffled data (Fig. 2.16). This difference between within-trial and between-trial sniff synchrony was not contingent on the mouse's slower nose speed during the ITI (Fig. 2.5A). Kinematic synchrony was the same when only periods of high-speed nose movement in the ITI are included in the analysis (Fig. 2.17). This reduction of kinematic synchrony when the mouse is not performing the task suggests that sniff synchronized movement is not an inevitable biomechanical accompaniment to fast sniffing, but rather reflects a strategic behavioral state. Further support for this

idea comes from analyzing time intervals when the mouse attempts to initiate a trial before the end of the inter-trial interval.

After such premature attempts at trial initiation, the mice execute sniff-synchronized movement, despite the absence of the experimenter-applied odor stimulus (Fig. 2.17). Lastly, sniff synchrony changes dramatically in the time interval between crossing the virtual decision line and entering the reward port, when odor is still present yet the animal has committed to a decision (Fig. 2.17). Taken together, our observations indicate that sniff synchronous movement is a proactive, odor-seeking strategy rather than a reactive, odor-gated reflex.

2.3.5 State space modeling finds recurring motifs that are sequenced diversely across mice

In our olfactory search paradigm, the overall rhythm of nose movement synchronizes with sniffing (Fig. 2.4 and 2.5), and yet the mice move through a different trajectory on every trial (Fig. 2.3D). Given this heterogeneity, it was not obvious to us how to best quantify common features of movement trajectories across trials and

---

Figure 2.5 (next page). Kinematic rhythms synchronize with the sniff cycle selectively during olfactory search. i-iii) Nose speed, yaw velocity, and z-velocity respectively (see Fig. 2.4 for definitions). **A)** Top. Color-plot showing movement parameter aligned to inhalation onset for within-trial sniffs taken before crossing the decision line. Taken from one mouse, one behavioral session. Dotted line at time 0 shows inhalation onset, the second line demarcates the end of the sniff cycle, sorted by duration. Data are taken from one behavioral session. Middle. Color-plot showing each movement parameter aligned to inhalation onset for inter-trial interval sniffs taken before the first attempt at premature trial initiation. Bottom. Sniff-aligned average of each movement parameter. Thin lines represent individual mice (n = 11), bolded lines and shaded regions represent the grand mean ± standard deviation. Green: within-trial sniffs, Pink: inter-trial interval sniffs. **B)** Normalized cross correlation between movement parameter and sniff signal for the same sniffs as above. **C)** Spectral coherence of movement parameter and sniff signal for the same sniffs as above.

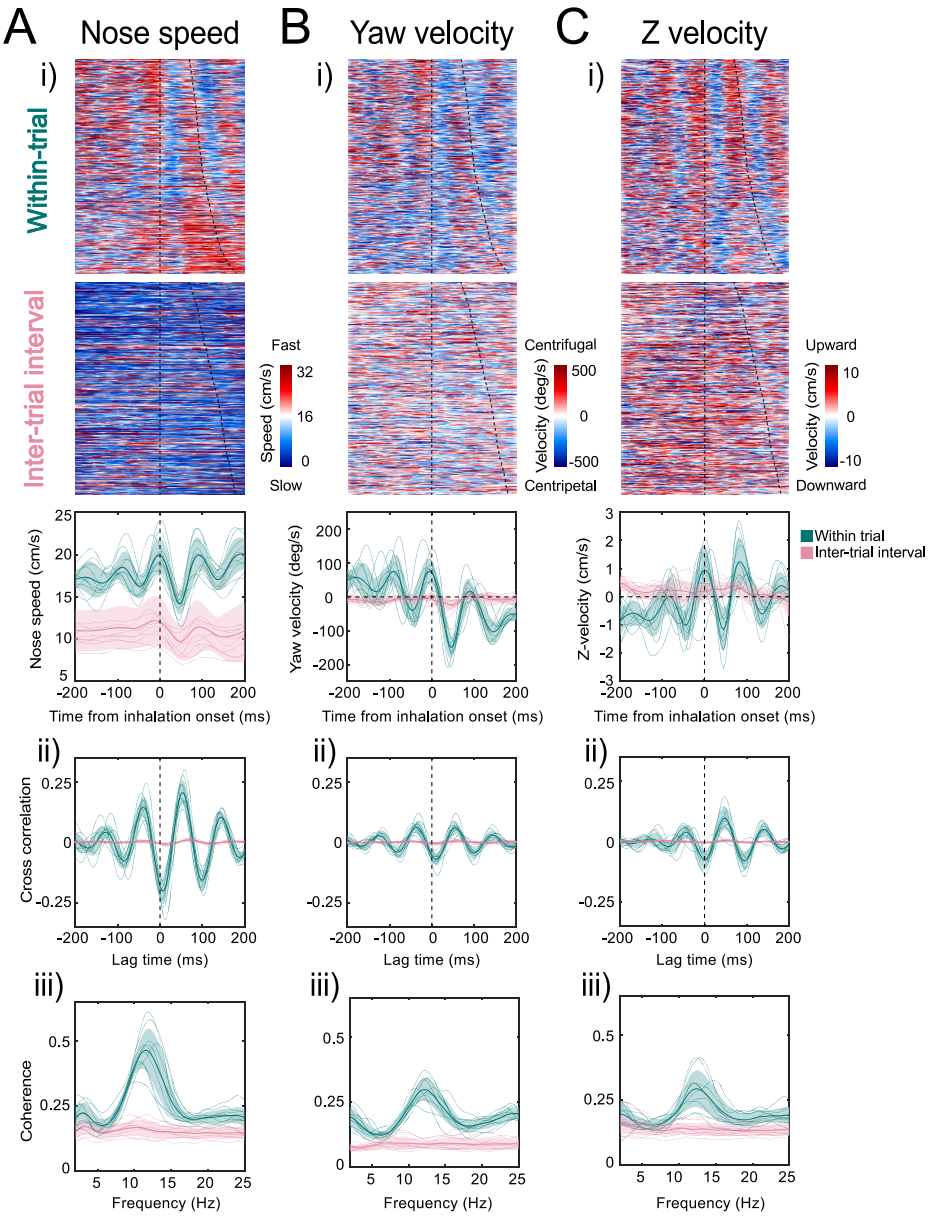

Figure 2.5.

subjects. Rather than guess at suitable features, we used an unsupervised machine learning tool, modeling the movement data with an Auto-Regressive Hidden Markov Model (AR-HMM) (Murphy, 2012; Poritz, 1982). This model parses continuous sequential data into a discrete set of simpler movement motif sequences, similarly to "Motion Sequencing" (MoSeq) (Wiltschko et al., 2015). We fit AR-HMMs to the

allocentric three-point coordinate data (front of snout, back of head, and center of mass) pooled across a subset of mice and trial conditions (See Methods and Fig. 2; e.g. 80:20, 90:30, nostril stitch). Models were then tested for their ability to explain a separate set of held-out trials (see Methods). These models defined discrete movement patterns, or "motifs", that recur throughout our dataset (e.g., Fig. 2.6A). We fit different AR-HMM models each constrained to find a particular number of motifs (between 6 and 100) and found that the cross-validated log-likelihood of these fits continued to rise up to 100 motifs (Fig. 2.18). For visualization, we will focus on a model with 16 states, which we narrow to 11, by excluding rare motifs that take up $<5\%$ of the assigned video frames (Fig. 2.6B and 6-figure supplement 1C,D). Models with more or fewer states gave equivalent results (Fig. 2.19-2.21).

The motifs extracted by this model have interpretable spatiotemporal trajectories on average (Fig. 2.6B; Video 6), although averaging masks considerable across-instance variability (Video 7). Across trials for a given mouse, motifs occurred in consistent but non-stereotyped sequences (Fig. 2.6C; Video 5). Across mice, the model identified consistent behavioral features as motifs (Fig. 2.22), but most mice were uniquely identifiable from how they sequenced motifs across trials. A classifier trained to decode mouse identity from the motif sequences on a trial by trial basis was able to perform above chance for 8 out of 9 mice (Fig. 2.6D; $p < 0.01$). Across the different concentration ratios (Fig. 2.2D), movement sequences were not statistically distinguishable (Fig. 2.6E). The only condition that gave distinguishable motif patterns were the odor omission trials (0:0), in which the mice made longer, more tortuous trajectories (Fig. 2.2C). Thus, although this model is sensitive enough to decode mouse identity (Fig. 2.6D), it does not detect stimulus-dependent modifications of sampling behavior, suggesting that the mice do not modify

their sampling behavior in a gradient-dependent manner, at least in the movement parameters we measured. This lack of modification ran counter to our expectations, because we reasoned that making the task harder would make the mice adjust their strategy to maintain high performance. We speculate that this absence of an adaptive strategy is due to impulsivity.

2.3.6 Movement motifs reveal two-state organization of olfactory search

Many behaviors have hierarchical structure that is organized at multiple temporal scales. Brief movements are grouped into progressively longer modules, and are ultimately assembled into purposive behavioral programs (Berman et al., 2016; Gallistel, 1982; Weiss, 1968).

Figure 2.6 (next page). Recurring movement motifs are sequenced diversely across mice and consistently across stimuli **A)** 8 example frames from one instance of a behavioral motif with tracking overlaid. **B)** Average motif shapes. Dots and lines show the average time course of posture for 8 frames of each of the 11 motifs (n = 9 mice). All instances of each motif are translated and rotated so that the head is centered and the head-body axis is oriented upward in the first frame. Subsequent frames of each instance are translated and rotated the same as the first frame. Time is indicated by color (dark to light). Background color in each panel shows the color assigned to each motif. **C)** Across-trial motif sequences for two behavioral sessions for one mouse. Trials are separated into trials where the mouse chose left and those in which the mouse chose right. Trials are sorted by duration. Both correct and incorrect trials are included. Color scheme as in B. **D)** Linear classifier analysis shows that mice can be identified from motif sequences on a trial by trial basis. Grayscale represents the fraction of trials from a given mouse (rows) that are decoded as belonging the data of a given mouse (columns). The diagonal cells represent the accuracy with which the decoded label matched the true label, while off-diagonal cells represent trials that were mislabeled by the classifier. Probabilities along rows sum to 1. Cells marked with asterisks indicate above chance performance (label permutation test, $p < 0.01$). **E)** Linear classifier analysis identifies odor omission trials above chance, but does not discriminate across odor concentration ratios (n = 9 mice). **F)** Cross-validated log-likelihood (evaluated on trials not used for model fitting) for fit AR-HMM models with different numbers of motifs, S, shows that model log-likelihood does not peak or plateau up to S = 100.

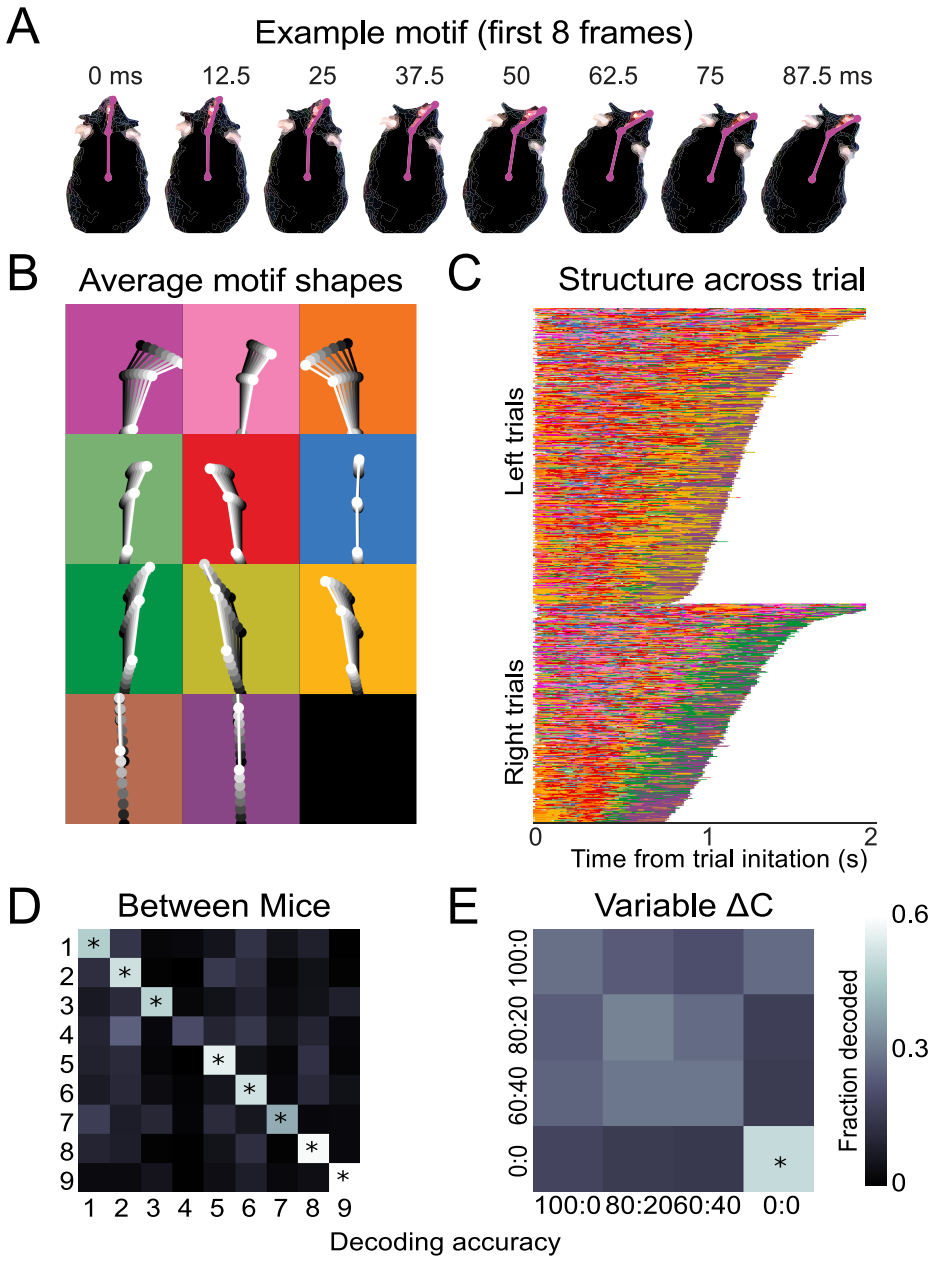

Figure 2.6.

Olfactory search programs in smaller organisms are often organized into two overarching states: move straight when concentration is increasing, and reorient when concentration is decreasing (Bargmann, 2006; Gomez-Marin et al., 2011; Kennedy and Marsh, 1974; Lockery, 2011; van Breugel and Dickinson, 2014; Vickers and Baker,

1994). We hypothesized that olfactory search motifs in mice are organized similarly. To reveal higher-order structure in the temporal organization of these motifs, we applied a clustering algorithm that minimizes the Euclidean distance between rows of the Markov transition matrix (i.e., purely based on the conditional probabilities of motifs following them). This clustering separated motifs into two groups (Fig. 2.7A), with several distinct properties. These properties were present in models with more or fewer states (Fig. 2.19-2.21). Based on these differences (see below), we label these groups as putative "investigation" and "approach" states. First, investigation and approach motifs cluster their onset times in the trial, with investigation motifs tending to occur early in the trial, while approach motifs tend to begin later (Fig. 2.7B). Grouping motifs into these higher-order states shows a consistent trial sequence, with trials beginning with investigation and ending with approach (Fig. 2.7C,D). Importantly, entering the approach state is not a final, ballistic commitment to a given water port – switches from approach back to investigation were common (Fig. 2.7C,D; Video 8). This pattern suggests that the mice are continuously integrating evidence about the odor gradient throughout their trajectory to the target. Second, these states correlated with distinct sniff rates and movement speeds. During investigation motifs, the mice moved more slowly and sniffed more rapidly, whereas the approach states were associated with faster movement and slower sniffing (Fig. 2.7E). Third, the sniff-synchronized kinematic rhythms (Fig. 2.4 and 5) were distinct in the two states (Fig. 2.7F; Kolmogorov-Smirnov test, $p < 0.01$). Specifically, nose speed and yaw velocity are more synchronized with sniffing during the investigation state (Fig. 2.7F). Given the consistent sequence from investigation to approach and given that mice sniff faster during the early part of trial, these differences in kinematic parameters could reflect across-trial tendencies instead of within-trial synchrony.

To test this possibility, we calculated the Kolmogorov-Smirnov statistic, which quantifies the difference between two cumulative distributions, for real and trial-shuffled data (Fig. 2.23). This analysis showed that nose speed and yaw velocity modulation exceeded what would be expected from across-trial tendencies (1000 shuffles, $p < 0.001$), while the z-velocity modulation did not ($p = 0.31$). Switches between the investigation and approach state mark behavioral inflection points that can be identified from trial to trial. We reason that these behavioral inflection points are a signature of key moments in the mouse's evolving decision process. Thus,

---

Figure 2.7 (next page). Behavioral motifs can be categorized into two distinct groups, which we putatively label as investigation (blue) and approach motifs (orange). **A)** Transition probability matrix. Grayscale represents the log probability with which a given motif (rows) will be followed by another (columns). Clustering by minimizing Euclidean distance between rows reveals two distinct blocks of motifs. We label the top-left block as "investigation" and the bottom-right block as "approach". **B)** Distribution of onset times for each motif, normalized by trial duration. Investigation motifs tend to occur early in trials, while approach motifs tend to occur later ($n = 9$ mice). **C)** Across-trial motif sequences for two behavioral sessions for one mouse, with motifs classified into investigation and approach. Trials are separated into correct trials (above) and incorrect trials (below). Motif sequences are sourced from the same data as Fig. 2.6C. **D)** Temporal details of investigation-approach transitions with overlaid sniff signal. Data come from a subset of trials shown in Fig. 2.7C. In the sniff signal, green represents inhalations, black represents the rest of the sniff. **E)** Investigation and approach motifs differ in nose speed and sniff rate. Individual markers represent one motif from one mouse. Marker shapes correspond to the individual mice ($n = 4$). Sniff rate and nose speed are normalized within mice. **F)** Investigation and approach motifs differ in the kinematic rhythms (same parameters as in Figures 4 and 5). Thin lines represent individual mice ($n = 4$), thick lines and shaded regions represent the grand mean $\pm$ standard deviation. Blue: within-trial sniffs, orange: inter-trial interval sniffs. Top. Nose speed modulation, defined by a modulation index (maxspeed-minspeed)/(max+min) calculated from the grand mean, is significantly greater for investigation motifs than approach motifs (Fig. 2.23; $p < 0.001$, permutation test). Middle: Yaw velocity modulation is significantly greater for investigation motifs than approach motifs (Fig. 2.23; $p < 0.001$, permutation test). Bottom: Z-velocity modulation does not significantly differ between approach motifs and investigation motifs ($p = 0.31$, permutation test).

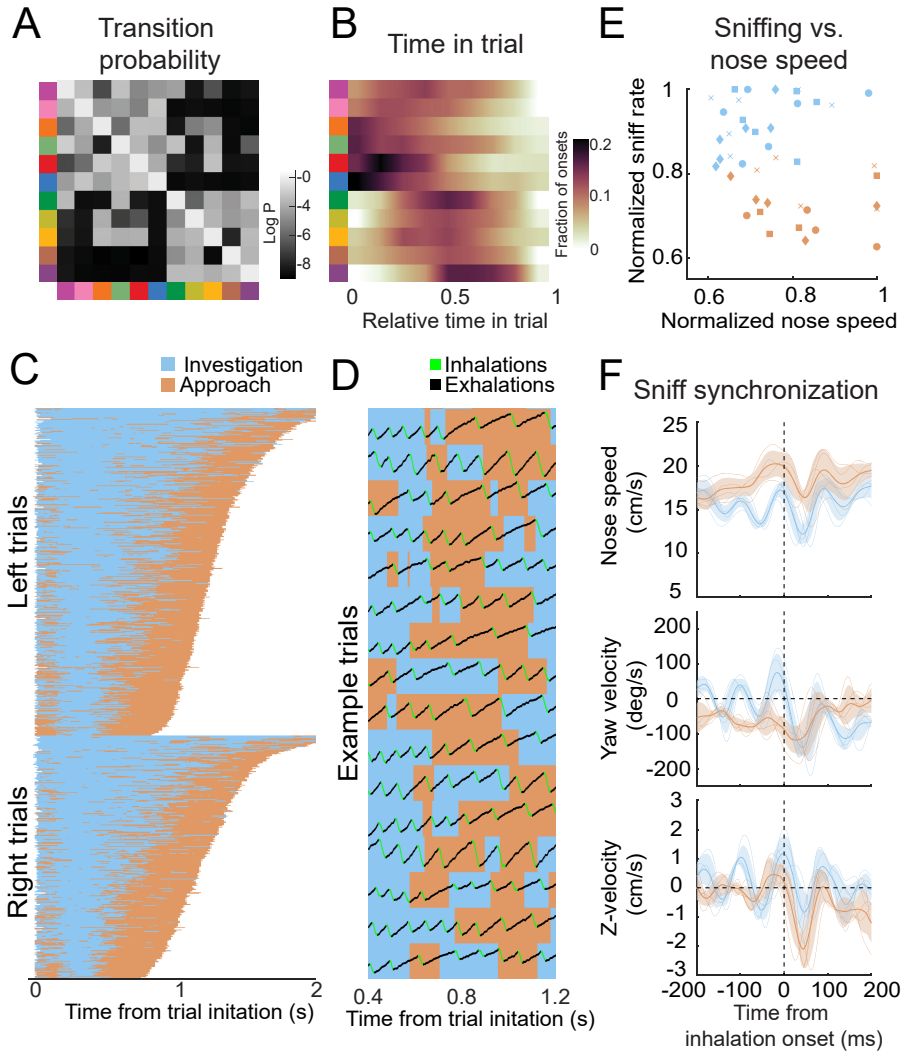

Figure 2.7.

our analysis can provide a framework for temporal alignment of diverse movement trajectories with simultaneously-recorded physiological data (Markowitz et al., 2018).

2.3.7 Investigation motif onsets are precisely locked to sniffing

If motif transitions correspond to relevant behavioral events, their temporal structure should correlate with the temporal structure of neural activity (Markowitz et al., 2018). During fast sniffing, respiration matches with the rhythms of head movement (Fig. 2.4 and 5), whisking, and nose twitches (Kurnikova et al., 2017;

Moore et al., 2013; Ranade et al., 2013). These motor rhythms correlate with activity in numerous brain regions, including brainstem, olfactory structures, hippocampus, amygdala, and numerous neocortical regions (Karalis and Sirota, 2018; Kay, 2005; Macrides et al., 1982; Vanderwolf, 1992; Yanovsky et al., 2014; Zelano et al., 2016). We hypothesized that movement motifs would lock with these behavioral and neural rhythms, so we aligned sniff signals with motif onset times. Importantly, the breath signal was not input to the model.

This alignment revealed a striking organization of motif sequences relative to the sniff rhythm. For example, the onset times of motif 6 (dark blue) occurred in a precise timing relationship with sniffing (Fig. 2.8A). To visualize the timing relationship between onsets of all motifs and sniffing, we calculated the equivalent of a peri-stimulus time histogram for inhalation times relative to the onset time of each motif, and took the grand mean across all mice (Fig. 2.8B; n = 4). Further, to determine how motif onset times are organized relative to the sniff cycle, for each motif we calculated a histogram of motif onset in sniff phase coordinates (Fig. 2.8C; relative position in the sniff cycle). Sharp peaks are apparent in both histograms for investigation motifs, and less so for approach motifs (quantified below; Fig. 2.8B,C). Importantly, these timing relationships are consistent across mice, with some motifs tending to occur early in the sniff cycle during inhalation, and others occurring later in the sniff cycle (Fig. 2.8D). Thus, parsing diverse movement trajectories into sequences of recurring movement motifs reveals additional sniff-synchronized kinematic structure in a consistent manner across mice.

Are motif onsets timed with respect to inhalation times, or do they coordinate with the entire sniff cycle? In other words, is motif onset probability more modulated in time or phase? To quantify the sniff synchronization of motif

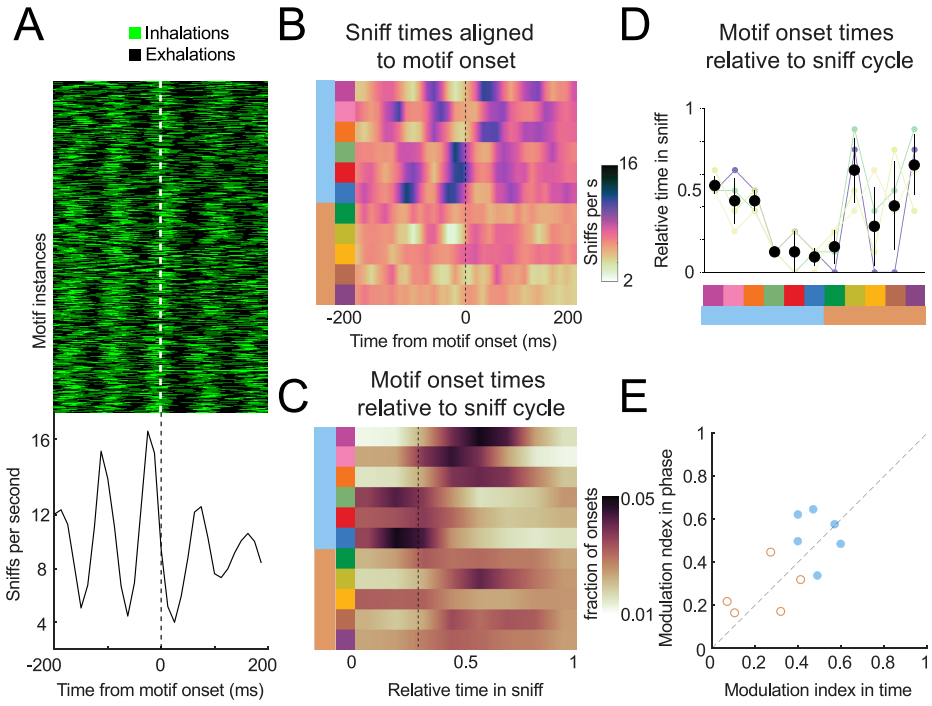

Figure 2.8. Motif onsets synchronize to the sniff cycle. **A)** Alignment of the sniff signal to an example motif. Top. Color scheme shows sniff cycles aligned to the onsets of motif 6 (blue). Motif instances are in chronological order. Green: inhalation, black: rest of sniff. Bottom. Peristimulus time histogram of inhalation times aligned to the onset of motif 6. **B)** Alignment of sniff signal to onset times of all motifs across mice (n = 4). Colormap represents the grand means for peristimulus time histograms of inhalation times aligned to the onset of motifs. **C)** Alignment of motif onset times in sniff phase. Colormap represents peristimulus time histograms of motif onsets (bin width = 12.5ms) times aligned to inhalation onset, with all sniff durations normalized to one. Dotted line shows the mean phase of the end of inhalation. **D)** Motifs alignment to sniff phase is consistent across mice. Thin lines represent individual mice, black points are means, and whiskers are ±1 standard deviation (n = 4 mice). **E)** Investigation motifs are more synchronized to the sniff cycle than approach motifs. Dots represent the modulation index in time on the x coordinates and in phase on the y-coordinates. Filled dots represent motifs that are significantly modulated in both time and phase ($p < 0.01$, permutation test). Half-filled dots represent motifs that are significantly modulated in time (left half filled) or phase (right half filled).

onset times, we calculated a modulation index (MI=(max-min)(max+min)) for each motif's across-mouse mean histogram (n = 4). To test whether these trial-by-trial modulation indices exceeded what would be expected from across-trial tendencies, we compared real and trial-shuffled data (Fig. 2.24). All investigation motifs were

significantly modulated for both time and phase coordinates (Fig. 2.8E; filled symbols, permutation test, $p < 0.001$), with some having higher MI in time, and others in phase. One approach motif was significantly modulated in time coordinates (Fig. 2.8E; right-half filled symbol, $p = 0.003$), while two approach motifs were significantly modulated in phase coordinates (Fig. 2.8E; left-half filled symbols, $p = 0.015$ and $p < 0.001$). Comparing the modulation indices between time and phase coordinates does not reveal a consistent pattern of modulation in time vs phase – some motifs had higher MI in phase, others in time. Thus, our data are inconclusive as to how motif onsets organize relative to the sniff cycle. Nevertheless, these analyses demonstrate that kinematic inflection points synchronize with breathing during olfactory search. Given that breathing synchronizes to other motor and brain rhythms, these motifs likewise correlate to the structure of activity of many neurons. Thus our analysis will be a useful tool to pinpoint behaviorally relevant activity in widespread brain regions.

2.3.8 Investigation and approach occupancy maps suggest a serial sniff comparison strategy

We propose that motif transition times indicate "decision points" at which the animal chooses its next move (Markowitz et al., 2018). The transitions between investigation and approach motifs are particularly relevant, since investigation motifs may correspond to an evidence-gathering state, while approach motifs may correspond to a reward-gathering state. What kind of sensory evidence guides transitions between investigation and approach? Although we cannot determine the precise odor inputs the mice acquire on a sniff-by-sniff basis, we reasoned that we could elucidate the search strategy by examining aggregate across-trial patterns in allocentric maps of investigation and approach occupancy.

As expected from the temporal structure of investigation and approach (Fig. 2.7C,D), the mice primarily investigate near the initiation port (Fig. 2.9A; n = 9 mice), and primarily approach close to the decision line (Fig. 2.9A; orange). Along the longitudinal axis of the arena, the two occupancy maps overlap in a region between initiation port and decision line (Fig. 2.9A; black) where overall occupancy also peaks (Fig. 2.3F). Along the lateral axis of the arena, on average the overlap region is roughly centered on the lateral midline between left and right sides (Fig. 2.9A). However, this centered position is not representative of the individual mice, which have their overlap region in different positions relative to the lateral midline, with some on the left, and others on the right (Fig. 2.25A). However, if trials are oriented such that the chosen side is always up in the occupancy maps, the overlap region is displaced toward the chosen side of the arena in all individual mice (Fig. 2.25B). Thus, the mice primarily switched states while located on the side they would ultimately choose. To quantify the overlap between states, we calculated a relative occupancy index, defined as the difference in investigation and approach occupancy divided by their sum (Fig. 2.9C; I.A.I). For this index, a bin where the mice primarily investigated has a positive value, while a bin primarily occupied during the approach state has a negative value. Along the longitudinal axis, most of the change in this index occurred between inflection points at 5 and 10 cm, which we define as a "transition zone" for the analyses below (Fig. 2.9D).

Along the lateral axis of the arena, I.A.I was quite variably distributed across mice, both for the entire occupancy map, and within the transition zone (Fig. 2.9E), consistent with the individual mouse occupancy maps (Fig. 2.25A). Orienting trials with respect to the chosen side demonstrates a clearer pattern, with primarily

investigation on the unchosen side, and primarily approach on the chosen side (Fig. 2.9F and 2.25B).

Occupancy maps allowed us to further evaluate hypotheses about the search strategy mice use in these conditions. One hypothetical strategy is that the mice memorize absolute concentrations across trials and compare each individual sniff to an internal threshold learned over previous trials (single-sniff hypothesis). Another possible strategy would be serial sniff comparison, where the mouse senses changes between sequential samples within individual trials (serial-sniff hypothesis). These hypotheses make distinct predictions about where the mouse should sample. For the single-sniff hypothesis, the most informative location to sample is directly downwind of the odor ports, where concentration differences between left and right trials are maximal (Fig. 2.12). For gradient sensing, the optimal location is instead across the lateral midline, where the gradients are sharpest (Fig. 2.12). We tested these

---

Figure 2.9 (next page). The allocentric spatial distribution of investigation and approach occupancy. **A)** Colormaps show two-dimensional histograms of the occupancy density (1 cm2 bins, n = 9 mice) with investigation density in blue, approach density in orange, and overlap shown by darker coloring (key in top-left corner). Histograms around the colormaps show the state occupancy projected onto the longitudinal (top) and lateral (left) axes of the arena. **B)** Occupancy distributions after the right-choice trials are flipped upward so that the chosen side is always facing up in the diagram. **C)** Relative usage is quantified with an investigation approach index (I.A.I.), defined as the difference between investigation and approach occupancy divided by their sum. Blue and orange triangles are visual aids that represent the I.A.I. **D)** Relative occupancy density of investigation and approach (I.A.I.), plotted along the longitudinal axis of the arena from the initiation port to the decision line. We define the region between 5 cm and 10 cm as a 'transition zone', in which most transitions between investigation and approach take place. Thin lines are individual mice (n = 9), thick line and shaded region are mean ± s.e.m. **E)** I.A.I. plotted along the lateral axis in real space (i.e., left-right orientation) for all occupancy throughout the arena (left) and for the transition zone only (right). Thin lines are individual mice (n = 9), thick line and shaded region are mean ± s.e.m. **F)** Same as e), but after the lateral axis has been reoriented so that the chosen side is always up.

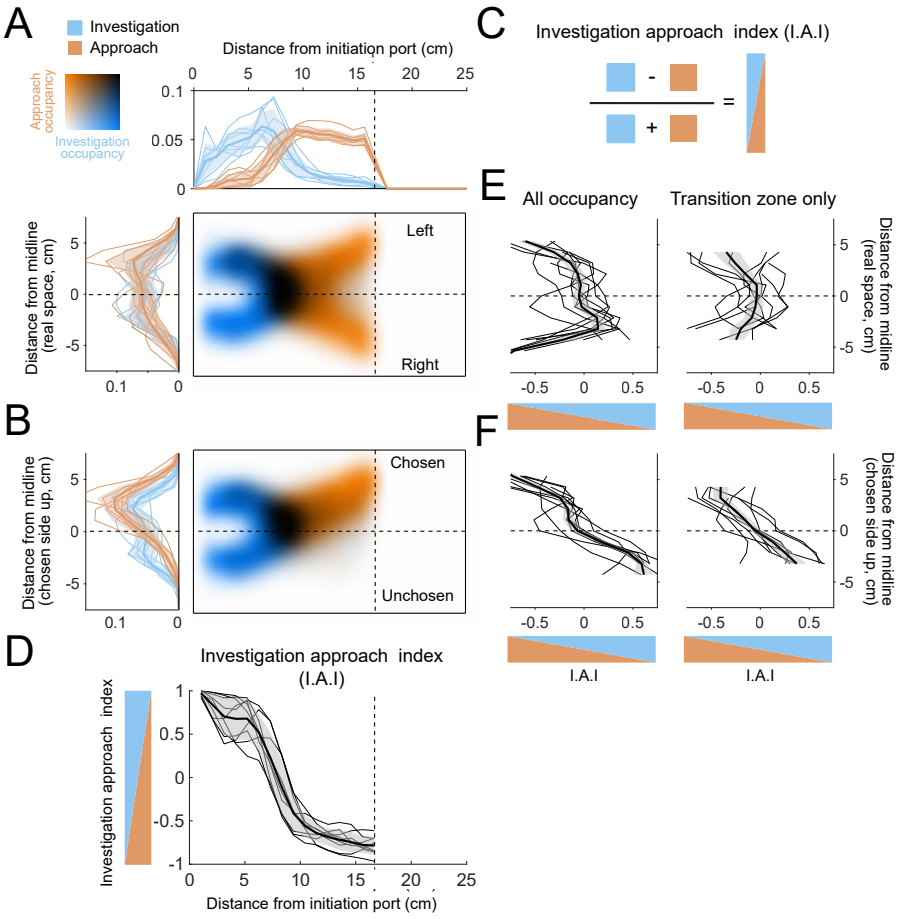

Figure 2.9.

predictions by comparing occupancy maps between correct and incorrect trials. For the single-sniff strategy, the mouse should get it correct more often when it investigates downwind of the odor ports, while a serial sniff hypothesis predicts that correct trials should show increased investigation at the midline.

Correct and incorrect trials yielded qualitatively similar occupancy maps (Fig. 2.10A). To quantify their differences, we first compared their occupancy indices along the longitudinal axis of the arena. (Fig. 2.10B,C). Correct trials featured significantly higher I.A.I (greater investigation) in the latter part of the transition zone, while past the decision zone the I.A.I was higher for incorrect trials (Fig. 2.10C; permutation

test, $p < 0.001$, $n = 9$ mice). Thus, increased investigation in the transition zone was associated with correct trials, while increased investigation near the decision line was associated with incorrect trials. This pattern suggests that investigation is not inherently advantageous to olfactory search irrespective of location. Instead, it matters where the mouse investigates, and some locations are less advantageous. Notably, absolute concentrations are most discriminable near the decision line (Fig. 2.12C), suggesting that mice may not be able to capitalize on this cue under these conditions.

We next quantified state occupancies along the lateral axis within the transition zone (Fig. 2.10D-F). Correct trials featured significantly increased investigation at and on the unchosen side of the midline relative to incorrect trials (Fig. 2.10G; permutation test, $p < 0.001$, $n = 9$ mice). By definition, occupancy of the unchosen side precedes a crossing of the midline to get to the chosen side. This suggests an advantage to sampling both sides of the midline, consistent with a serial-sniff gradient sensing strategy. Further, investigation more laterally, downwind of the odor port, was increased on incorrect trials, suggesting that sampling this location was not advantageous for task performance, contrary to the single-sniff absolute concentration hypothesis. Approach occupancy showed a different pattern, with significantly higher approach at and around the midline on incorrect trials, and a significant increase in approach occupancy closer to the chosen water port (Fig. 2.10H; permutation test, $p < 0.001$). Consistent with these observations, on correct trials I.A.I showed significant elevation at the midline and into the unchosen side of the arena, while increased I.A.I of the chosen side was associated with incorrect trials (Fig. 2.10I; permutation test, $p < 0.001$). Altogether, these results suggest that it is advantageous to sample both sides of the midline in this task, consistent with the serial-sniff hypothesis.

An important consideration in interpreting these results pertains to the construction of our task. Before choosing a side, the mice have to turn out of the initiation port in one direction or the other on every trial. On some trials they stay and choose the side of the first turn, while on other trials they switch and choose the other side. A single-sniff hypothesis predicts that if the mouse happens to turn first toward the correct side, it will tend to encounter above threshold concentrations during the turn, and should therefore tend to transition to approach without crossing

---

Figure 2.10 (next page). Occupancy maps indicate an advantage for investigation of both sides. **A)**Colormaps show two-dimensional histograms of the occupancy density (1 cm2 bins, n = 9 mice) with investigation density in blue, approach density in orange, and overlap shown by darker coloring (key in top-left corner). Histograms around the colormaps show the state density projected onto the longitudinal (top) and lateral (left) axes of the arena. Left: correct trials. Right: incorrect trials. **B)** Investigation approach index (I.A.I.) for correct (purple) and incorrect (green) trials. Thick lines and shaded region are mean ± s.e.m., thin lines are individual mice. **C)** Difference in I.A.I. between correct and incorrect trials along the longitudinal axis (2.5 cm bins, n = 9). Thick line is the across-mouse mean difference, thin gray lines are 1000 permutations in which correct and incorrect trial labels were scrambled. **D)** Investigation occupancy along the lateral axis, within the transition zone (5–10 cm longitudinal) for correct and incorrect trials. Thick lines and shaded region are mean ± s.e.m., thin lines are individual mice. **E)** Approach occupancy along the lateral axis, within the transition zone (5–10 cm longitudinal) for correct and incorrect trials. Thick lines and shaded region are mean ± s.e.m., thin lines are individual mice. **F)** I.A.I. along the lateral axis, within the transition zone (5–10 cm longitudinal) for correct and incorrect trials. Thick lines and shaded region are mean ± s.e.m., thin lines are individual mice. **G)** Difference in investigation occupancy between correct and incorrect trials along the lateral axis, within the transition zone. Thick blue line is the across-mouse mean difference, thin blue lines are 1000 permutations in which correct and incorrect trial labels were scrambled. **H)** Difference in approach occupancy between correct and incorrect trials. Thick orange line is the across-mouse mean difference, thin orange lines are 1000 permutations in which correct and incorrect trial labels were scrambled. **I)**Difference in I.A.I. between correct and incorrect trials. Thick orange line is the across-mouse mean difference, thin orange lines are 1000 permutations in which correct and incorrect trial labels were scrambled.

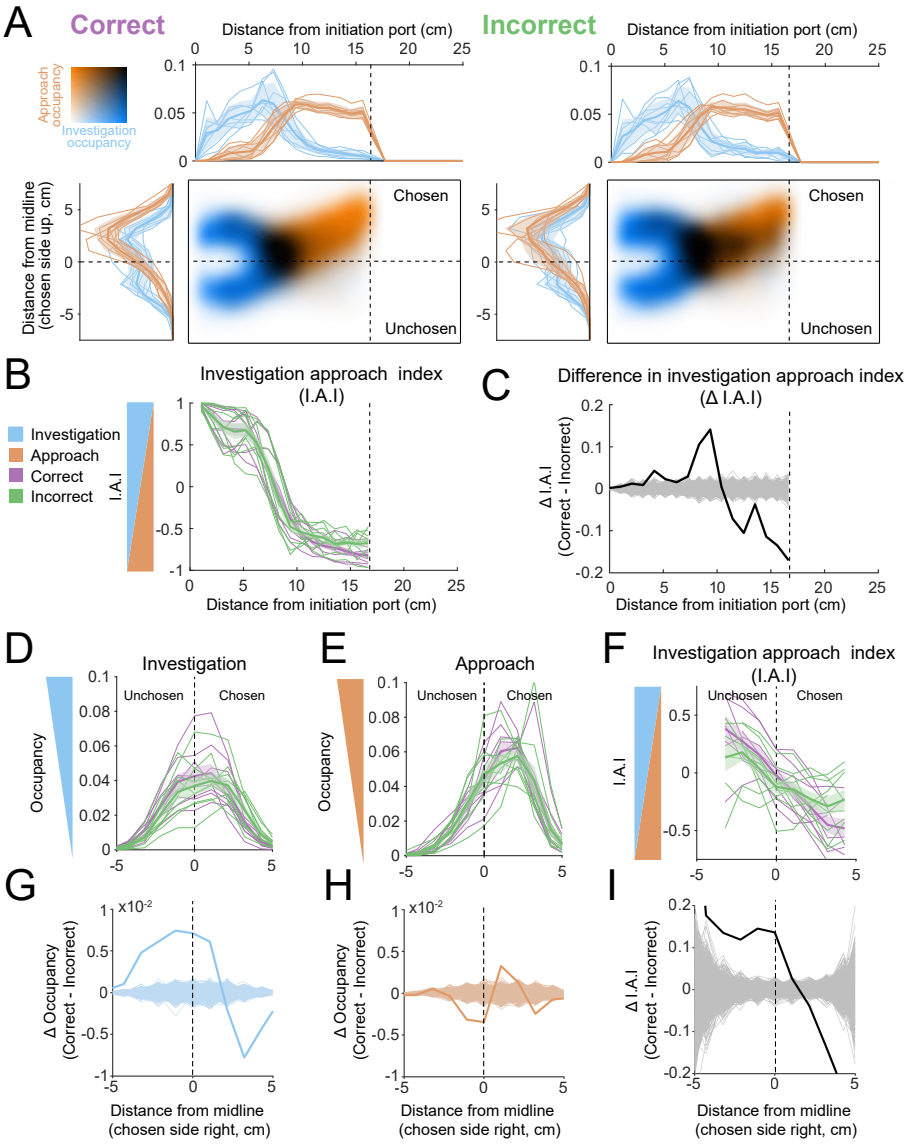

Figure 2.10.

the midline. However, if the mouse turns first to the incorrect side, threshold crossings will tend not to occur and the mouse can initiate approach before crossing the midline.

Thus, this hypothesis predicts that correct vs incorrect occupancy differences should occur at different positions along the lateral axis for stay and switch trials. To test this prediction, we performed the same analyses separately for stay and switch trials. Although not identical, both stay and switch trials showed significantly

increased investigation at and on the unchosen side of the midline for correct trials (Fig. 2.26). This analysis demonstrates that the apparent advantage of sampling across the midline is not an artefact of the asymmetry between switch and stay trials. Taken together, investigation and approach occupancy mapping provides further evidence suggesting that mice use a serial-sniff strategy to sense gradient cues in this task (Catania, 2013).

2.4 Discussion

This study elucidates sensory computations and movement strategies for olfactory search by freely-moving mice. Mice learn our behavioral task in days, after which they perform approximately 150 trials daily, sometimes for months. Task performance worsens for shallower odor gradients at a fixed absolute concentration, but is unaffected by varying absolute concentrations at a fixed concentration gradient. Taken together, these results show that mice can navigate noisy gradients formed by turbulent odor plumes. This gradient-guided search is robust to perturbations including novel odorant introduction and naris occlusion. These results give insight into sensory computations for olfactory search and constrain the possible underlying neural mechanisms.

Mice perform this task with a strategic behavioral program. During search, mice synchronize rapid three-dimensional head movements with fast sniffing. This synchrony is not a default accompaniment of fast sniffing – synchrony is absent when the mice are not searching. Movement trajectories are not stereotyped, but vary considerably across trials. To manage this complexity, we took an unsupervised computational approach to parse heterogeneous trajectories into a small number of movement motifs that recur across trials and subjects. This analysis captures common movement features across mice, but individual mice can be identified by how they

sequence these motifs. Our model was not constrained to find structure at a specific timescale, and consequently identified very brief, simple motifs. To find higher-order temporal structure in the data, we clustered motifs by their transition probabilities, which revealed two clear categories, putatively corresponding to investigation and approach. Investigation motifs tend to be executed early in the trial, and entail slower movement, faster sniffing, and more sniff synchrony than approach motifs. Even so, approach motifs are not ballistic commitments to an answer – switches from approach to investigation occurred on many trials. Lastly, the onset times of motifs were precisely locked to sniffing, with investigation motifs starting at characteristic phases of the sniff cycle.

The allocentric structure of investigation and approach suggests that the investigation state is not inherently advantageous. Rather, where the mouse investigates matters for performance. This dependence of performance on location indicates the spatial distribution of informative features in this olfactory scene. Notably, incorrect trials feature more investigation directly downwind of the odor source, along the axis of maximal odor concentration, which would be optimal if the mouse were using a single-sniff, absolute concentration strategy (Fig. 2.12C). Thus, these analyses provide further evidence that the mice do not capitalize on absolute concentration information to guide performance in this task. Instead, correct trials feature more investigation at and across the axis of maximal odor gradient (Fig. 2.12D), reminiscent of an object localization strategy observed in Egyptian fruit bats. When approaching an object, these bats do not center their sonar beams directly at the object, but rather point them off axis, so that the maximum slope of the acoustic profile intersects the object (Yovel et al., 2010)). Likewise, in this task mice do not gain an advantage by centering their sniffing directly downwind of the odor sources,

but rather perform best when they investigate the location of the steepest slope of the odor gradient, consistent with a serial-sniff gradient sensing strategy. Thus, our unsupervised computational analysis of airborne odor tracking supports the idea that sampling off axis can be an optimal strategy for localization across diverse sensory systems and species (Yovel et al., 2010).

Olfactory navigation can be either guided or gated by odor (Baker et al., 2018). Some organisms operate in a regime where diffusion forms smooth chemical gradients, in which classical chemotaxis strategies can be effective (Bargmann, 2006; Berg, 2000; Gomez-Marin and Louis, 2012; Lockery, 2011). In contrast, other organisms, such as flying insects, often operate in a highly turbulent regime where concentration gradients are not reliably informative (Crimaldi et al., 2002; Murlis et al., 1992; Riffell et al., 2008). By design, mice in our task operate in an intermediate regime, where turbulent odor plumes close to the ground form noisy gradients (Gire et al., 2017; Riffell et al., 2008). By varying the absolute concentration and the concentration difference between the two sides, we tested whether performance in this regime is guided or gated by odor. Because behavior varies with the gradient and not the absolute concentration (Fig. 2.2C-E), we have shown that mice are guided by gradient cues in this regime. Further, performance is higher when the mice sample both sides of the midline, suggesting that they sense the gradient by comparing sniff sequences across time.

Our naris occlusion experiments demonstrate that performance is statistically indistinguishable with naris occlusion, suggesting that stereo olfaction does not play a major role in our task. This finding contrasts with previous studies of olfactory navigation in a different regime: following a depositional odor trail. In these studies, stereo manipulations had small but significant effects on performance, and

led to changes in movement strategy (Jones and Urban, 2018; Khan et al., 2012). Importantly, a study of olfactory search in moles showed that stereo reversal did not affect navigation at a distance from the target, but reversed turning behavior in the target's immediate vicinity (Catania, 2013). These results suggest that stereo cues may be informative near a source, where gradients are steep, but that stereo cues play less of a role at a greater distance from the source where gradients are more shallow. In this more distant condition, serial sniff comparisons have been hypothesized as a potential sensory computation for odor gradient following (Catania, 2013). We propose that our task design, in which mice must commit to a side at a distance from the source, forces mice out of the stereo regime, and into the serial sniff comparison regime. Neurons sensitive to sniff-to-sniff odor concentration changes have been observed in the olfactory bulb of head-fixed mice (Parabucki et al., 2019), providing a potential physiological mechanism for this sensory computation.

On the other hand, physiological mechanisms revealed in head-fixed mice may not generalize to the freely-moving search condition. The external stimulus obtained by moving the nose through a noisy gradient differs dramatically from the square odor pulses delivered during head-fixed or odor-poke olfactory tasks. Further, the sniff statistics we observe in our mice are qualitatively faster than those reported in head-fixed mice under most conditions (Bolding and Franks, 2017; Shusterman et al., 2011; Wesson et al., 2009). One exception is that mice sniff fast in response to a novel odor (Wesson et al., 2009). Such fast stimulation impacts the responsiveness of olfactory sensory neurons (Esclassan et al., 2012; Ghatpande and Reisert, 2011; Verhagen et al., 2007). In addition to the temporal properties of odor transduction, short- and long-term synaptic and network plasticity mechanisms will influence the olfactory bulb's responses during fast sniffing (Beshel et al., 2007; D'iaz-Quesada

et al., 2018; Gupta et al., 2015; Jordan et al., 2018; Mandairon and Linster, 2009; Patterson et al., 2013; Zhou et al., 2020). Without tapping into the fast sniffing regime, the understanding we can gain from head-fixed studies in olfaction will be incomplete at best. In the future, it will be necessary to complement well-controlled reductionist behavioral paradigms with less-controlled, more natural paradigms like ours.

Mice execute a strategic behavioral program when searching, synchronizing fast sniffing with three-dimensional head movements at a tens of milliseconds timescale. It has long been known that rodents investigate their environment with active sniffing and whisking behaviors (Kepecs et al., 2006; Wachowiak, 2011; Welker, 1964). More recent work has established that under some conditions sniffing locks with whisking, nose twitches, and head movement on a cycle-by-cycle basis (Kurnikova et al., 2017; Moore et al., 2013; Ranade et al., 2013). Sniffing also synchronizes with brain oscillations not only in olfactory regions, but also in hippocampus, amygdala, and neocortex (Karalis and Sirota, 2018; Kay, 2005; Macrides et al., 1982; Vanderwolf, 1992; Yanovsky et al., 2014; Zelano et al., 2016). Respiratory central pattern generators may coordinate sampling movements to synchronize sensory dynamics across modalities with internal brain rhythms (Kleinfeld et al., 2014). Further, locomotor and facial movement, which are often synchronized to respiration, drive activity in numerous brain regions, including primary sensory areas (McGinley et al., 2015a; Musall et al., 2019; Niell and Stryker, 2010; Stringer et al., 2019b). Why are respiration and other movements correlated with activity in seemingly unrelated sensory regions? In the real world, sensory receptors operate in closed loop with movement (Ahissar and Assa, 2016; Gibson, 1966). Consequently, sensory systems must disambiguate self-induced stimulus dynamics from changes in the environment.

Further, active sampling movements can provide access to sensory information that is not otherwise available to a stationary observer (Gibson, 1962; Schroeder et al., 2010; Yarbus, 1967). Widespread movement-related signals may allow the brain to compensate for and capitalize on self-induced stimulus dynamics (Poulet and Hedwig, 2006; Sommer and Wurtz, 2008; Sperry, 1950; von Holst and Mittelstaedt, 1950; Webb, 2004). Our work advances understanding of how sensation and movement interact during active sensing.

Rigorously quantifying the behavior of freely-moving animals is more feasible than ever, thanks to recent developments in machine vision, deep learning, and probabilistic generative modeling (Datta et al., 2019; Gomez-Marin et al., 2014; Mathis and Mathis, 2020), as our work shows. In particular, the motifs we have defined provide a compact description of the behavior, while still capturing the idiosyncrasies of individual mice. Importantly, these motifs can be grouped into two larger-scale behavioral states that we putatively call "investigation" and "approach". Two-state search strategies are common across phylogeny (Bargmann, 2006; Berg, 2000; Kennedy and Marsh, 1974; Lockery, 2011; van Breugel and Dickinson, 2014; Vickers and Baker, 1994). In smaller organisms, state switches have provided a useful behavioral readout for understanding the neural mechanisms of odor-guided behavior (Bi and Sourjik, 2018; Larsch et al., 2015; Baker et al., 2018). Here, we have shown that where switches between investigation and approach occur in allocentric space can reveal the location of informative features in an olfactory scene. The transition points between 'investigation' and 'approach' serve as a principled template against which to compare neural activity. Our work thus establishes a framework for studying neural mechanisms of active sensing in an unrestrained mammal.

2.5 Methods and Materials

Custom-written task control, analysis, and visualization code is available at https://github.com/SmearLab/Freely-moving-olfactory-search (Findley et al, 2021).

2.5.1 Animals: Housing and Care

All experimental procedures were approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Oregon and are compliant with the National Institutes of Health Guide to the Care and Use of Laboratory Animals. C57BL/6J mice (2-14 months old) from the Terrestrial Animal Care Services (TeACS) at University of Oregon (19 male, 7 female) were used for behavioral experiments. Mice were housed individually in plastic cages with bedding and running wheels provided by TeACS. Mice were fed standard rodent chow ad libitum, and were water restricted, receiving a daily allotment (1mL - 1.5mL) of acidified or chlorinated water. Animal health was monitored daily, and mice were taken off water restriction if they met the 'sick animal' criteria of a custom IACUC-approved health assessment.

2.5.2 Behavioral Assay Design

2.5.2.1 Arena and Task Structure

Mice were trained to perform a two-choice behavioral task where they must locate an odor source for a water reward. This 15 x 25cm behavioral arena was largely custom designed in lab (all designs available upon request). The behavioral arena contains a custom designed and 3D-printed honeycomb wall through which continuous clean air is delivered to the arena and a latticed wall opposite to the honeycomb allowing airflow to exit the arena. Two odor tubes (Cole-Parmer Instrument Company, #06605-27) are embedded inside the honeycomb wall and consistently deliver either clean or odorized air. There are three nose pokes in the arena: one trial initiation poke and two reward pokes. The initiation poke is

embedded inside the latticed wall (where airflow exits) and is poked to initiate trials. The left and right reward pokes are embedded in the left and right arena walls against the honeycomb airflow delivery and are used for water reward delivery. Mice initiate odor release by entering the initiation poke. If the mouse locates the odor source successfully (by entering the quadrant of the arena containing the correct odor port), water (6μL - 8μL) is available at the corresponding nose poke. An intertrial interval of 4 seconds is then initiated. If the mouse goes to the incorrect side, water is not made available and they must wait an increased intertrial interval of 10 seconds.

2.5.2.2 Odor Delivery

Odor is delivered to the arena using two custom designed and built olfactometers. For a single olfactometer, air and nitrogen are run through separate mass flow controllers (MFCs) (Alicat Scientific, #MC-100SCCM-RD) that can deliver 1000mL/min and 100mL/min at full capacity, respectively. We can use these MFCs to control the percentage of total nitrogen flow (100mL/min) that runs through liquid odorant. Consequently, we can approximately control the amount of odor molecules in the resulting odorized air stream. Total flow is maintained at 1000mL/min (for example, if we are delivering 80mL/min of nitrogen, we will deliver 920mL/min of air). Nitrogen MFC output is directed through a manifold (NResearch Incorporated, #225T082) with embedded solenoids that direct flow to one of four possible vials. These vials contain odorant diluted in mineral oil or are empty. To odorize air, nitrogen is directed through a vial containing liquid odorant. The nitrogen aerosolizes the odorant and combines with airflow MFC output at the exit point of the manifold. If nitrogen is directed through an empty vial, un-odorized nitrogen will combine with airflow at the exit point. The resulting combined flow of air and nitrogen is then directed to a final valve (NResearch Incorporated, #SH360T042). Odorized air

continuously runs to exhaust until this final valve is switched on at which point clean air is directed to exhaust and odorized air to the behavioral assay. Therefore, we can control the percentage of odorized flow (using the MFCs), the presence or absence of odorized flow (using the vials and solenoids), and the flow of odorized air to the assay (using the final valve). There are two olfactometers (one for each odor port), which are calibrated weekly to match outputs using a photoionization detector (PID).

2.5.2.3 Video tracking

We use a Pointgrey Fly Capture Chameleon 3.0 camera (FLIR Integrated Imaging Solutions Inc, #CM3-U3-13Y3C) for video tracking. We capture frames at 80Hz at 1200x720 pixel resolution. All real-time tracking is executed using a custom Bonsai program. We isolate the mouse's centroid by gray scaling a black mouse on a white background and finding the center of the largest object. We track head position by applying red paint on the mouse's implant between the ears and thresholding the real-time HSV image to identify the center of the largest red shape. We can then identify nose position by calculating the extremes of the long axis of the mouse shape and isolating the extreme in closer proximity to the head point. These three points are sent to python at 80 Hz for real-time tracking in our assay. We use this real-time tracking to determine successful odor localization; if the mouse enters the quadrant of the arena that contains the correct odor port, it has answered correctly. Bonsai is an open source computer vision software available online (Lopes et al., 2015), and our custom code is available upon request.

For more rigorous behavioral analysis, we increased our tracking accuracy by using the open source tracking software Deeplabcut (Mathis et al., 2018; M. W. Mathis and Mathis, 2020). All Deeplabcut tracking occurred offline following experimentation.

2.5.2.4 Sniff Recordings

We record sniffing using intranasally implanted thermistors (TE Sensor Solutions, #GAG22K7MCD419; see Methods: Surgical Procedures). These thermistors are attached to pins (Assmann WSW Components, #A-MCK-80030) that can be connected to an overhead commutator (Adafruit, #736) and run through a custom-built amplifier (Texas Instruments, #TLV2460, amplifier circuit design available upon request).

2.5.2.5 Software

All behavioral experiments were run using custom code in Python, Bonsai, and Arduino. Behavioral boards designed at Janelia Research Farms that use Arduino software and hardware were used to control all hardware. Bonsai was used to execute real-time tracking of animals, and Python was used run the assay, communicate with Arduino and Bonsai, and save data during experiments. All programs used are open source, and all custom code is available upon request.

2.5.3 Surgical Procedures

For all surgical procedures, animals were anesthetized with 3% isoflurane; concentration of isoflurane was altered during surgery depending on response of the animal to anesthesia. Incision sites were numbed prior to incision with 20mg/mL lidocaine.

2.5.3.1 Thermistor Implantation

To measure respiration during behavior, thermistors were implanted between the nasal bone and inner nasal epithelium of mice. Following an incision along the midline, a small hole was drilled through the nasal bone to expose the underlying epithelium 2mm lateral of the midline in the nasal bone. The glass bead of the thermistor was then partially inserted into the cavity between the nasal bone and

the underlying epithelium. Correct implantation resulted in minimal damage the nasal epithelium. The connector pins were fixed upright against a 3cm headbar (custom designed and 3D printed) placed directly behind the animals' ears and the thermistor wire was fixed in place using cyanoacrylate. The headbar was secured against a small skull screw (Antrin Minature Specialties, #B002SG89OI) implanted above cerebellum. A second skull screw was placed at the juncture of the nasal bones to secure the anterior portion of the implant. All exposed skull and tissue were secured and sealed using cyanoacrylate. At the end of surgery, a small amount of fluorescent tempera red paint (Pro Art, #4435-2) was applied to the center of the headbar for tracking. Immediately following surgery, animals received 0.1mg/kg buprenorphine followed by 3 days of 0.03mg/kg ketoprofen. All but 9 mice were implanted prior to training. Mice that were implanted post-training were taken off water restriction at least 2 days prior to surgery and were not placed back on water restriction for at least one week following all analgesic administration.

2.5.3.2 Naris Occlusion

To test the necessity of stereo olfaction as a sampling strategy, we occluded the nostrils of C57BL/6J mice using 6-0 gauge surgical suture (SurgiPro, #MSUSP5698GMDL). Mice were given 0.03mg/kg ketoprofen and topical lidocaine on the nostril prior to induction. Suture was either pulled through the upper lip of the nostril and maxillary region to fully occlude the desired nostril or looped at the upper lip of the nostril for a sham stitch. Commercially available VetBond was applied to protect the suture knot. To ensure full occlusion, a small water droplet was placed on the occluded nostril. The absence of bubbles or seepage indicated a successful occlusion. Occlusion was re-tested in the same manner directly before each

experimental session. All stitches were removed within a week of application, and animals were stitched a total of 3 times per nostril.

2.5.4 Behavioral Training

All mice were trained to locate an odor source from one of two possible sources in the olfactory arena. Mice were removed from training and future experiments if they lost sniff signal or did not exceed 50 trials/perform above 60% correct in 15 sessions. The training process was divided into four primary stages.

2.5.4.1 Water Sampling

Mice were trained to alternate between the three pokes in the behavioral arena. Water (5 - 8µL) was made available at the nose pokes in the following order: initiation port, left reward port, initiation port, right reward port (repeat). Mice were trained in this task for 30 minutes per session until the mouse completed 70 iterations. This took mice 2-9 sessions. Data are only shown for 19 mice, because earlier iterations of the system did not save training data.

2.5.4.2 Odor Association

Mice were trained in the same sequence as water sampling. However, in odor association, water availability was removed from the initiation poke, and odor was released from whichever side water was available. Therefore, the mouse must initiate water availability by poking the initiation poke and then is further guided to the correct reward port by odor release. This task taught mice to initiate trials using the initiation poke and to associate odor with reward. However, in this step, odor is not required for reward acquisition as the task alternates left and right trials. Mice were trained in this task for 30 minutes per session until the mouse completed 70 iterations. This took mice 1-5 sessions. Data are only shown for 19 mice, because earlier iterations of the system did not save training data.

2.5.4.3 100:0 condition

Mice were given the same task as odor association, but with odor now randomly being released from the left or right odor port following an initiation poke. 10% of these trials were randomly 0:0 condition trials. To correctly answer, animals had to enter the quadrant of the arena (as tracked by the overhead camera) where odor was being released. If they answered correctly, water was made available at the reward port on the corresponding side. If they answered incorrectly, water was not made available and the mouse received an increased intertrial interval. Mice were trained in this task for 40 minutes per session until they exceeded 80% accuracy, which took 1-4 sessions (n = 26).

2.5.4.4 80:20 condition

When trials were initiated in this task, odor was released from both odor ports, but at differing concentrations. The animal had to enter the quadrant containing the odor port releasing the higher concentration. In this case, 80 means that the nitrogen MFC was set to 80mL/min on one olfactometer (see Methods: Behavioral Assay). Therefore, one odor port would release roughly 80% of the total possible odorant concentration. If one olfactometer was set to 80, then the other olfactometer would be set to 20 in this condition. 10% of these trials were randomly 0:0 condition. Mice were trained in this task for 40 minutes per session, taking 1-9 sessions to exceed 60% performance (n = 24).

2.5.5 Behavioral Experiments

2.5.5.1 Variable ΔC, Constant $|C|$

This experiment tested how performance and sampling strategy changes with task difficulty. In this experiment, mice performed a two-choice behavioral task where they located an odor source for a water reward at varying concentration differences

between the two ports. This experiment interleaved several possible conditions: 100:0 (all odor released from one port or the other), 80:20 (odor is released from both ports at different concentrations: 80% of the total possible airborne concentration and 20% of the total possible airborne concentration), and 60:40 (60% and 40%). Additionally, there was a control condition where all system settings were the same as the 80:20 condition, but nitrogen flow was directed through a clean vial so that the final flow was not odorized. 10% of the total number of trials were the 0:0 condition. Mice ran 40 minute experimental sessions and totaled 5-50 sessions (n = 19). These experiments were run with 1% liquid dilution of pinene.

2.5.5.2 Novel Odorant

This experiment tested how mice generalized our olfactory search task. A subset of mice were run with 1% liquid dilution of vanillin, which, unlike pinene, does not activate the trigeminal system (n = 3).

2.5.5.3 Constant ΔC, Variable $|C|$

This experiment tested if the animals use a thresholding strategy based on a fixed concentration threshold to solve the localization task. We ran this experiment using air dilution delivering the concentration ratios 90:30 and 30:10 interleaved randomly (n = 5). Mice ran 40 minute sessions and we analyzed data from the first session.

2.5.5.4 Naris Occlusion

This experiment tested the necessity of stereo olfaction in our localization task. Mice were run in the interleaved experiment (see above) initially. However, after observing no differences between concentration groups, we continued this experiment running mice in the 80:20 and 0:0 conditions only. Mice were run in one of five categories: left occlusion, left sham stitch, right occlusion, right sham stitch, and no stitch (see Methods: Surgical Procedures). Mice ran 40 minute experimental sessions

and totaled 5-30 sessions (n = 13). Stitches were always removed after 4 days. These experiments were run using 1% pinene dilutions.

2.5.6 Mapping the Olfactory Environment

We used a photoionization detector (PID, Aurora Scientific Inc, #201A) to capture real time odor concentration at a grid of 7x5 sampling locations in the assay. Using vials of 50% liquid dilution of pinene, we captured 15 two-second trials per sampling point. Odor maps were generated using the average concentration detected across all trials at each location. These maps were smoothed via interpolation across space. Discriminability maps in Figs. 1-supplement 2C and D were calculated with ROC analysis on the PID data (Green and Swets, 1966). To generate the distributions, each two second trial was divided into 25 ms chunks (approximately the mean inhalation duration during the task). For each space bin, the mean value of each 25 ms chunk was compiled into a distribution of odor concentration values for right and left trials (the different gradient conditions were pooled for this analysis). To map concentration gradient discriminability, 25 ms samples from each bin were assembled into a pseudosample, such that each sampling position had a concentration value. The gradient angle in each bin of this pseudosample was then calculated (imgradient function in MATLAB), and compiled into a distribution of angles for right and left trials (the different gradient conditions were pooled for this analysis). For both absolute concentration and gradient maps, the area under the ROC curve was calculated for each bin, scaled to between -1 and 1, and absolute valued, and these were assembled into a map and smoothed. Values are thresholded and shown at low bit depth (8 grayscale values) to facilitate perception of where the auROC values are highest.

2.5.7 Data Analysis

Analyses of odormaps, sniffing, DLC tracking, and motif sequences were performed in MATLAB. Inhalation and exhalation times were extracted by finding peaks and troughs in the temperature signal after smoothing with a 25ms moving window. Sniffs with duration less than the 5th percentile and greater than the 95th percentile were excluded from analysis. For alignment of movement with sniffing, tracking and motif sequences were shifted forward in time by 25ms (two frames), the temporal offset revealed by video calibration (Fig. 1).

2.5.7.1 Figure 1

Odormaps were visualized by smoothing the PID sampling grid with a gaussian, and colored using Cubehelix (Green, 2011).

2.5.7.2 Figure 2

Sessions where mice performed less than 60% correct on 80:20 (90:30 for Constant ΔC, Variable |C|), were less than 80 trials, or had any missing folders or files were excluded. Trials longer than 10s were excluded. Percent correct was calculated by dividing the correct trials by total trials in a single session and was averaged across all sessions, all mice. Trial duration was measured between nose poke initiation and reward poke and was averaged across all trials, all sessions, all mice. Tortuosity was measured by dividing the total path length by the shortest possible path length and was averaged across all trials, all sessions, all mice.

Statistical tests were performed in python using the scipy package (Jones et al., 2001). A binomial test was used to test statistical significance of above chance performance. Wilcoxon rank-sum tests were used for all group comparisons with pairwise comparisons for more than two groups. Two group comparisons were tested

using all trials pooled together and pairwise comparisons of three groups or more were tested across mice using individual mouse averages.

2.5.7.3 Figure 3

Occupancy and sniff rate colormaps were generated by down-sampling the tracking data to a 50x30 grid of bins (0.5 cm2). Occupancy colormaps are a 2D histogram of the nose position data. Sniff rate histograms were generated by dividing the sniff count in each position bin by the corresponding bin in the occupancy histogram. Both histograms were gaussian-smoothed, and colored using Cubehelix (Green, 2011). Grand means are shown in Fig. 2.3F,G, while individual mouse occupancy heatmaps are shown in Fig. 2.4. Maps were colored using Cubehelix (Green, 2011).

2.5.7.4 Figure 4 and 5

Nose speed, yaw velocity, and z velocity were calculated from the 3-point position time-series generated by Deeplabcut. For analysis, a 400ms window centered on each inhalation time was extracted from the kinematic timeseries. Colormaps in Fig. 2.4 show traces surrounding individual sniffs, while colormaps were generated using Bluewhitered (Childress, 2020). For within-trial sniffs, only those inhaled before the decision line were included. The inter-trial interval sniffs are taken from the time of reward port entry to the time of the first initiation port entry in the inter-trial interval. For cross-correlation and coherence analysis, we aligned the time-series of sniffing and kinematic parameters from the entire trial, or from the interval between reward and initiation port in the inter-trial interval. Tracking glitches were excluded by discarding trials or inter-trial intervals that contained frames with nose speed above a criterion value (100 pixels per frame).

2.5.7.5 Figure 6

Average motif shapes were generated from the mean positions of the nose, head, and body points from the first 8 frames of every instance of a given motif as determined by the AR-HMM. Decoding analysis is described in the following section.

2.5.7.6 Figure 7

The transition probability matrix was clustered by minimizing Euclidean distance between rows. For analyses separating investigation and approach sniffs, sniffs were defined as investigation or approach sniffs based on the state at the inhalation time. Colors for investigation and approach were selected from the Josef Albers painting, Tautonym, (B) (Albers, 1944).

2.5.7.7 Figure 8

Figures are generated by motif-onset triggered averages of inhalation times determined as described above. Fig. 2.8B,C are the grand mean of the motif onset-triggered average for each motif. Maps were colored using Cubehelix (Green, 2011). Sniff phase (relative time in sniff) was determined by dividing the motif onset latency from inhalation by the total duration (i.e., inhalation time to inhalation time) of each sniff. Modulation index was calculated as the difference between maximum and minimum instantaneous sniff rate, divided by the sum (max-minmax+min).

2.5.7.8 Figures 9, 9-supplement 1, 10, and 10-supplement-1

Investigation and approach occupancy maps were generated by down-sampling the tracking data to a 25x15 grid of bins (1 cm2). Occupancy maps are a 2D histogram of the nose position data, compiled separately for investigation and approach frames (see below for details of ARHMM analysis). In plots where the data are re-oriented with respect to the choice, the lateral axis of all right-choice trials has been flipped so that the trajectories always end on the left side (top side in the displayed occupancy

maps). Both histograms were normalized to the total occupancy in a given bin (i.e., investigation + approach), gaussian-smoothed, and merged and colored using a scheme adapted from fluorescence microscopy. Grand means (n = 9) are shown in Figs. 2.9A,B, 10A, and S16, while individual mouse mean occupancy maps are shown in Fig. 2.25. Investigation approach index (I.A.I) is calculated as the difference between investigation and approach occupancies over their sum, for a given bin. I.A.I is taken from histograms which are the projection of the 2D maps onto the longitudinal or lateral axes. The "transition zone" is defined as the region between 5-10 cm from the longitudinal axis origin (i.e., the initiation port), and lateral axis histograms are taken from within this region in Figs. 10D-F and S16. Correct-incorrect occupancy and index differences are grand mean of the individual mouse differences in Figs. 2.10C,G-I and S16. These differences are evaluated statistically against a null distribution generated by scrambling the correct and incorrect trial labels 1000 times and re-running this analysis. Importantly, these shuffles are performed within mice before taking the post-shuffle grand means, so that these null distributions incorporate both within-mouse and across-mouse variability.

2.5.7.9 Sniff synchronization

Sniff cycles were compared with kinematics to determine the extent of movement modulation at individual sampling points. Individual sniffs were cross-correlated with each kinematic signal (i.e. nose speed) at -200ms from inhalation onset to +200ms from inhalation onset. To further determine synchrony between the two signals, we measured the coherence of signal oscillation between sniff signals and individual kinematic measurements at -200ms from inhalation onset to +200ms from inhalation onset.

2.5.8 Auto-Regressive Hidden Markov Model

Let $\mathbf{x}_t$ denote the 6-dimensional vector of nose-head-body coordinates at video-frame t (sampled at 80 Hz), with components $(x_{\text{nose}}, y_{\text{nose}}, x_{\text{head}}, y_{\text{head}}, x_{\text{body}}, y_{\text{body}})$. We fit an auto-regressive hidden Markov model (AR-HMM) to mouse trajectory data, $\mathbf{x}_t^{(i)}$, across trials (indexed by (i)) from 13 out of 15 mice (two mice were excluded *a priori* due to low task performance). These mice performed olfactory search under the following experimental conditions: variable ΔC, constant $/C/$ (9 mice); naris occlusion (7 mice); and constant ΔC, variable $/C/$ experiments (5 mice).

2.5.8.1 The generative view

Viewed as a generative model (that generates simulated data), the AR-HMM has two "layers": a layer of hidden discrete states (corresponding to discrete movement motifs), and an observed layer which is the continuous trajectory $\mathbf{x}_t$. We denote the temporal sequence of discrete states by $\mathbf{z}_t$. In each time-step, $\mathbf{z}_t \in \{1, 2, \ldots, S\}$, *i.e.* it is one of an S number of states, or movement motifs. The discrete hidden states evolve in time according to a Markov chain: going from time-step t to $t+1$, the discrete state may change to another state according to a transition probability matrix $\pi_{\mathbf{z}_1,\mathbf{z}_2}$, which denotes the conditional probability of switching to $\mathbf{z}_2$ having started in $\mathbf{z}_1$. The probability distribution over the initial state, $\mathbf{z}_{t=1}$, of the Markov chain at the start of each trial was taken to be the uniform distribution.

Now suppose for time steps t_1 to t_2 (inclusive) the discrete layer remained in state z. The continuous or autoregressive part of the model dictates that, over this time interval, the continuous trajectory, $\mathbf{x}_t$, evolves according to a linear autoregressive (AR) process. The parameters of this AR process can be different in different states or motifs, z. In other words, $\mathbf{x}_t$ is governed by

$$\mathbf{x}_t = A_{\mathbf{z}}\mathbf{x}_{t-1} + \mathbf{b}_z + \boldsymbol{\varepsilon}_t \qquad t_1 \leq t \leq t_2 \tag{2.1}$$

where $A_{\mathbf{z}}$ is a 6×6 matrix and $\mathbf{b}_z$ is a 6×1 vector, and the noise vector $\boldsymbol{\varepsilon}_t$ is sampled from the multi-variate zero-mean gaussian distribution $\mathcal{N}(0, Q_z)$ where Q_z is a 6×6 noise covariance matrix. Moreover, the parameters $A_{\mathbf{z}}$, $\mathbf{b}_z$, and Q_z depend on the discrete state $\mathbf{z}$, and in general are different in different discrete states. The simple stochastic linear dynamics described by Eq. (2.1) can describe simple motions of the mouse, such as turning left/right, dashing towards a certain direction, freezing (when $A_{\mathbf{z}}$ is the identity matrix and $\mathbf{b}_z$ is zero), etc. The switches between these simple behaviors allow the model to generate complex trajectories.

The AR-HMM model is an example of a model with latent variables, which in this case are the discrete state sequence $\mathbf{z}_t^{(i)}$ in each trial. The model, as a whole, is specified by the set of parameters $(\pi, \{A_z\}, \{\mathbf{b}_z\}, \{Q_z\})$, which we will denote by $\boldsymbol{\theta}$. For a d-dimensional trajectory ($d = 6$ here) and S states, comprises $S(S-1) + S(d^2 + d + d(d+1)/2) = S(S - 1 + 3d(d+1)/2)$ parameters.

2.5.8.2 Model fits

Models with latent variables are often fit using the expectation-maximization (EM) algorithm which maximizes the likelihood of the model in terms of the parameters $\boldsymbol{\theta} \equiv (\pi, \{A_z\}, \{\mathbf{b}_z\}, \{Q_z\})$ for a given set of observed data $\{\mathbf{x}_t^{(i)}\}$. In this work, we did not use the EM algorithm, but adopted a fully Bayesian approach in which both the hidden variables and the model parameters were inferred by drawing samples from their posterior distribution (**?**). The posterior distribution combines the model likelihood and Bayesian priors imposed on its parameters, according to Bayes' rule. If we denote the joint-likelihood of observed trajectories, $\{\mathbf{x}_t^{(i)}\}$, and the latent variables, $\{z_t^{(i)}\}$, by $P(\{\mathbf{x}_t^{(i)}, z_t^{(i)}\}|\boldsymbol{\theta})$ and the prior distribution over model parameters

by $P(\boldsymbol{\theta})$, then up to normalization, the joint posterior distribution of latent variables and model parameters is given by

$$P(\{z_t^{(i)}\}, \boldsymbol{\theta} | \{\mathbf{x}_t^{(i)}\}) \propto P(\{(\mathbf{x}_t^{(i)}, z_t^{(i)})\} | \boldsymbol{\theta}) P(\boldsymbol{\theta}) \tag{2.2}$$

For the AR-HMM model, the (logarithm of the) joint log-likelihood is given by

$$\log P(\{\mathbf{x}_t^{(i)}, z_t^{(i)}\} | \boldsymbol{\theta}) = \sum_i \sum_{t=2}^{T_i} \left[\log \pi_{z_{t-1}^{(i)}, z_t^{(i)}} + \log \mathcal{N}(\mathbf{x}_t^{(i)} | A_{z_t^{(i)}} \mathbf{x}_{t-1}^{(i)} + \mathbf{b}_{z_t^{(i)}}, Q_{z_t^{(i)}}) \right] \tag{2.3}$$

where T_i is the length of trial i, and we use the notation $\mathcal{N}(\mathbf{x} | \boldsymbol{\mu}, Q) = e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}} Q^{-1} (\mathbf{x}-\boldsymbol{\mu})} / \sqrt{|2\pi Q|}$ to denote the density at point $\mathbf{x}$ of a multivariate gaussian with mean vector $\boldsymbol{\mu}$ and covariance matrix Q.

We imposed loose conjugate priors on the model parameters, which were factorized over the parameters of the AR process, $(\{A_z\}, \{\mathbf{b}_z\}, \{Q_z\})$, in different discrete states z, and the different rows of the Markov transition matrix, π. On the rows of π, we imposed Dirichlet distribution priors with uniform distribution means, and concentration hyperparameter α, which was set to 4. We imposed matrix normal inverse-Wishart priors on the AR parameters, independently for different discrete states. Under this prior, the noise covariance Q_z has an inverse Wishart distribution with a "scale matrix" hyperparameter, which was set to the $d \times d$ $(= 6 \times 6)$ identity matrix, and a "degrees-of-freedom" scalar hyperparameter set to $d + 2 = 8$. Conditional on Q_z, the remaining AR parameters, $(A_z, \mathbf{b}_z)$, have a joint multivariate normal distribution under the prior, which can be specified by the prior mean and joint prior covariance matrix of A_z and b_z. The prior means of A_z and $\mathbf{b}_z$ were set to the $d \times d$ identity matrix and the d-dimensional zero vector, respectively, while the

prior covariance matrix of the concatenation $(A_z, \mathbf{b}_z)$ was given by the tensor product of Q_z and the $(d+1) \times (d+1)$ $(= 7 \times 7)$ identity matrix (equivalently, under this prior, b_z and different columns of A_z are independent and uncorrelated, while each of these column vectors has a prior covariance equal to the (prior) AR noise covariance, Q_z).

Bayesian model inference was carried out by sampling from (instead of maximizing) the joint posterior distribution of the model parameters and latent state variables conditioned on the observed trajectory data, Eq. (2.2). We did this by Gibbs sampling (an example of Markov Chain Monte Carlo; not to be confused with the Markov Chain in the AR-HMM model), which works in a manner conceptually similar to the EM algorithm: it switches between sampling $z_t^{(i)}$ in all trials, conditioned on previously sampled parameters, and then sampling the parameters $\boldsymbol{\theta}$ given the previous sample of $\{z_t^{(i)}\}$. To carry out this model inference procedure, we used the Python package developed by M. J. Johnson and colleagues, publicly available at `https://github.com/mattjj/pyhsmm`.

We ran the Gibbs sampler for 300 iterations, and burned the first 200 samples, retaining 100. We used the remaining samples to obtain the posterior probabilities of hidden discrete states at each time step of each trial (by calculating the frequency of different state in that time step and trial, across the retained Gibbs samples), as well as posterior expectation of the model parameters (by calculating their averages over the retained Gibbs samples). We refer to the AR-HMM with parameters given by these latter posterior expectations as the "fit model".

2.5.8.3 Model selection

We fit AR-HMM's with different numbers of states (motifs), S, to mouse trajectory data pooled across animals. To evaluate the statistical goodness-of-fit

of these fit model and select the best S (the number of states or motifs), we evaluated the log-likelihoods of fit models on trajectory data from a held-out set of trials, not used for model fitting. The corresponding plot of log-likelihoods is shown in Figure 6F. As seen, the log-likelihood keeps increasing with S, up to S = 100. This shows that, up to at least S = 100, additional motifs do have utility in capturing more variability in mouse trajectories. These variabilities may include differences in movement across mice, as well as movement variations in the same mouse but across different trials or different instances of the same movement; for example, a clockwise head turn executed with different speeds in different instances or trials. In the AR-HMM model, the AR observation distribution of a given Markov state corresponds to a very simple (linear) dynamical system which cannot capture many natural and continuous variations in movement, such as changes in movement speed. Nevertheless AR-HMM models with higher S can capture such variations with more precision, by specializing different discrete Markov states, with different AR distributions, to movement motifs of different mice, or, for example, to capture different speeds of the same qualitative movement motif.

The goal for this modelling was to give a compact description of recurring movement features across animals and conditions, suitable for visualization and alignment. For these purposes, the goodness-of-fit did not provide a suitable criterion, because the log-likelihood plots did not peak or plateau even at very large numbers of states. Guided by visual inspection, we thus chose the model with S = 16, for the main figures (Fig. 2.6-8). Although this was a somewhat arbitrary choice, we show that the findings in Figures 6, 7, and 8 do not depend on the choice of S – models with S = 6, 10, or 20 gave equivalent results (Fig. 2.19-2.21).

2.5.8.4 MAP sequences

The Gibbs sampling algorithm which we used for model inference yields (timewise marginal) Maximum A Posteriori (MAP) estimates of the latent variables $\{\mathbf{z}_t^{(i)}\}$, as follows. Using the Gibbs samples for the latent variables we can estimate the posterior probability of the mouse being in any of the S states in any given time-step of a given trial. We made MAP sequences by picking, at any time step and trial, the state with the highest posterior probability. The inferred MAP motifs tended to have high posterior probability, which exceeded 0.8 in 66.2% of all time-steps across the 17195 trials in the modeled dataset.

2.5.9 Decoding analysis

We decoded experimental conditions and animal identities from single-trial MAP motif sequences inferred using the AR-HMM model. Specifically, we trained multi-class decoders with linear decision boundaries (Linear Discriminant Analysis) to decode the above categorical variables from the single-trial empirical state transition probability matrices derived from the MAP sequence of each trial. If $\hat{z}_t^{(i)}$ is the motif MAP sequence for trial i, the empirical transition probability, $\hat{\pi}_{a,b}^{(i)}$, from state a to state b ($a, b \in \{1, \ldots, K\}$), for that trial was calculated by:

$$\hat{\pi}_{a,b}^{(i)} \equiv \frac{n_{a,b}^{(i)}}{\sum_{c=1}^{K} n_{a,c}^{(i)}} \tag{2.4}$$

$$n_{a,b}^{(i)} \equiv \sum_{t=1}^{T^{(i)}-1} \mathbb{I}(\hat{z}_t = a)\mathbb{I}(\hat{z}_{t+1} = b) \tag{2.5}$$

where $T^{(i)}$ is the length of trial i, and $\mathbb{I}(\cdot)$ is an indicator function, returning 1 or 0 when its argument is true or false, respectively.

We used the decoder to either classify experimental condition or mouse identity, in different trials (Fig. **??**D,E). For decoders trained to classify the trials' experimental condition, we used pooled data across mice. For decoders trained to classify mouse identity, we only used data from the 80-20 odor condition. Data was split into training and test dataset in a stratified 5-fold cross-validation manner, ensuring equal proportions of trials of different types in both datasets. The trial type was the combination of left vs right decision, experimental condition, and mouse identity.

To calculate the statistical significance of decoding accuracies, we performed an iterative shuffle procedure on each fold of the cross-validation. In each shuffle, the training labels which the classifer was trained to decode were shuffled randomly across trials of the training set, and the classifer's accuracy was evaluated on the unshuffled test data-set. This shuffle was performed 100 times to create a shuffle distribution of decoding accuracies for each fold of the cross-validation. From these distributions we calculated the z-score of decoding accuracy for each class in each cross-validation fold. These z-scores were then averaged across the folds of cross-validation and used to calculate the overall p-value of the decoding accuracy obtained on the original data.

2.6 Bridge to Chapter III

In this chapter, we investigated the behavioral strategies mice employ during olfactory search. Studying behavior as a stand-alone project is a necessary first step in describing what an animal is actually doing moment to moment before recording from sensory neural populations. We developed the methods necessary to accomplish this for a freely moving paradigm like we have here, but is also applicable in a more restricted head-fixed set-up. This method segmented the mouse's behavior into

identifiable and reoccurring behavioral states, which we can use to align to periods of interest in neural activity. In Chapter III, we do something similar with locomotion. We analyze neural activity during periods when the animal is at rest and when the animal is moving and found an acceleration of visual processing in the neural data. While this chapter looked at active sensation in the olfactory system, Chapter III looks at active sensation in the visual system.

2.7 Supplemental Figures

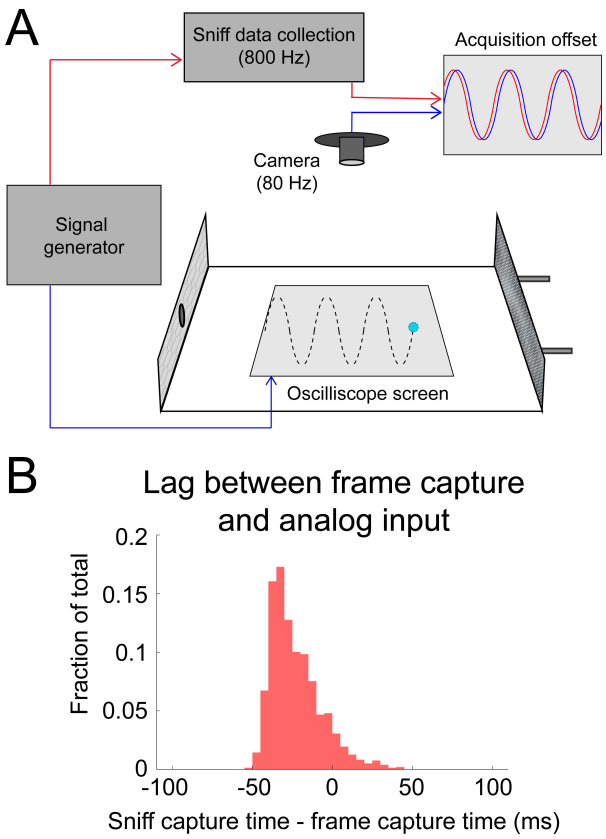

Figure 2.11. Calibrating alignment of video frames with sniff signal. **A)** Sinusoidal signals (5, 8, 10, and 15 Hz) were simultaneously sent to the analog input channel (used to capture sniffing) and to a phosphor-display oscilloscope (Tektronix). The display of the oscilloscope was reflected by mirrors to allow it to be video-captured inside the behavioral arena. **B)** The timing relationship is given by the lag between peaks in the analog input channel and the vertical peaks in the position of the oscilloscope trace. Analog input led video frames by 23.5 ± 15.7 ms (mean ± sd; approximately two frames at 80 frames/second).

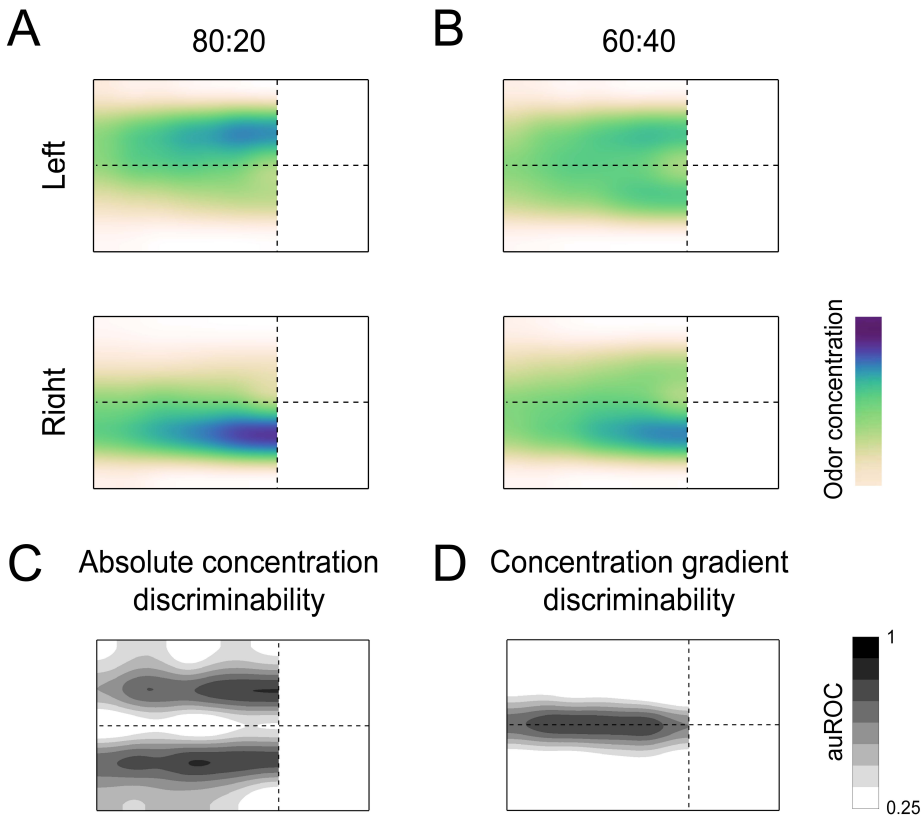

Figure 2.12. Characterizing the odor stimulus conditions. **A)** Color maps of average odor concentration across 15 two-second trials captured by a 7x5 grid of sequential photoionization detector recordings. - Each row represents trial type (left correct or right correct). 80:20 odor condition (see Methods: Behavioral Training: 80:20). **B)** Same as A, for the 60:40 odor condition (see Methods: Interleaved: 60:40). **C)** Absolute concentration discriminability map based on PID recordings (see methods). Darker shades indicate regions where absolute concentrations are most discriminable according to ROC analysis. Essentially, these regions downwind of the odor ports have the largest differences in absolute concentration between left and right trials. **D)** Concentration gradient discriminability map based on PID recordings. Darker shades indicate regions where odor concentration gradient angles are most discriminable according to ROC analysis. Essentially, this region around the lateral midline of the arena has the largest differences in concentration gradient between left and right trials.

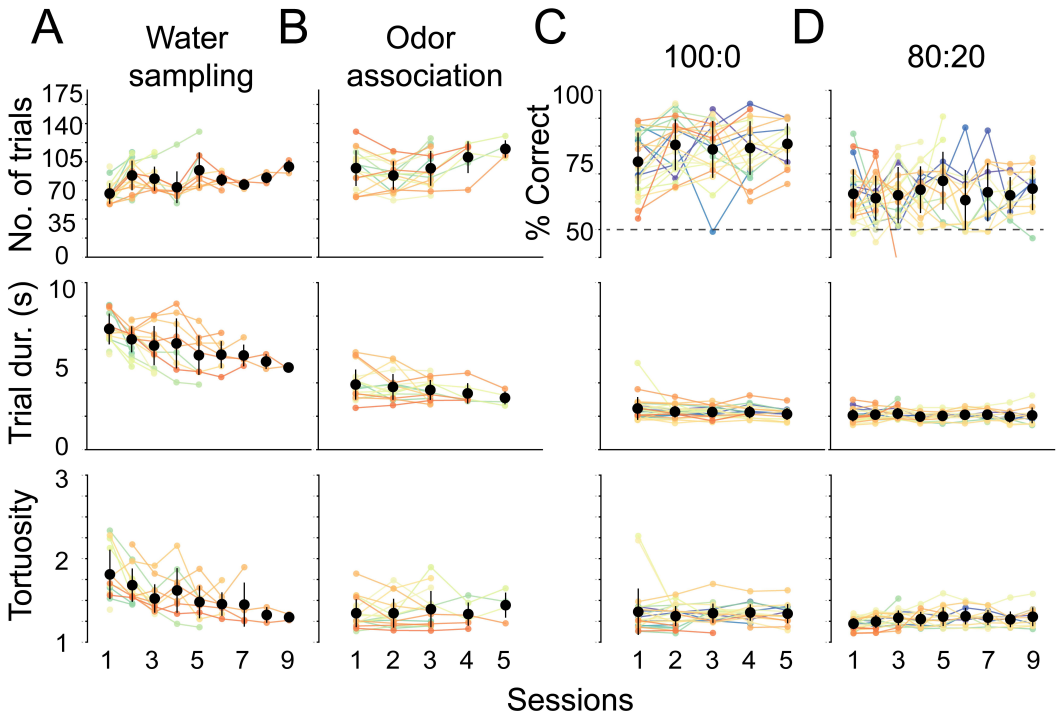

Figure 2.13. Session statistics across trainer sessions. Individual mice are depicted by colored lines, average across mice are black points, and whiskers are ±1 standard deviation from the average across mice. Mice 2054-2062 did not have trainer 1 and 2 recorded (this accounts for increasing n) and mice were commonly removed from the experiment if they lost sniff signal (this accounts for the reducing n). Above. Number of trials performed or percent of correct trials. Middle. Average trial duration. Below. Average trial path tortuosity (total path length/shortest possible path length). **A)** Session statistics for the first trainer, water sampling (n = 19). **B)** Session statistics for the second trainer, odor association (n = 19). **C)** Session statistics for the third trainer, 100:0 or olfactory search (n = 26). Mice perform above 70% in first session. **D)** Session statistics for final training step, 80:20, that preceded experiments shown in Figure 2 (n = 24).

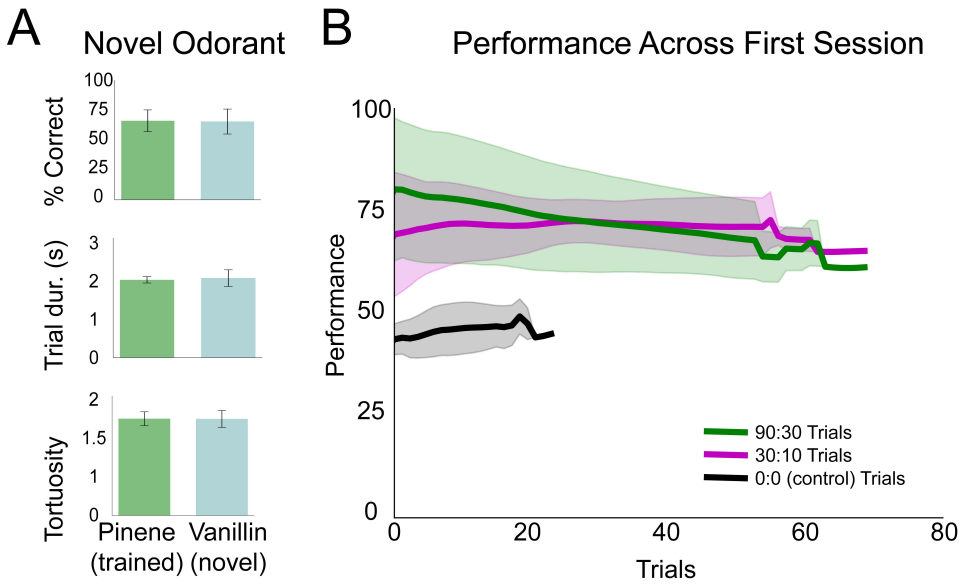

Figure 2.14. Mice generalize search task to novel odorants and variable |C| session. **A)** Performance, trial duration, and trial tortuosity (total path length/shortest possible path length) for the last session of pinene training in 80:20 and the first session of vanillin in 80:20 across mice (n = 3). **B)** Grouped by stimulus condition (90:30, 30:10, 0:0), each line represents the rolling average across mice (window = 10) for the first session (n = 5). Shaded regions represent ±1 standard deviation.

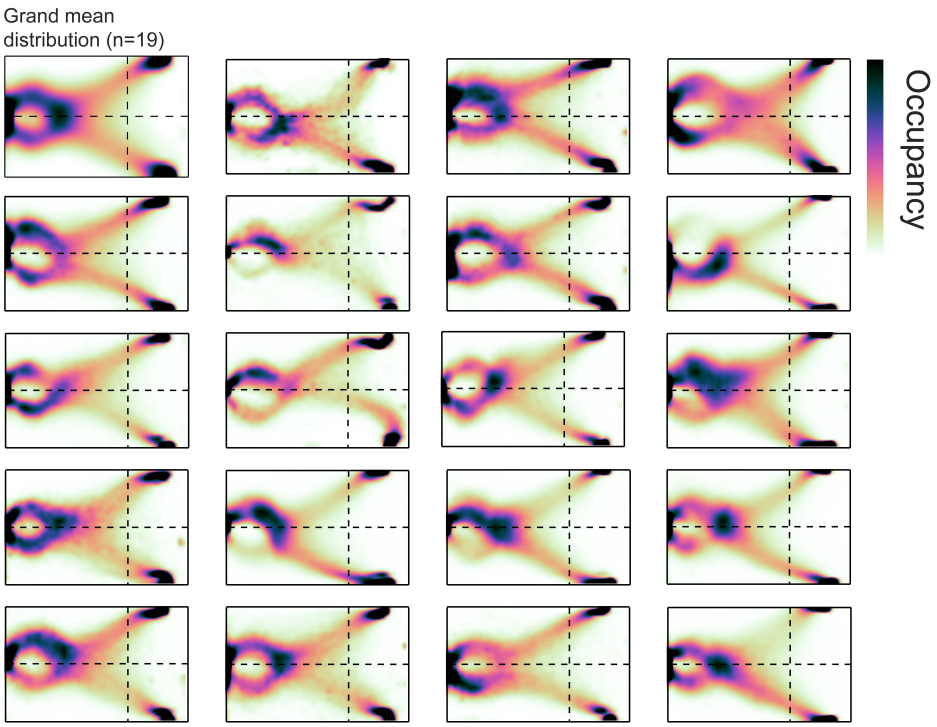

Figure 2.15. Idiosyncratic occupancy distributions across individual mice. Two-dimensional histogram of occupancy (fraction of frames spent in each 0.5cm^2 bin).

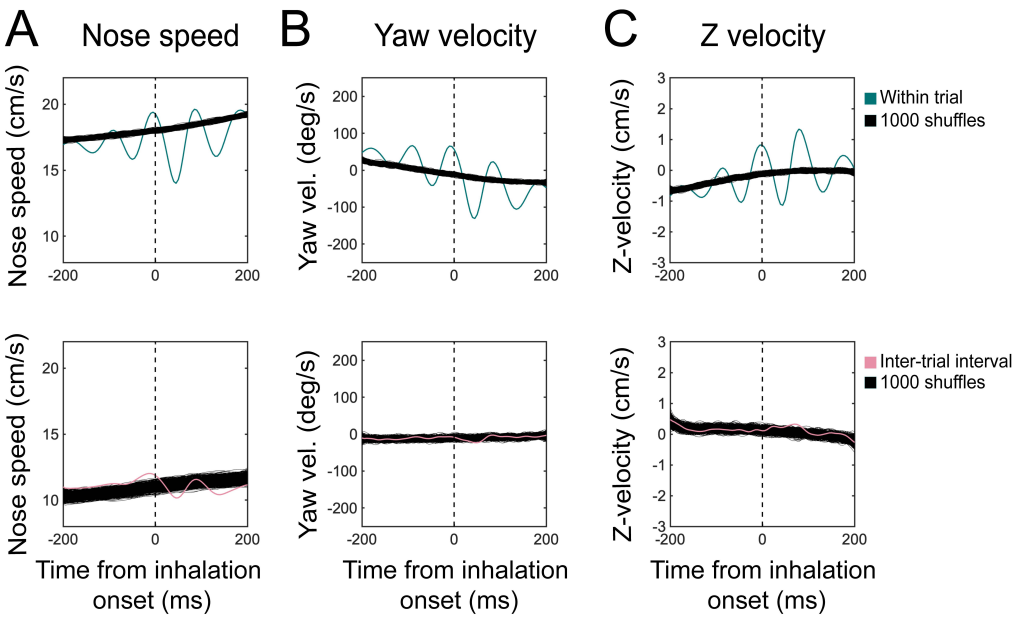

Figure 2.16. Sniff synchronization shuffle test. Sniff-aligned grand mean (n = 11 mice) of **A)** nose speed, **B)** yaw velocity, and **C)** Z-velocity for within-trial (Top) and inter-trial interval (Bottom) sniffs, overlaid on 1000 iterations of trial-shuffled grand means. Thin black lines represent individual iterations, all of which are shown.

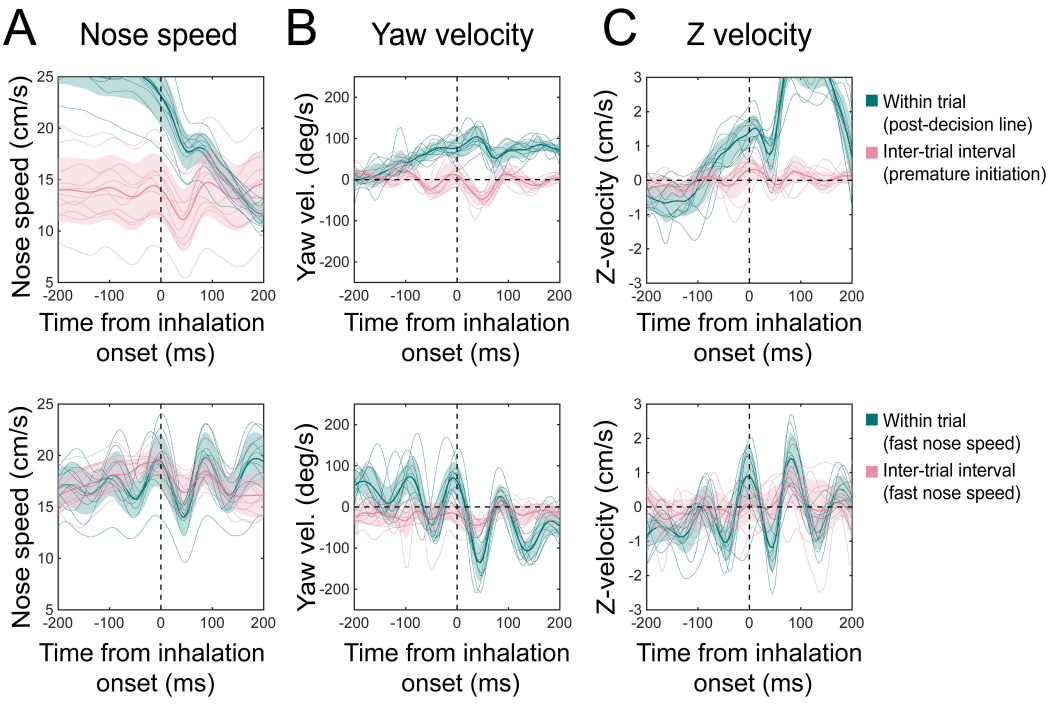

Figure 2.17. Kinematic rhythms for premature initiations during the intertrial interval and between decision line and reward port during trials. Sniff-aligned average of **A)** nose speed, **B)** yaw velocity, and **C)** Z-velocity. Thin lines represent individual mice (n = 11), bolded lines and shaded regions represent the grand mean ± standard deviation. Top. Green: within-trial sniffs from the time between crossing the decision line and entering the reward port. Pink: inter-trial interval sniffs from the time between the first premature trial initiation attempt and the successful initiation of the next trial. Bottom. Green: within-trial sniffs at nose speeds above the threshold 15 cm/s. Pink: inter-trial interval sniffs at nose speeds above the threshold 15 cm/s.

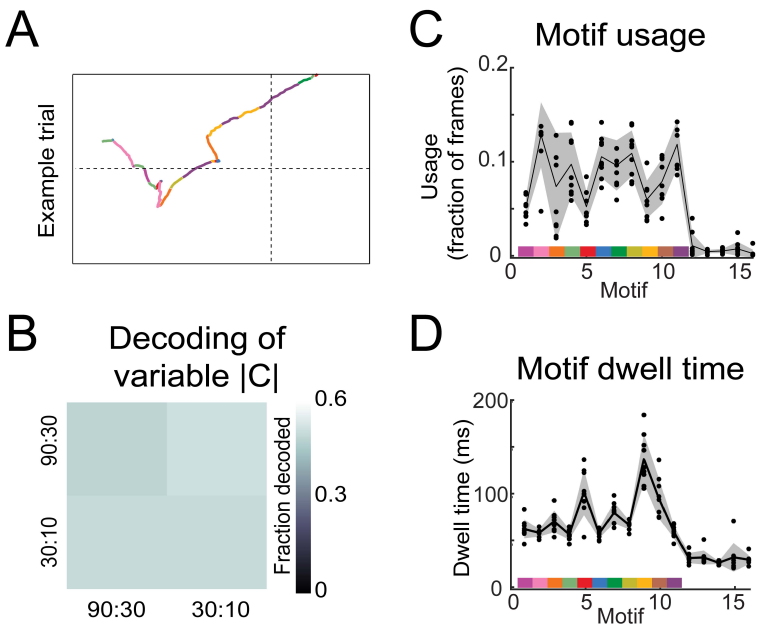

Figure 2.18. Motif statistics and examples and linear decoder results for 80:20 experiments.**A)** Example nose trace across a single trial colored by motif identity. **B)** Linear decoder (Fig. 3.6; Methods: Linear decoding section) results for Variable |C| experiments (n = 5). **C)** Fraction of motif usage across all mice (n = 8) for the model with S = 16. Black points are individual mice, black line is average across mice, and shaded region is ±1 standard deviation. Colors on x axis represent motifs used in analysis (Fig. 3.6) and y axis are fractions of frames that motif occupies. **D)** The average dwell time of each motif across all mice (n = 8) for the model with S = 16. Black points are individual mice, black line is average across mice, and shaded region is ±1 standard deviation. Colors on x axis represent motifs used in analysis (Fig. 3.6) and y axis are fractions of frames that motif occupies.

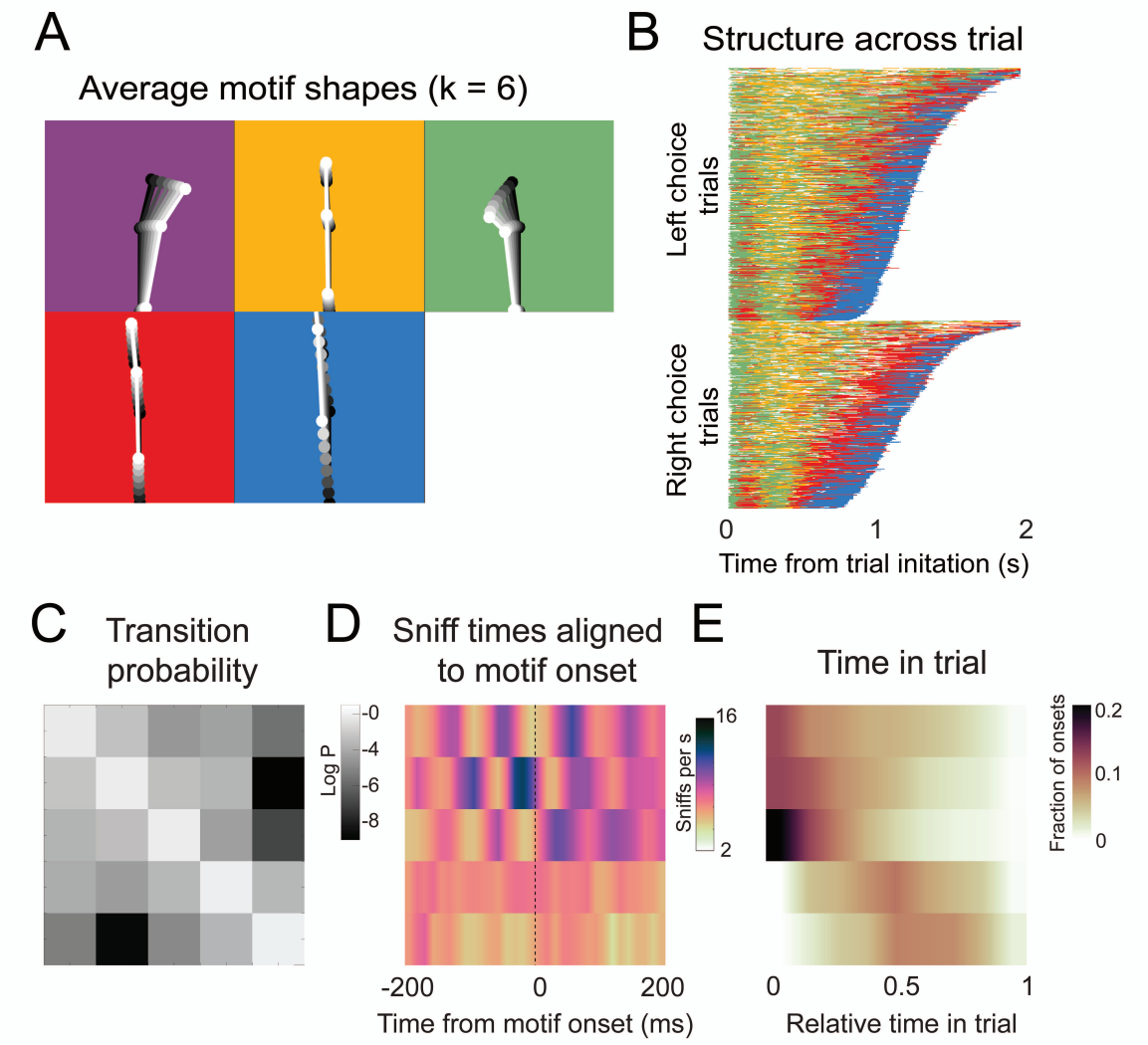

Figure 2.19. Motif shapes, sequences, transition matrices, and sniff synchronization for an AR-HMM capped at a maximum of 6 states

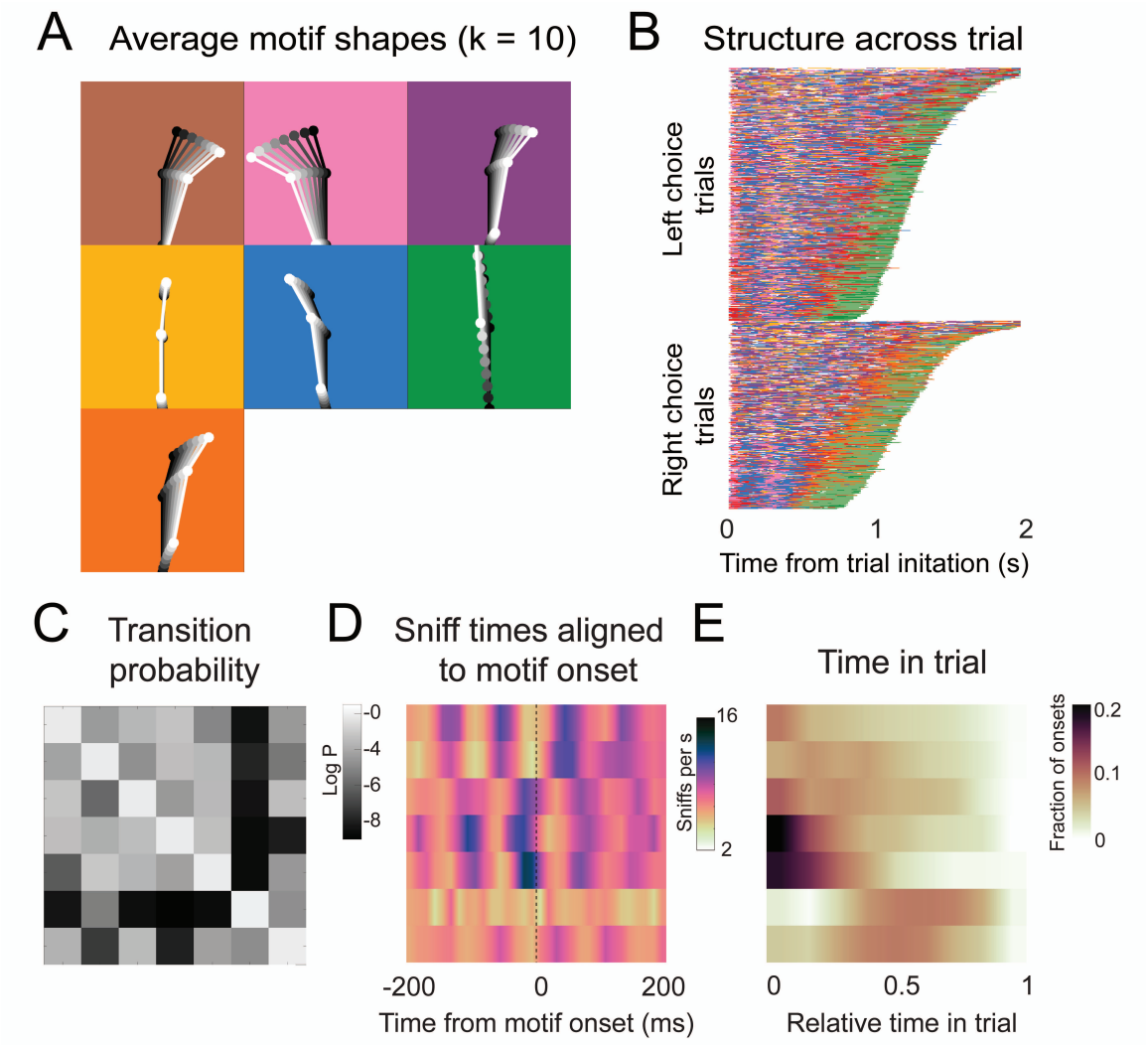

Figure 2.20. Motif shapes, sequences, transition matrices, and sniff synchronization for an AR-HMM capped at a maximum of 10 states

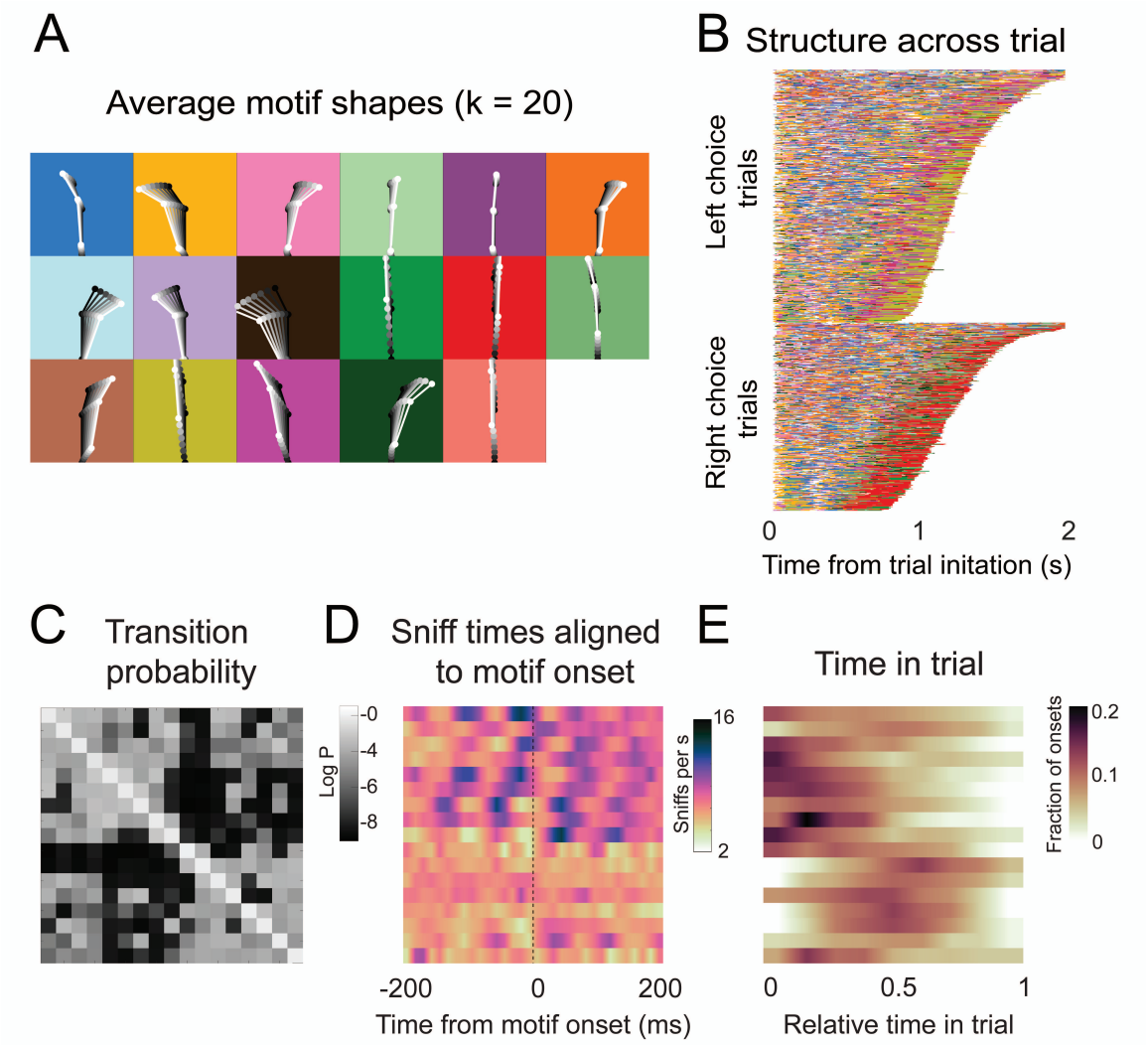

Figure 2.21. Motif shapes, sequences, transition matrices, and sniff synchronization for an AR-HMM capped at a maximum of 20 states

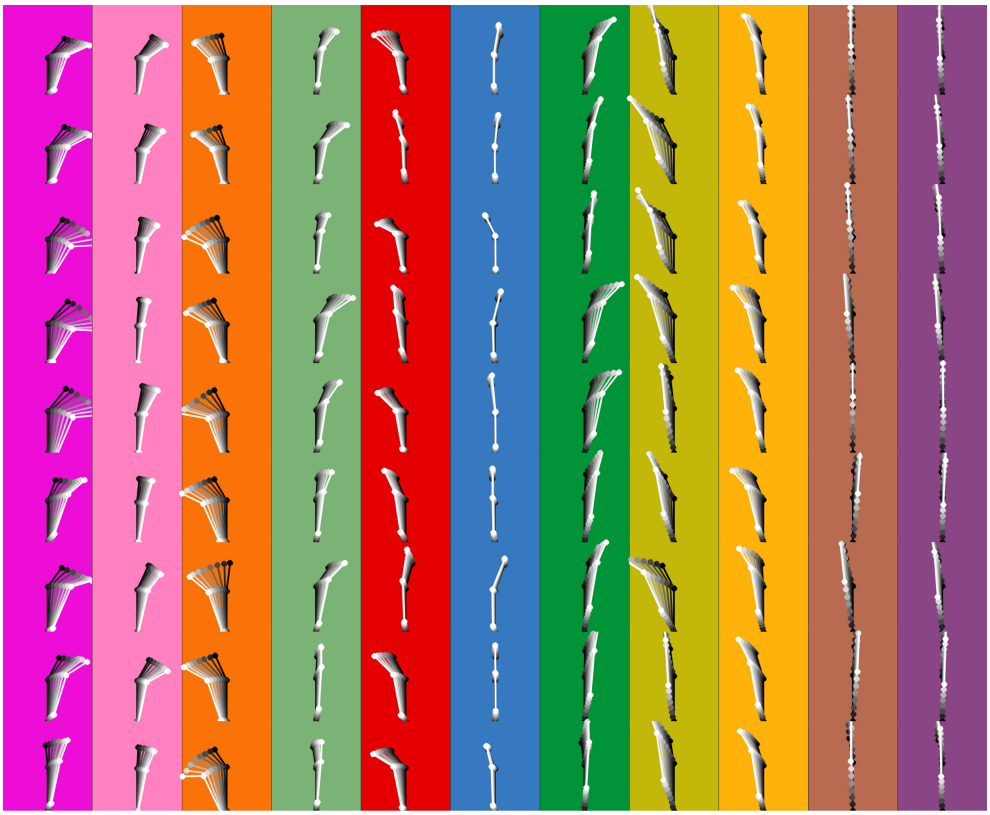

Figure 2.22. Average first 8 frames of the 11 commonly used motifs across individuals. Dots and lines show the average time course of posture for 8 frames of each of the 11 motifs. All instances of each motif are translated and rotated so that the head is centered and the head-body axis is oriented upward in the first frame. Subsequent frames of each instance are translated and rotated the same as the first frame. Time is indicated by color (dark to light). Each color/column represents a single motif and each row an individual mouse.

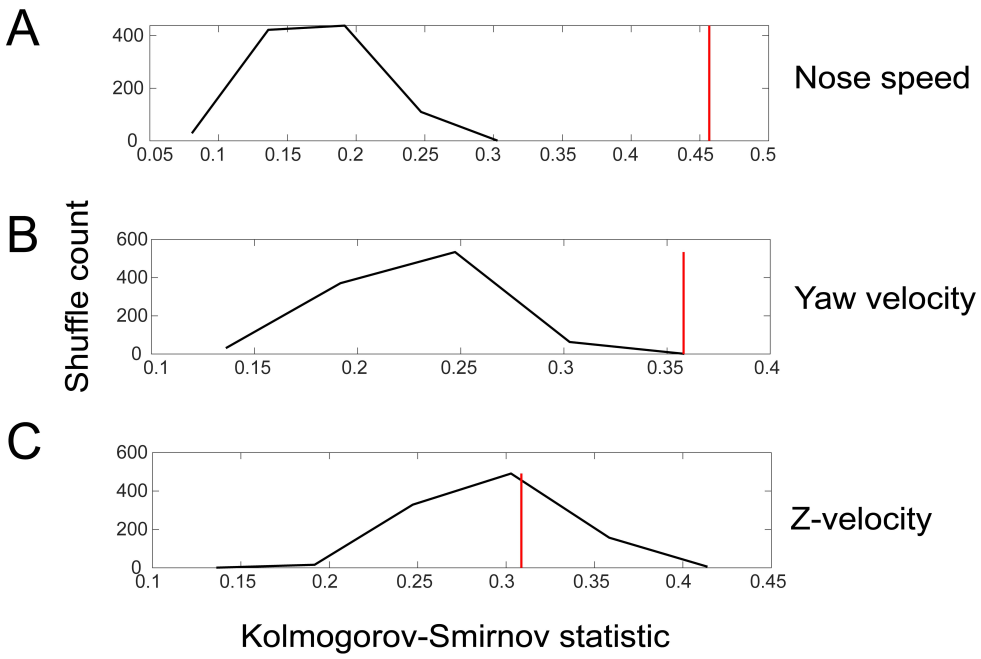

Figure 2.23. Shuffle test for the difference in sniff synchronization between investigation and approach motifs for movement parameters. We quantified the difference by calculating the Kolmogorov-Smirnov statistic for the comparison between the sniff triggered averages in the two states, first for real data, and then for 1000 iterations of trial-shuffled data. Red shows the value for the real data, while the black histogram plots the distribution of Kolmogorov-Smirnov statistic for the 1000 iterations.

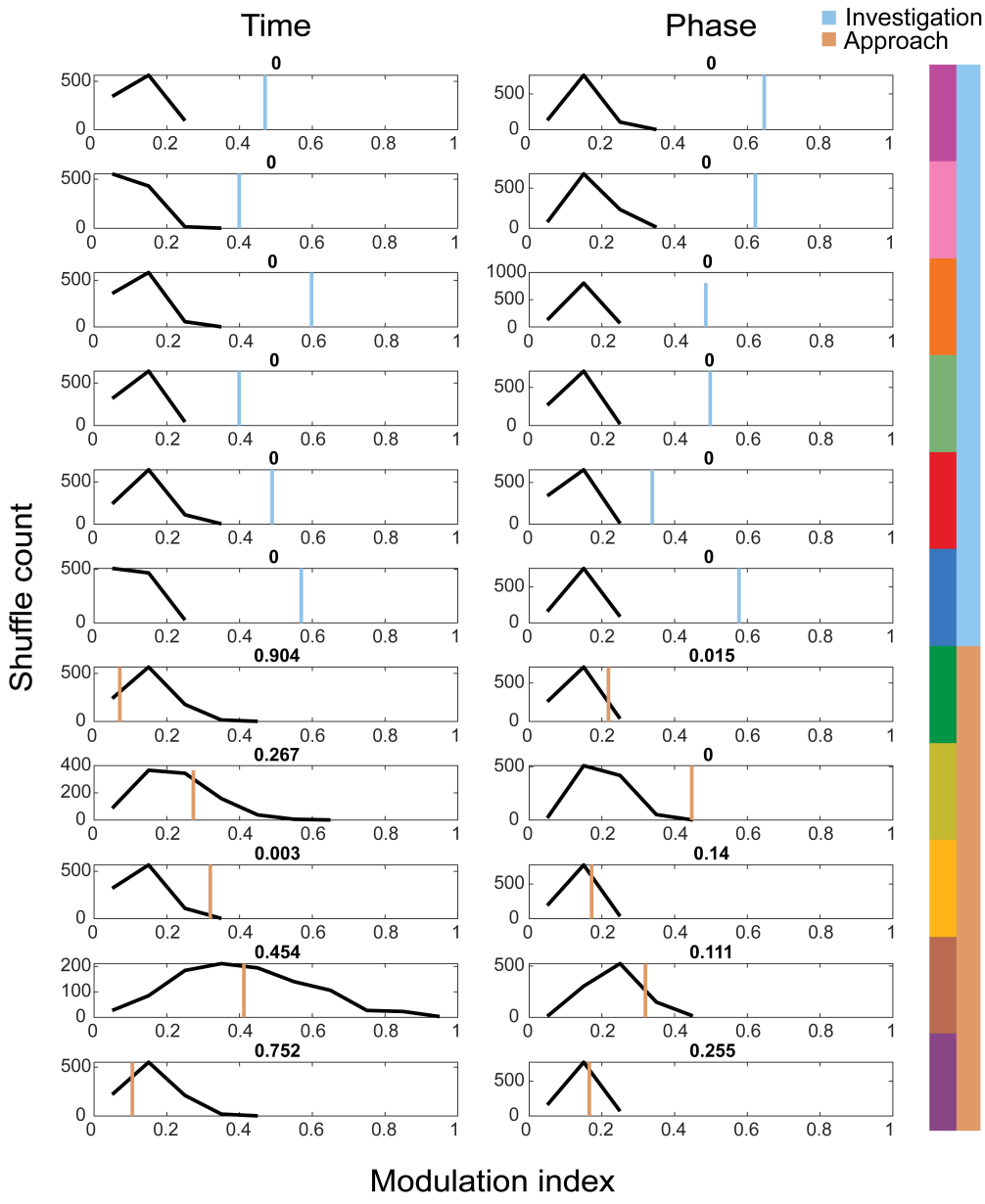

Figure 2.24. Shuffle test for sniff synchronization of motif onset for investigation and approach motifs. We calculated a modulation index (MI = (max - min)(max + min) for each motifs' across-mouse mean histogram (n = 4), and calculated the same for 1000 trial-shuffled across-mouse mean histograms. Blue and orange lines give the value from the real data, while the black histogram shows the distribution of MI across shuffle iterations.

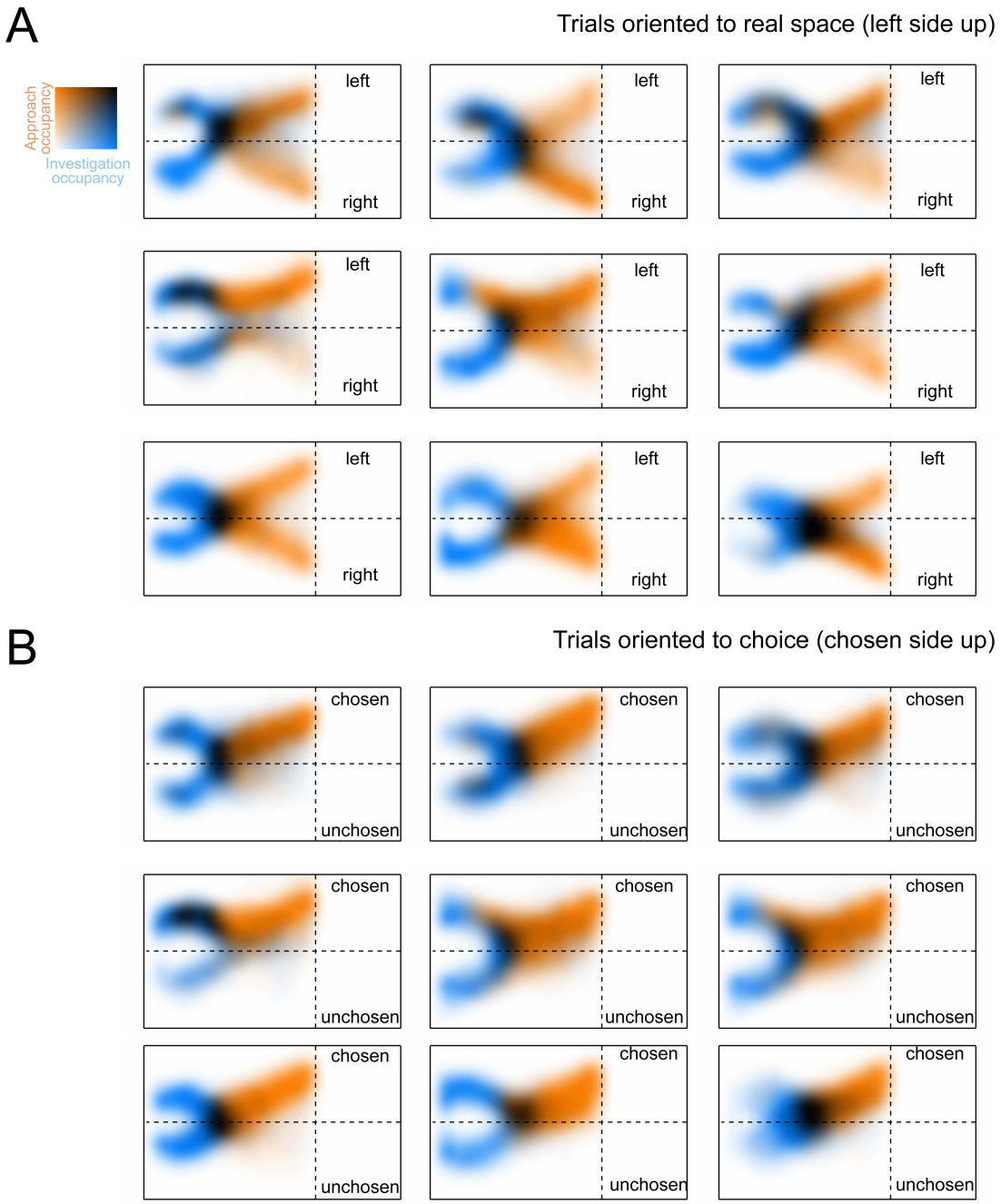

Figure 2.25. The allocentric spatial distribution of investigation and approach occupancy for individual mice. **A)** Colormaps show two-dimensional histograms of the occupancy density (1 cm^2 bins, n = 9 mice), with investigation density in blue, approach density in orange, and overlap shown by darker coloring (key in top left corner). Each map corresponds to a single mouse. Trials are oriented with respect to real space, such that the left side of the arena faces up in the figure. **B)** Same as a, except that trials are re-oriented so that the chosen side always faces up in the figure.

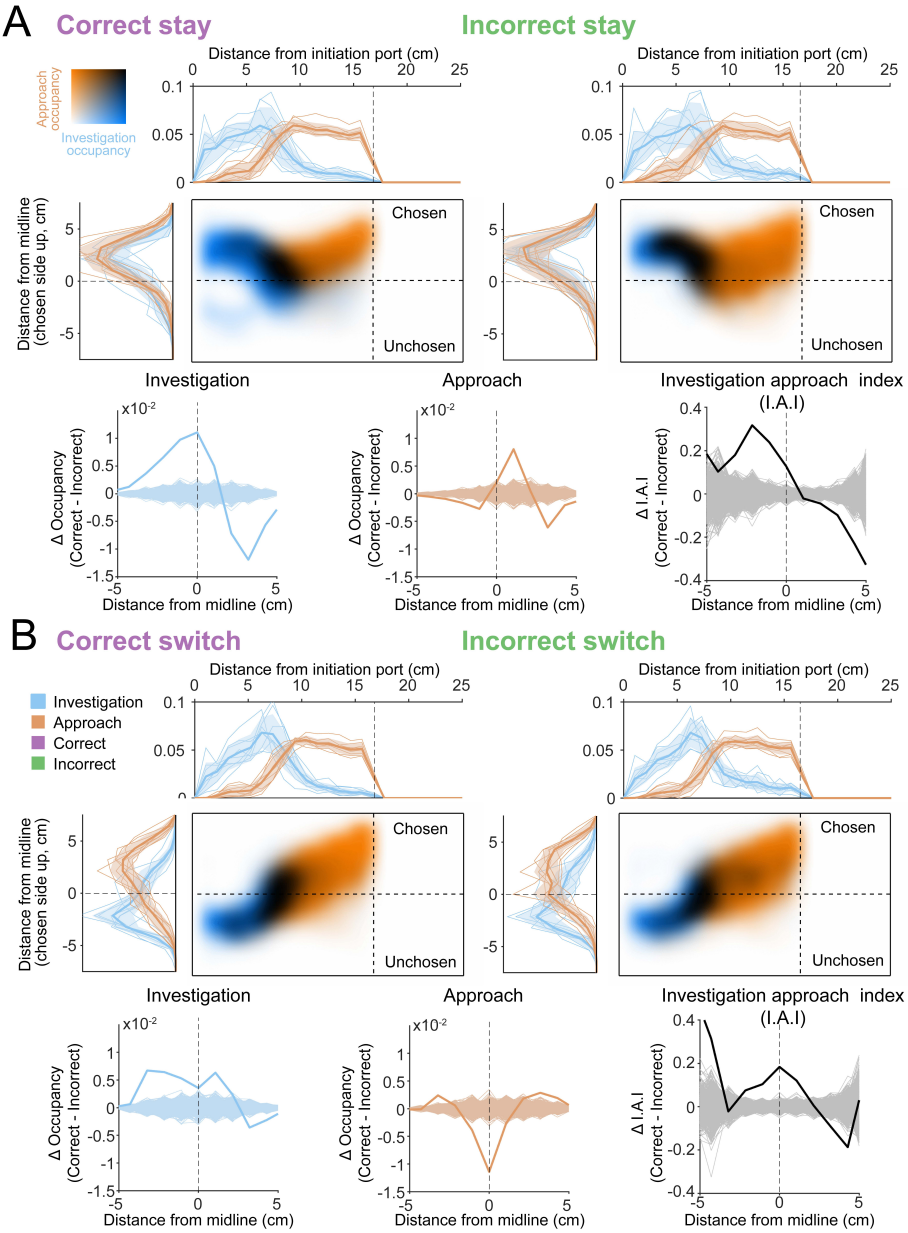

Figure 2.26. Occupancy maps indicate an advantage for investigation of both sides for both stay trials and switch trials. **A)** Occupancy map analysis for "stay" trials (trials where the mouse chooses the side it first turned to, analyses as in Fig. 2.10). **B)** Occupancy map analysis for "switch" trials (trials where the mouse chooses the opposite side from its first turn).

CHAPTER III

STATE-DEPENDENT REGULATION OF CORTICAL PROCESSING SPEED VIA GAIN MODULATION

3.1 Author contributions

Originally published as David G Wyrick, Luca Mazzucato. State-dependent regulation of cortical processing speed via gain modulation. Journal of Neuroscience, 2021. 41(18):3988–4005. LM designed research. LM and DW performed research; LM and DW wrote the paper; DW analyzed data.

3.2 Introduction

Animals respond to the same stimulus with different reaction times depending on the context or the behavioral state. Faster responses may be elicited by expected stimuli or when the animal is aroused and attentive (Niemi and Näätänen, 1981). Slower responses may occur in the presence of distractors or when the animal is disengaged from the task (Grueninger and Pribram, 1969; Treisman and Gelade, 1980; Desimone and Duncan, 1995). Experimental evidence suggests that neural correlates of these contextual modulations occur early in the cortical hierarchy, already at the level of the primary sensory cortex (Jaramillo and Zador, 2010; Samuelsen et al., 2012). During the waking state, levels of arousal, attention, and task engagement vary continuously and are associated with ongoing and large changes in the activity of neuromodulatory systems (Lee et al., 2014; Pinto et al., 2013; Fu et al., 2014) as well as cortico-cortical feedback pathways (Guo et al., 2014; Chen et al., 2017; Nelson et al., 2013; Leinweber et al., 2017; Zhang et al., 2014). Activation of these pathways modulate the patterns of activity generated by cortical circuits and may affect their

information-processing capabilities. However, the precise computational mechanism underlying these flexible reorganizations of cortical dynamics remains elusive.

Variations in behavioral and brain state, such as arousal, engagement and body movements may act on a variety of timescales, both slow (minutes, hours) and rapid (seconds or subsecond), and spatial scales, both global (pupil diameter, orofacial movements) and brain subregion-specific; and they can be recapitulated by artificial perturbations such as optogenetic, chemogenetic or electrical stimulation. These variations have been associated with a large variety of seemingly unrelated mechanisms operating both at the single cell and at the population level. At the population level, these mechanisms include modulations of low and high frequency rhythmic cortical activities (McGinley et al., 2015b); changes in noise correlations (Cohen and Maunsell, 2009; Dadarlat and Stryker, 2017); and increased information flow between cortical and subcortical networks (McGinley et al., 2015b). On a cellular level, these variations have been associated with modulations of single-cell responsiveness and reliability (Dadarlat and Stryker, 2017); and cell-type specific gain modulation (McGinley et al., 2015b). These rapid, trial-by-trial modulations of neural activity may be mediated by neuromodulatory pathways, such as cholinergic and noradrenergic systems (Lee et al., 2014; Pinto et al., 2013; Fu et al., 2014; Reimer et al., 2016), or more precise cortico-cortical projections from prefrontal areas towards primary sensory areas (Guo et al., 2014; Chen et al., 2017; Nelson et al., 2013; Leinweber et al., 2017; Zhang et al., 2014). The effects of these cortico-cortical projections can be recapitulated by optogenetic activation of glutamatergic feedback pathways (Zagha et al., 2015). In the face of this wide variety of physiological pathways, is there a common computational principle underlying the effects they elicit on sensory cortical circuits?

A natural way to model the effect of activating a specific pathway on a downstream circuit is in the form of a perturbation to the downstream circuit's afferent inputs or recurrent couplings (Mazzucato et al., 2019; Huang et al., 2019). Here, we will present a theory explaining how these perturbations control the information-processing speed of a downstream cortical circuit. Our theory shows that the effects of perturbations that change the statistics of the afferents or the recurrent couplings can all be captured by a single mechanism of action: intrinsic gain modulation, where gain is defined as the rate of change of the intrinsic input/output transfer function of a neuron measured during periods of ongoing activity. Our theory is based on a biologically plausible model of cortical circuits using clustered spiking network (Amit and Brunel, 1997). This class of models capture complex physiological properties of cortical dynamics such as state-dependent changes in neural activity, variability (Litwin-Kumar and Doiron, 2012; Deco and Hugues, 2012; Mazzucato et al., 2015, 2016; Rostami et al., 2020) and information-processing speed (Mazzucato et al., 2019). Our theory predicts that gain modulation controls the intrinsic temporal dynamics of the cortical circuit and thus its information processing speed, such that decreasing the intrinsic single-cell gain leads to faster stimulus coding.

We tested our theory by examining the effect of locomotion on visual processing in the visual hierarchy. We found that locomotion decreased the intrinsic gain of visual cortical neurons in the absence of stimuli in freely running mice. The theory thus predicted a faster encoding of visual stimuli during running compared to rest, which we confirmed in the empirical data. Our theoretical framework links gain modulation to information-processing speed, providing guidance for the design and interpretation of future manipulation experiments by unifying the changes in brain

state due to behavior, optogenetic, or pharmacological perturbations, under the same shared mechanism.

3.3 Methods

3.3.1 Spiking network model

Architecture. We modeled the local cortical circuit as a network of $N = 2000$ excitatory (E) and inhibitory (I) neurons (with relative fraction $n_E = 80\%$ and $n_I = 20\%$) with random recurrent connectivity (Fig. 3.2). Connection probabilities were $p_{EE} = 0.2$ and $p_{EI} = p_{IE} = p_{II} = 0.5$. Nonzero synaptic weights from pre-synaptic neuron j to post-synaptic neuron i were $J_{ij} = j_{ij}/\sqrt{N}$, with j_{ij} sampled from a gaussian distribution with mean $j_{\alpha\beta}$, for $\alpha, \beta = E, I$, and standard deviation δ^2. E and I neurons were arranged in p clusters. E clusters had heterogeneous sizes drawn from a gaussian distribution with a mean of $N_E^{clust} = 80$ E-neurons and 20% standard deviation. The number of clusters was then determined as $p = \text{round}(n_E N(1 - n_{bgr})/N_E^{clust})$, where $n_{bgr} = 0.1$ is the fraction of background neurons in each population, i.e., not belonging to any cluster. I clusters had equal size $N_I^{clust} = \text{round}(n_I N(1 - n_{bgr}/p)$. Clusters were defined by an increase in intra-cluster weights and a decrease in inter-cluster weights, under the constraint that the net input current to a neuron would remain unchanged compared to the case without clusters. Synaptic weights for within-cluster neurons where potentiated by a ratio factor $J_{\alpha\beta}^+$. Synaptic weights between neurons belonging to different clusters were depressed by a factor $J_{\alpha\beta}^-$. Specifically, we chose the following scaling: $J_{EI}^+ = p/(1 + (p-1)/g_{EI})$, $J_{IE}^+ = p/(1 + (p-1)/g_{IE})$, $J_{EI}^- = J_{EI}^+/g_{EI}$, $J_{IE}^- = J_{IE}^+/g_{IE}$ and $J_{\alpha\alpha}^- = 1 - \gamma(J_{\alpha\alpha}^+ - 1)$ for $\alpha = E, I$, with $\gamma = f(2 - f(p+1))^{-1}$, where $f = (1 - n_{bgr})/p$ is the fraction of E neurons in each cluster. Within-cluster E-to-E synaptic weights were further multiplied by cluster-specific factor equal to the ratio between the average cluster

Model parameters for clustered network simulations		
Parameter	Description	Value
j_{EE}	mean E-to-E synaptic weights $\times\sqrt{N}$	0.6 mV
j_{IE}	mean E-to-I synaptic weights $\times\sqrt{N}$	0.6 mV
j_{EI}	mean I-to-E synaptic weights $\times\sqrt{N}$	1.9 mV
j_{II}	mean I-to-I synaptic weights $\times\sqrt{N}$	3.8 mV
j_{E0}	mean E-to-E synaptic weights $\times\sqrt{N}$	2.6 mV
j_{I0}	mean I-to-I synaptic weights $\times\sqrt{N}$	2.3 mV
δ	standard deviation of the synaptic weight distribution	20%
J_{EE}^{+}	Potentiated intra-cluster E-to-E weight factor	14
J_{II}^{+}	Potentiated intra-cluster I-to-I weight factor	5
g_{EI}	Potentiation parameter for intra-cluster I-to-E weights	10
g_{IE}	Potentiation parameter for intra-cluster E-to-I weights	8
r_{ext}	Average baseline afferent rate to E and I neurons	5 spks/s
V_E^{thr}	E-neuron threshold potential	1.43 mV
V_I^{thr}	I-neuron threshold potential	0.74 mV
V^{reset}	E- and I-neuron reset potential	0 mV
τ_m	E- and I-neuron membrane time constant	20 ms
τ_{refr}	E- and I-neuron absolute refractory period	5 ms
τ_s	E- and I-neuron synaptic time constant	5 ms

Table 1. Parameters for the clustered network used in the simulations.

size N_E^{clust} and the size of each cluster, so that larger clusters had smaller within-cluster couplings. We chose network parameters so that the cluster timescale was 100 ms, as observed in cortical circuits (Jones et al., 2007; Mazzucato et al., 2015, 2019). Parameter values are in Table 1.

Neuronal dynamics. We modeled spiking neurons as current-based leaky-integrate-and-fire (LIF) neurons whose membrane potential V evolved according to the dynamical equation

$$\frac{dV}{dt} = -\frac{V}{\tau_m} + I_{rec} + I_{ext} \ ,$$

where τ_m is the membrane time constant. Input currents included a contribution I_{rec} coming from the other recurrently connected neurons in the local circuit and an external current $I_{ext} = I_0 + I_{stim} + I_{pert}$ (units of mV s^{-1}). The first term

$I_0 = N_{ext} J_{\alpha 0} r_{ext}$ (for $\alpha = E, I$) is a constant term representing input to the E or I neuron from other brain areas and $N_{ext} = n_E N p_{EE}$; while I_{stim} and I_{pert} represent the incoming sensory stimulus or the various types of perturbation (see Stimuli and perturbations below). When V hits threshold V_α^{thr} (for $\alpha = E, I$), a spike is emitted and V is then held at the reset value V^{reset} for a refractory period τ_{refr}. We chose the thresholds so that the homogeneous network (i.e.,where all $J_{\alpha\beta}^{\pm} = 1$) was in a balanced state with average spiking activity at rates $(r_E, r_I) = (2, 5)$ spks/s (Amit and Brunel, 1997; Mazzucato et al., 2019). Post-synaptic currents evolved according to the following equation

$$\tau_{syn} \frac{dI_{rec}}{dt} = -I_{rec} + \sum_{j=1}^{N} J_{ij} \sum_k \delta(t - t_k) \ ,$$

where τ_s is the synaptic time constant, J_{ij} are the recurrent couplings and t_k is the time of the k-th spike from the j-th presynaptic neuron. Parameter values are in Table 1.

Sensory stimuli. We considered two classes of inputs: sensory stimuli and perturbations. In the "evoked" condition (Fig. 3.4a), we presented the network one of four sensory stimuli, modeled as changes in the afferent currents targeting 50% of E-neurons in stimulus-selective clusters; each E-cluster had a 50% probability of being selective to a sensory stimulus (mixed selectivity). In the first part of the paper (Fig. 3.1-3.6, I-clusters were not stimulus-selective. Moreover, in both the unperturbed and the perturbed stimulus-evoked conditions, stimulus onset occurred at time $t = 0$ and each stimulus was represented by an afferent current $I_{stim}(t) = I_{ext} r_{stim}(t)$, where $r_{stim}(t)$ is a linearly ramping increase reaching a value $r_{max} = 20\%$ above baseline at $t = 1$. In the last part of the paper (Fig. 3.7), we introduced a new stimulation

protocol where visual stimuli targeted both E and I clusters in pairs, corresponding to thalamic input onto both excitatory and inhibitory neurons in V1 (Zhuang et al., 2013; Miska et al., 2018; Kloc and Maffei, 2014; Khan et al., 2018). Each E-I cluster pair had a 50% probability of being selective to each visual stimulus. If a E-I cluster pair was selective to a stimulus, then all E neurons and 50% of I neurons in that pair received the stimulus. The time course of visual stimuli was modeled as a double exponential profiles with rise and decay times of (0.05, 0.5)s, and peak equal to a 20% increase compared to the baseline external current.

External perturbations. We considered several kinds of perturbations. In the perturbed stimulus-evoked condition (Fig. 3.4b), perturbation onset occurred at time $t = -0.5$ and lasted until the end of the stimulus presentation at $t = 1$ with a constant time course. We also presented perturbations in the absence of sensory stimuli ("ongoing" condition, Fig. 3.2-3.3); in that condition, the perturbation was constant and lasted for the whole duration of the trial (5s). Finally, when assessing single-cell responses to perturbations, we modeled the perturbation time course as a double exponential with rise and decay times $[0.1, 1]$s (Fig. 3.6). In all conditions, perturbations were defined as follows:

- δmean(E), δmean(I): A constant offset $I_{pert} = zI_0$ in the mean afferent currents was added to all neurons in either E or I populations, respectively, expressed as a fraction of the baseline value I_0 (see Neuronal dynamics above), where $z \in [-0.1, 0.2]$ for E neurons and $z \in [-0.2, 0.2]$ for I neurons.

- δvar(E), δvar(I): For each E or I neuron, respectively, the perturbation was a constant offset $I_{pert} = zI_0$, where z is a gaussian random variable with zero mean and standard deviation σ. We chose $\sigma \in [0, 0.2]$ for E neurons and $\sigma \in [0, 0.5]$ for I neurons. This perturbation did not change the mean afferent current

but only its spatial variance across the E or I population, respectively. We measured the strength of these perturbations via their coefficient of variability $CV(\alpha) = \sigma_\alpha/\mu_\alpha$, for $\alpha = E, I$, where σ and $\mu = I_0$ are the standard deviation and mean of the across-neuron distribution of afferent currents.

- δAMPA: A constant change in the mean $j_{\alpha E} \to (1 + z)j_{\alpha E}$ synaptic couplings (for $\alpha = E, I$), representing a modulation of glutamatergic synapses. We chose $z \in [-0.1, 0.2]$.

- δGABA: A constant change in the mean $j_{\alpha I} \to (1 + z)j_{\alpha I}$ synaptic couplings (for $\alpha = E, I$), representing a modulation of GABAergic synapses. We chose $z \in [-0.2, 0.2]$.

The range of the perturbations were chosen so that the network still produced metastable dynamics for all values.

Inhibition stabilization. We simulated a stimulation protocol used in experiments to test inhibition stabilization (Fig. 3.2c). This protocol is identical to the δmean(I) perturbation during ongoing periods, where the perturbation targeted all I neurons with an external current $I_{pert} = zI_0$ applied for the whole length of 5s intervals, with $z \in [0, 1.2]$ and 40 trials per network and 10 networks for each value of the perturbation.

Simulations. All data analyses, model simulations, and mean field theory calculations were performed using custom software written in MATLAB, C and Python. Simulations in the stimulus-evoked conditions (both perturbed and unperturbed) comprised 10 realizations of each network (each network with different realization of synaptic weights), with 20 trials for each of the 4 stimuli. Simulations in the ongoing condition comprised 10 different realizations of each network, with

Model parameters for the reduced two-cluster network		
Parameter	Description	Value
j_{EE}	mean E-to-E synaptic weights $\times\sqrt{N}$	0.8 mV
j_{EI}	mean I-to-E synaptic weights $\times\sqrt{N}$	10.6 mV
j_{IE}	mean E-to-I synaptic weights $\times\sqrt{N}$	2.5 mV
j_{II}	mean I-to-I synaptic weights $\times\sqrt{N}$	9.7 mV
j_{E0}	mean E-to-E synaptic weights $\times\sqrt{N}$	14.5 mV
j_{I0}	mean I-to-I synaptic weights $\times\sqrt{N}$	12.9 mV
J^{+}_{EE}	Potentiated intra-cluster E-to-E weight factor	11.2
r^{ext}	Average baseline afferent rate to E and I neurons	7 spk/s
V^{thr}_{E}	E-neuron threshold potential	4.6 mV
V^{thr}_{I}	I-neuron threshold potential	8.7 mV
τ_s	E- and I-neuron synaptic time constant	4 ms
n_{bgr}	Fraction of background E neurons	65%

Table 2. Parameters for the simplified two-cluster network used for the mean-field theory analysis (the remaining parameters are in Table 1.

40 trials per perturbation. Each network was initialized with random synaptic weights and simulated with random initial conditions in each trial. Sample sizes were similar to those reported in previous publications (Mazzucato et al., 2015, 2016, 2019). Dynamical equations for the leaky-integrate-and-fire neurons were integrated with the Euler method with a 0.1ms step. Code to simulate the model with δvar(E) perturbation and to perform the decoding analysis on the Allen Institute Neuropixel dataset has been uploaded to: `https://github.com/mazzulab/cortical_processing_speed`. Code to reproduce the full set of perturbations investigated in this paper are available upon request to the corresponding author.

3.3.2 Mean field theory

We performed a mean field analysis of a simplified two-cluster network for leaky-integrate-and-fire neurons with exponential synapses, comprising $p + 2$ populations for $p = 2$ (Amit and Brunel, 1997; Mazzucato et al., 2019): the first p representing

the two E clusters, the last two representing the background E and the I population. The infinitesimal mean μ_n and variance σ_n^2 of the postsynaptic currents are:

$$\begin{aligned}
\mu_n &= \tau_m\sqrt{N}\left(n_E p_{EE} j_{EE}\left(f J^+_{EE} r_n + J^-_{EE}(\sum_{l=1}^{p-1} r_l + (1-pf) r_E^{bgr}) + \frac{j_{E0}}{j_{EE}} r_{ext}\right)\right) \\
&\quad -\tau_m\sqrt{N}\left(n_I p_{EI} j_{EI} r_I\right) \ , \\
\mu_{bgr} &= \tau_m\sqrt{N}\left[n_E p_{EE} j_{EE}\left(J^-_{EE}\sum_{l=1}^{p} r_l + (1-pf) r_E^{bgr} + \frac{j_{E0}}{j_{EE}} r_{ext}\right) - n_I p_{EI} j_{EI} r_I\right] \ , \\
\mu_I &= \tau_m\sqrt{N}\left[n_E p_{IE} j_{IE}\left(f\sum_{l=1}^{p} r_l + (1-pf) r_E^{bgr}\right) - n_I p_{II}(j_{II} r_I + j_{I0} r_{ext})\right] \qquad (3.1)
\end{aligned}$$

$$\begin{aligned}
\sigma_n^2 &= \tau_m\sqrt{N}\left(n_E p_{EE} j_{EE}^2\left(f(J^+_{EE})^2 r_n + (J^-_{EE})^2(\sum_{l=1}^{p-1} r_l + (1-pf) r_E^{bgr})\right)\right) \\
&\quad -\tau_m\sqrt{N}\left(n_I p_{EI} j_{EI}^2 r_I\right) \ , \\
\sigma_{bgr}^2 &= \tau_m\sqrt{N}\left[n_E p_{EE} j_{EE}^2\left((J^-_{EE})^2\sum_{l=1}^{p} r_l + (1-pf) r_E^{bgr}\right) - n_I p_{EI} j_{EI}^2 r_I\right] \ , \\
\sigma_I &= \tau_m\sqrt{N}\left[n_E p_{IE} j_{IE}^2\left(f\sum_{l=1}^{p} r_l + (1-pf) r_E^{bgr}\right) - n_I p_{II} j_{II}^2 r_I\right] \ , \qquad (3.2)
\end{aligned}$$

where $r_n, r_l = 1, \ldots, p$ are the firing rates in the p E-clusters; r_E^{bgr}, r_I, r_{ext} are the firing rates in the background E population, in the I population, and in the external current. Other parameters are described in Architecture and in Table 2. The network attractors satisfy the self-consistent fixed point equations:

$$r_l = F_l[\mu_l(\mathbf{r}), \sigma_l^2(\mathbf{r})] \ , \qquad (3.3)$$

where $\mathbf{r} = (r_1, \ldots, r_p, r_{bgr}, r_I)$ and $l = 1, \ldots, p, bgr, I$, and F_l is the current-to-rate transfer function for each population, which depend on the condition. In the absence of perturbations, all populations have the LIF transfer function

$$F_l(\mu_l, \sigma_l) = \left(\tau_{refr} + \tau_m\sqrt{\pi}\int_{H_l}^{\Theta_l} e^{u^2}[1 + \mathrm{erf}(u)]\right)^{-1} , \tag{3.4}$$

where $H_l = (V^{reset} - \mu_l)/\sigma_l + ak$ and $\Theta_l = (V_l^{thr} - \mu_l)/\sigma_l + ak$. $k = \sqrt{\tau_s/\tau_m}$ and $a = |\zeta(1/2)|/\sqrt{2}$ are terms accounting for the synaptic dynamics (Fourcaud and Brunel, 2002). The perturbations δvar(E) and δvar(I) induced an effective population transfer function F^{eff} on the E and I populations, respectively, given by (Mazzucato et al., 2019):

$$F_\alpha^{pert}(\mu_\alpha, \sigma_\alpha) = \int DzF_\alpha(\mu_\alpha + z\sigma_z\mu_\alpha^{ext}, \sigma_\alpha^2) , \tag{3.5}$$

where $\alpha = E, I$ and $Dz = dz\ \exp\left(-z^2/2/\sqrt{2\pi}\right)$ is a gaussian measure of zero mean and unit variance, $\mu_\alpha^{ext} = \tau_m\sqrt{N}n_\alpha p_{\alpha 0} j_{\alpha 0} r_{ext}$ is the external current and σ_z is the standard deviation of the perturbation with respect to baseline, denoted CV(E) and CV(I). Stability of the fixed point equation 3.3 was defined with respect to the approximate linearized dynamics of the instantaneous mean m_l and variance s_l^2 of the input currents (Mazzucato et al., 2015, 2019):

$$\tau_s\frac{dm_l}{dt} = -m_l + \mu_l(r_l) \ ; \quad \tau_s\frac{ds_l^2}{2dt} = -s_l^2 + \sigma_l^2(r_l) \ ; \quad r_l = F_l(m_l(\mathbf{r}), s_l^2(\mathbf{r})) , \tag{3.6}$$

where μ_l, σ_l^2 are defined in 3.1-3.2 and F_l represents the appropriate transfer function 3.4 or 3.5. Fixed point stability required that the stability matrix

$$S_{lm} = \frac{1}{\tau_s}\left(\frac{\partial F_l(\mu_l, \sigma_l^2)}{\partial r_m} - \frac{\partial F_l(\mu_l, \sigma_l^2)}{\partial \sigma_l^2}\frac{\partial \sigma_l^2(\mathbf{r})}{\partial r_m} - \delta_{lm}\right) , \tag{3.7}$$

was negative definite. The full mean field theory described above was used for the comprehensive analysis of Fig. 3.3. For the schematic of Fig. 3.3c, we replaced the LIF transfer function 3.4 with the simpler function $\tilde{F}(\mu_E) = 0.5(1 + \tanh(\mu_E))$ and the δvar(E) perturbation effect was then modeled as $\tilde{F}^{eff}(\mu) = \int Dz\tilde{F}(\mu_E + z\sigma_z\mu_{ext})$.

Effective mean field theory for a reduced network. To calculate the potential energy barrier separating the two network attractors in the reduced two-cluster network, we used the effective mean field theory developed in (Mascaro and Amit, 1999; Mattia et al., 2013; Mazzucato et al., 2019). The idea is to first estimate the force acting on neural configurations with cluster firing rates $\mathbf{r} = [\tilde{r}_1, \tilde{r}_2]$ outside the fixed points (3.3), then project the two-dimensional system onto a one-dimensional trajectory along which the force can be integrated to give an effective potential E. In the first step, we start from the full mean field equations for the $P = p+2$ populations in 3.3, and obtain an effective description of the dynamics for q populations "in focus" describing E clusters ($q = 2$ in our case) by integrating out the remaining $P-q$ out-of-focus populations describing the background E neurons and the I neurons ($P - q = 2$ in our case). Given a fixed value $\tilde{\mathbf{r}} = [\tilde{r}_1, \ldots, \tilde{r}_q]$ for the q in-focus populations, one obtains the stable fixed point firing rates $\mathbf{r}' = [r'_{q+1}, \ldots, r'_P]$ of the out-of-focus populations by solving their mean field equations

$$r'_\beta(\tilde{\mathbf{r}}) = F_\beta[\mu_\beta(\tilde{\mathbf{r}}, \mathbf{r}'), \sigma^2_\beta(\tilde{\mathbf{r}}, \mathbf{r}')] \ , \tag{3.8}$$

for $\beta = q + 1, \ldots, P$, as function of the in-focus populations $\tilde{\mathbf{r}}$, where stability is calculated with respect to the condition (3.7) for the reduced $(q + 1, \ldots, P)$ out-of-focus populations at fixed values of the in-focus rates $\tilde{\mathbf{r}}$. One then obtains a relation between the input $\tilde{\mathbf{r}}$ and output values $\tilde{\mathbf{r}}^{out}$ of the in-focus populations by inserting

the fixed point rates of the out-of-focus populations calculated in (3.8):

$$r_\alpha^{out}(\tilde{\mathbf{r}}) = F_\alpha[\mu_\alpha(\tilde{\mathbf{r}}, \mathbf{r}'(\tilde{\mathbf{r}})), \sigma_\alpha^2(\tilde{\mathbf{r}}, \mathbf{r}'(\tilde{\mathbf{r}}))] \ , \tag{3.9}$$

for $\alpha = 1, \ldots, q$. The original fixed points are $\tilde{\mathbf{r}}^*$ such that $\tilde{r}_\alpha^* = r_\alpha^{out}(\tilde{\mathbf{r}^*})$.

Potential energy barriers and transfer function gain. In a reduced network with two in-focus populations $[\tilde{r}_1, \tilde{r}_2]$ corresponding to the two E clusters, one can visualize Eq. (3.9) as a two-dimensional force vector $\tilde{\mathbf{r}} - \mathbf{r}^{out}(\tilde{\mathbf{r}})$ at each point in the two-dimensional firing rate space $\tilde{\mathbf{r}}$. The force vanishes at the stable fixed points A and B and at the unstable fixed point C between them (Fig. 3.3c). One can further reduce the system to one dimension by approximating its dynamics along the trajectory between A and B as (Mascaro and Amit, 1999):

$$\tau_s \frac{d\tilde{r}}{dt} = -\tilde{r} + r^{out}(\tilde{r}) \ , \tag{3.10}$$

where $y = r^{out}(\tilde{r})$ represents an effective transfer function and $\tilde{r} - r^{out}(\tilde{r})$ an effective force. We estimated the gain g of the effective transfer function as $g = 1 - \frac{r^{out}(\tilde{r}_{min}) - r^{out}(\tilde{r}_{min})}{\tilde{r}_{min} - \tilde{r}_{max}}$, where $\tilde{r}_{min}$ and $\tilde{r}_{max}$ represent, respectively, the minimum and maximum of the force (see Fig. 3.3c). From the one-dimensional dynamics (3.10) one can define a potential energy via $\frac{\partial E(\tilde{r})}{\partial r} = \tilde{r} - r^{out}(\tilde{r})$. The energy minima represent the stable fixed points A and B and the saddle point C between them represents the potential energy barrier separating the two attractors. The height Δ of the potential energy barrier is then given by

$$\Delta = \int_A^C d\tilde{r}[\tilde{r} - r^{out}(\tilde{r})] \ , \tag{3.11}$$

which can be visualized as the area of the curve between the effective transfer function and the diagonal line (see Fig. 3.3).

3.3.3 Experimental data

We tested our model predictions using the open-source dataset of neuropixel recordings from the Allen Institute for Brain Science (Siegle et al., 2019). We focused our analysis on experiments where drifting gratings were presented at four directions (0 45 90 135 and one temporal frequency (2 Hz). Out of the 54 sessions provided, only 7 sessions (6 male, 1 female) had enough trials per behavioral condition to perform our decoding analysis. Of these 7 sessions, 3 were recorded in wild-type male mice, 3 in transgenic male mice, and 1 in a transgenic female mouse. Neural activity from the visual cortical hierarchy was collected and, specifically: primary visual cortex (V1) in 5 of these 7 sessions, with a median value of 75 neurons per session; lateral visual area (LM): 6 sessions, 47 neurons; anterolateral visual area (AL): 5 sessions, 61 neurons; posteromedial visual area (PM): 6 sessions, 55; anteromedial visual area (AM): 7 sessions, 48 neurons. We matched the number and duration of trials across condition and orientation and combined trials from the drifting gratings repeat stimulus set, and drifting grating contrast stimulus set. To do this, we combined trials with low-contrast gratings (0.08, 0.1, 0.13, 0.2) and trials with high-contrast gratings (0.6, 0.8, 1) into separate trial types to perform the decoding analysis, and analyzed the interval $[-0.25, 0.5]$ seconds aligned to stimulus onset.

For evoked activity, running trials were classified as those where the animal was running faster than 3 cm/s for the first 0.5 seconds of stimulus presentation. During ongoing activity, behavioral periods were broken up into windows of 1 second. Periods of running or rest were classified as such if 10 seconds had elapsed without a behavioral change. Blocks of ongoing activity were sorted and used based on the

length of the behavior. Out of the 54 sessions provided, 14 sessions had enough time per behavioral condition (minimum of 2 minutes) to estimate single-cell transfer functions. Only neurons with a mean firing rate during ongoing activity greater than 5Hz were included in the gain analysis (2119 out of 4365 total neurons).

3.3.4 Stimulus decoding

For both the simulations and data, a multi-class decoder was trained to discriminate between four stimuli from single-trial population activity vectors in a given time bin (Jezzini et al., 2013). To create a timecourse of decoding accuracy, we used a sliding window of 100ms (200ms) in the data (model), which was moved forward in 2ms (20ms) intervals in the data (model). Trials were split into training and test data-sets in a stratified 5-fold cross-validated manner, ensuring equal proportions of trials per orientation in both data-sets. In the model, a leave-2-out cross-validation was performed. To calculate the significance of the decoding accuracy, an iterative shuffle procedure was performed on each fold of the cross-validation. On each shuffle, the training labels were shuffled and the classifer accuracy was predicted on the unshuffled test data-set. This shuffle was performed 100 times to create a shuffle distribution to rank the actual decoding accuracy from the unshuffled decoder against and to determine when the mean decoding accuracy had increased above chance. This time point is what we referred to as the latency of stimulus decoding. To account for the speed of stimulus decoding (the slope of the decoding curve), we defined the Δ-Latency between running and rest as the average time between the two averaged decoding curves from 40% up to 80% of the max decoding value at rest.

3.3.5 Firing rate distribution match

To control for increases of firing rate due to locomotion (Fig. 3.7b), we matched the distributions of population counts across the trials used for decoding in both

behavioral conditions. This procedure was done independently for each sliding window of time along the decoding time course. Within each window, the spikes from all neurons were summed to get a population spike count per trial. A log-normal distribution was fit to the population counts across trials for rest and running before the distribution match (Fig 3.8a left). We sorted the distributions for rest and running in descending order, randomly removing spikes from trials in the running distribution to match the corresponding trials in the rest distribution (Fig 3.8a right). By doing this, we only removed the number of spikes necessary to match the running distribution to rest distribution. For example, trials where the rest distribution had a larger population count, no spikes were removed from either distribution. Given we performed this procedure at the population level rather than per neuron, we checked the change in PSTH between running and rest conditions before and after distribution matching (Fig 3.8b).

3.3.6 Single-cell gain

To infer the single-cell transfer function in simulations and data, we followed the method originally described in (Recanatesi et al., 2020) (see also (Lim et al., 2015; Pereira and Brunel, 2018) for a trial-averaged version). We estimated the transfer function on ongoing periods when no sensory stimulus was present. Briefly, the transfer function of a neuron was calculated by mapping the quantiles of a standard gaussian distribution of input currents to the quantiles of the empirical firing rate distribution during ongoing periods (Fig. 3.3e). We then fit this transfer function with a sigmoidal function. The max firing rate of the neuron in the sigmoidal fit was bounded to be no larger than 1.5 times that of the empirical max firing rate, to ensure realistic fits. We defined the gain as the slope at the inflection point of the sigmoid.

3.3.7 Single-cell response and selectivity

We estimated the proportion of neurons that were significantly excited or inhibited by cortical state perturbations in the model (Fig. 3.6) or locomotion in the data (Fig. 3.7) during periods of ongoing activity, in the absence of sensory stimuli. In the model, we simulated 40 trials per network, for 10 networks per each value of the perturbation. Only in this perturbation-response condition in the absence of sensory stimuli, the perturbation was modeled as a double exponential with rise and decay times $[0.2, 1]$ (Fig. 3.3a), with 0.5s of unperturbed activity preceding perturbation onset. In the data, we binned the spike counts in 500ms windows for each neuron after matching sample size between rest and running conditions, and significant difference between the ongoing and perturbed epochs was assessed with a rank-sum test.

We estimated single neuron selectivity to sensory stimuli in each condition from the average firing rate responses $r_i^a(t)$ of the i-th neuron to stimulus a in trial t. For each pair of stimuli, selectivity was estimated as

$$d'(a,b) = \frac{\text{mean}\left[r(t)^a\right] - \text{mean}\left[r(t)^b\right]}{\sqrt{\frac{1}{2}\left(\text{var}[r(t)^a] + \text{var}[r(t)^b]\right)}} \ ,$$

where mean and var are estimated across trials. The d' was then averaged across stimulus pairs.

3.4 Results

To elucidate the effect of state changes on cortical dynamics, we modeled the local circuit as a network of recurrently connected excitatory (E) and inhibitory (I) spiking neurons. Both E and I populations were arranged in clusters (Amit and Brunel, 1997; Litwin-Kumar and Doiron, 2012; Mazzucato et al., 2015; Schaub et al., 2015; Mazzucato et al., 2019), where synaptic couplings between neurons in the

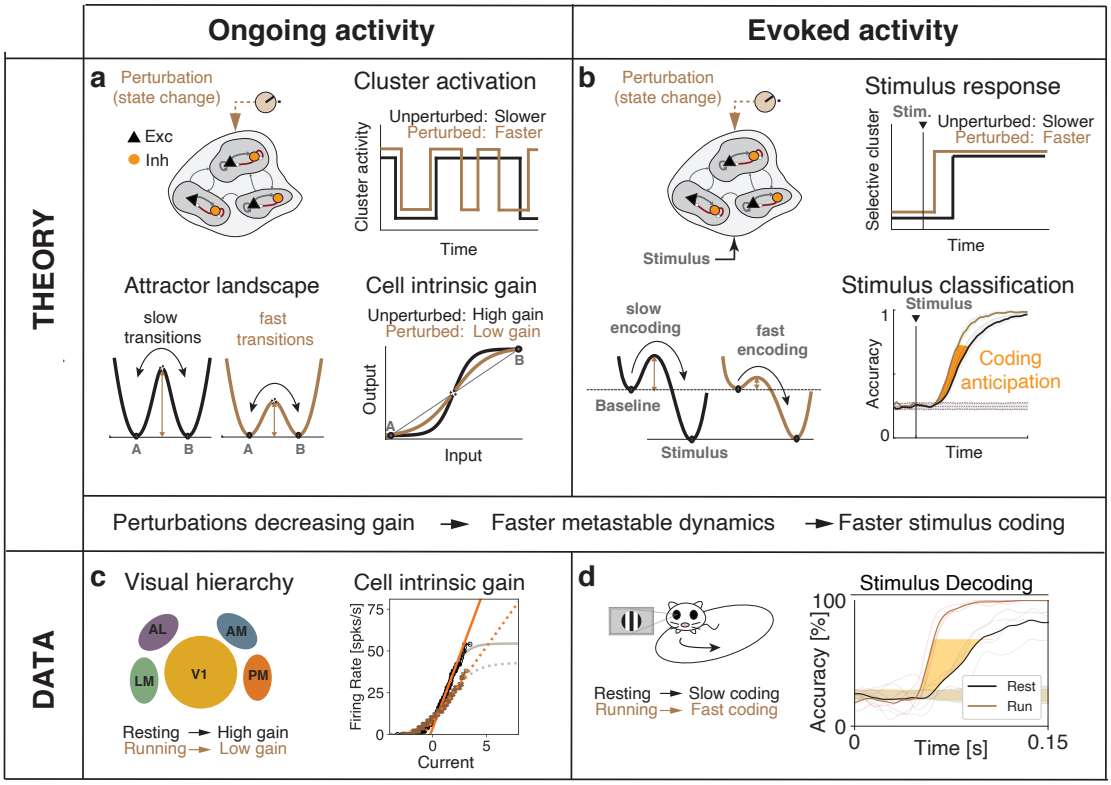

Figure 3.1. Conceptual summary of the main results. **a)** In a network model of sensory cortex featuring clusters of excitatory and inhibitory neurons with metastable dynamics, state changes are induced by external perturbations controlling the timescale of cluster activation during ongoing activity. The neural mechanism underlying timescale modulation is a change in the barrier height separating attractors, driven by a modulation of the intrinsic gain of the single-cell transfer function. **b)** During evoked activity, onset of stimulus encoding is determined by the activation latency of stimulus-selective cluster. External perturbations modulate the onset latency thus controlling the stimulus processing speed. The theory shows that the effect of external perturbations on stimulus-processing speed during evoked activity (right) can be predicted by the induced gain modulations observed during ongoing activity (left). **c)** Locomotion induced changes in intrinsic gain in the visual cortical hierarchy during darkness periods. **d)** Locomotion drove faster coding of visual stimuli during evoked periods, as predicted by the induced gain modulations observed during ongoing activity.

same cluster were potentiated compared to neurons in different clusters, reflecting the empirical observation of cortical assemblies of functionally correlated neurons (Fig. 3.2a; Kiani et al., 2015; Song et al., 2005; Perin et al., 2011; Lee et al., 2016). In the absence of external stimulation (ongoing activity), this E-I clustered network

generates rich temporal dynamics characterized by metastable activity operating in inhibition stabilized regime, where E and I post-synaptic currents track each other achieving tight balance (Fig. 3.2c). A heterogeneous distribution of cluster sizes leads to a lognormal distributions of firing rates (Fig. 3.2b). Network activity was characterized by the emergence of the slow timescale of cluster transient activation, with average activation lifetime of $\tau = 106 \pm 35$ ms (hereby referred to as "cluster timescale," Fig. 3.2b), much larger than single neuron time constant (20ms; Litwin-Kumar and Doiron, 2012; Mazzucato et al., 2015).

Figure 3.2 (next page). Biological plausible model of cortical circuit. **a)** Left: Schematics of the network architecture. A recurrent network of E (black triangles) and I (red circles) spiking neurons arranged in clusters is presented sensory stimuli targeting subsets of E clusters, in different cortical states implemented by perturbations. *Inset* shows a membrane potential trace from representative E neuron. Right: Synaptic couplings J_{ij} for a representative clustered network, highlighting the block diagonal structure of potentiated intra-cluster synaptic weights for both E and I clusters, and the background E and I populations (bgr). Cluster size was heterogeneous (inset). **b)** Representative neural activity during ongoing periods; tick marks represent spike times of E (black) or I (red) neurons. The network dynamics is metastable with clusters transiently activity for periods of duration τ. Inset: The cumulative distributions of single-cell firing rates (in the representative network are lognormal (blue: empirical data; orange: lognormal fit). **c)** When increasing the inhibitory drive (afferent current to the I population, same as δmean(I) perturbation), both E and I firing rates decrease (black and red curve in right panel, mean$\pm$s.e.m. across 10 simulated networks), highlighting the paradoxical effect, signature of the inhibition stabilized regime (Tsodyks et al., 1997). Beyond δmean(I)=50% the E population shuts down and the I population rebounds (dashed vertical line). **d)** Left: State-changing perturbation affecting the mean of the afferent currents to E populations (knobs represent changes in afferent to three representative E cells compared to the unperturbed state). Right: Histogram of afferent inputs to E-cells in the perturbed state (brown, all neurons receive a 10% increase in input) with respect to the unperturbed state (grey). **e)** Left: State-changing perturbation affecting the variance of afferent currents to E populations. Right: In the perturbed state (brown), each E-cell's afferent input is constant and sampled from a normal distribution with mean equal to the unperturbed value (grey) and 10% CV.

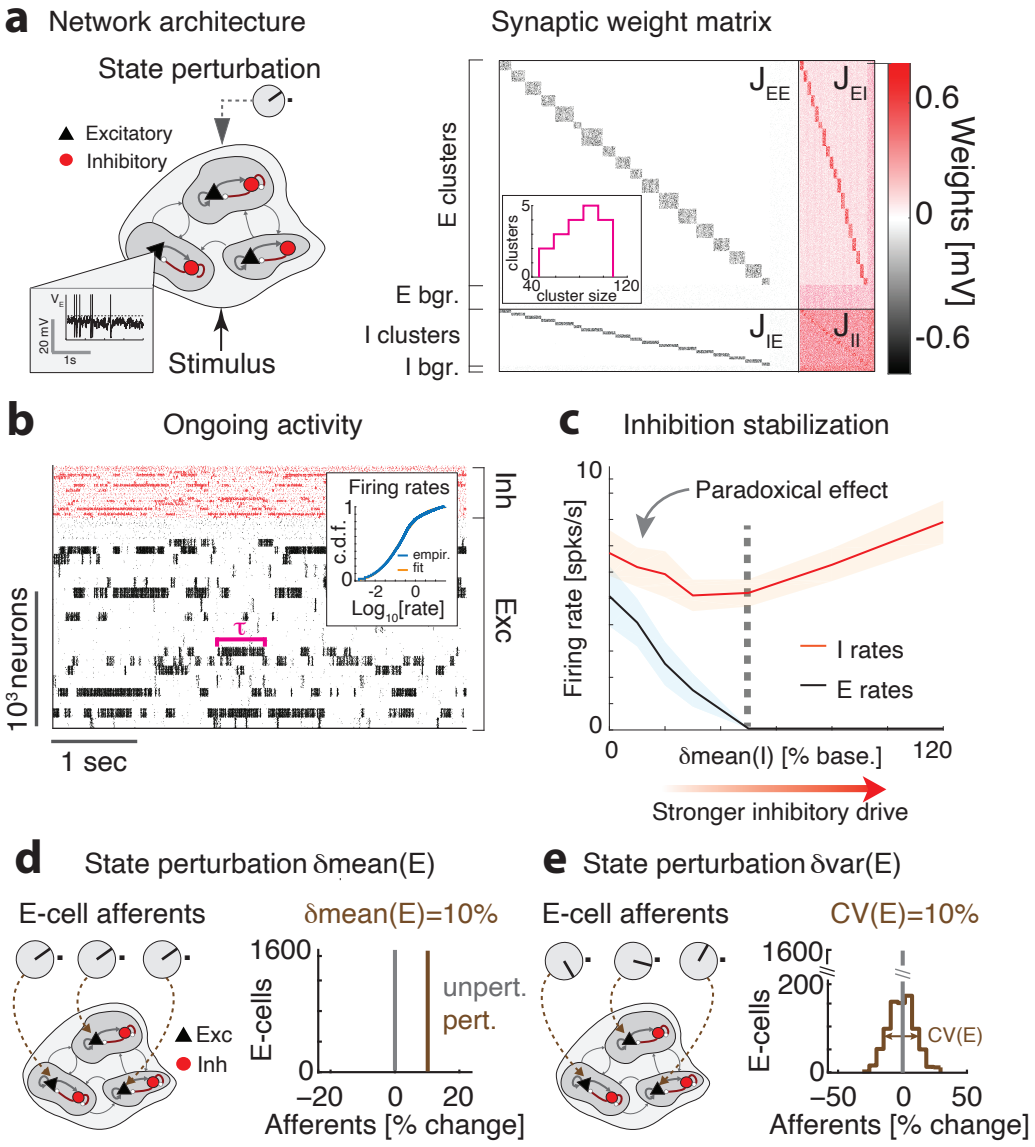

Figure 3.2.

To investigate how changes in cortical state may affect the network dynamics and information processing capabilities, we examined a vast range of state-changing perturbations (Fig. 3.2d-e, Table 3). State changes were implemented as perturbations of the afferent currents to cell-type specific populations, or as perturbations to the synaptic couplings. The first type of state perturbations

δmean(E) affected the mean of the afferent currents to E populations (Fig. 3.2d). E.g., a perturbation δmean(E)=10% implemented an increase of all input currents to E neurons by 10% above their unperturbed levels. The perturbation δmean(I) affected the mean of the afferent currents to I populations in an analogous way. The second type of state perturbations δvar(E) affected the across-neuron variance of afferents to E populations. Namely, in this perturbed state, the afferent current to each neuron in that population was sampled from a normal distribution with zero mean and fixed variance (Fig. 3.2e, measured by the coefficient of variability CV(E)=var(E)/mean(E) with respect to the unperturbed afferents). This perturbation thus introduced a *spatial variance* across neurons in the cell-type specific afferent currents, yet left the mean afferent current into the population unchanged. The state perturbation δvar(I) affected the variance of the afferent currents to I populations analogously. In the third type of state perturbations δAMPA or δGABA, we changed the average GABAergic or glutamatergic (AMPA) recurrent synaptic weights compared to their unperturbed values. We chose the range of state perturbations such that the network still retained non-trivial metastable dynamics within the whole range. We will refer to these state changes of the network as simply perturbations, and should not be confused with the presentation of the stimulus. We first established the effects of perturbations on ongoing network dynamics, and used those insight to explain their effects on stimulus-evoked activity.

3.4.1 State-dependent regulation of the network emergent timescale

A crucial feature of neural activity in clustered networks is metastable attractor dynamics, characterized by the emergence of a long timescale of cluster activation whereby network itinerant activity explores the large attractor landscape (Fig. 3.2b). We first examined whether perturbations modulated the network's metastable

dynamics and introduced a protocol where perturbations occurred in the absence of sensory stimuli ("ongoing activity").

We found that perturbations strongly modulated the attractor landscape, changing the repertoire of attractors the network activity visited during its itinerant dynamics (Fig. 3.3a). Changes in attractor landscape were perturbation-specific. Perturbations increasing δmean(E) (δmean(I)) induced a consistent shift in the repertoire of attractors: larger perturbations led to larger (smaller) numbers of co-active clusters. Surprisingly, perturbations that increased δvar(E) (δvar(I)), led to network configurations with larger (smaller) sets of co-activated clusters. This effect occurred despite the fact that such perturbations did not change the mean afferent input to the network. Perturbations affecting δAMPA and δGABA had similar effects to δmean(E) and δmean(I), respectively.

We then examined whether perturbations affected the cluster activation timescale. We found that perturbations differentially modulated the average cluster activation timescale τ during ongoing periods, in the absence of stimuli (Fig. 3.3b). In particular, increasing δmean(E), δvar(E), or δAMPA led to a proportional acceleration of the network metastable activity and shorter τ; while increasing δmean(I), δvar(I) or δGABA induced the opposite effect with longer τ. Changes in τ were congruent with changes in the duration of intervals between consecutive activations of the same cluster (cluster inter-activation intervals).

3.4.2 Changes in cluster timescale are controlled by gain modulation

What is the computational mechanism mediating the changes in cluster timescale, induced by the perturbations? We investigated this question using mean field theory, where network attractors, defined by sets of co-activated clusters, are represented as potential wells in an attractor landscape (Mascaro and Amit, 1999;

Litwin-Kumar and Doiron, 2012; Mazzucato et al., 2015; Mattia and Sanchez-Vives, 2012; Mazzucato et al., 2019). Let us illustrate this in a simplified network with two clusters (Fig. 3.3c). Here, the attractor landscape consists of two potential wells, each well corresponding to a configuration where one cluster is active and the other is inactive. When the network activity dwells in the attractor represented by the left potential well, it may escape to the right potential well due to internally generated variability. This process will occur with a probability determined by the height Δ of the barrier separating the two wells: the higher the barrier, the less likely the transition (Hänggi et al., 1990; Litwin-Kumar and Doiron, 2012; Mattia and Sanchez-Vives, 2012; Mazzucato et al., 2019).

Mean field theory thus established a relationship between the cluster timescale and the height of the barrier separating the two attractors. We found that perturbations differentially control the height of the barrier Δ separating the two attractors (Fig. 3.3d), explaining the changes in cluster timescale observed in the simulations (Fig. 3.3b).

Since reconstruction of the attractor landscape requires knowledge of the network's structural connectivity, the direct test of the mean field relation between changes in attractor landscape and timescale modulations is challenging. We thus aimed at obtaining an alternative formulation of the underlying neural mechanism only involving quantities directly accessible to experimental observation. Using mean field theory, one can show that the double potential well representing the two attractors can be directly mapped to the effective transfer function of a neural population (Mascaro and Amit, 1999; Mazzucato et al., 2019; Mattia et al., 2013). One can thus establish a direct relationship between changes in the slope (hereby referred to as "gain") of the intrinsic transfer function estimated during ongoing

periods and changes in the barrier height Δ separating metastable attractors (see Fig. 3.3c-d and Methods). In turn, this implies a direct relationship between gain modulation, induced by the perturbations, and changes in cluster activation timescale. In particular, perturbations inducing steeper gain will increase well depths and barrier heights, and thus increase the cluster timescale, and vice versa. Using mean field theory, we demonstrated a complete classification of the differential effect of all perturbations on barrier heights and gain (Fig. 3.3d).

We then proceeded to verify these theoretical predictions, obtained in a simplified two-cluster network, in the high dimensional case of large networks with several

---

Figure 3.3 (next page). Linking gain modulation to changes in cluster timescale.**a)** Clustered network activity during a representative unperturbed ongoing trial hops among different metastable attractors (grey box: attractor with 3 co-active clusters). Right: Perturbations strongly modulate the attractor landscape (color-coded curves: frequency of occurrence of network attractors with different number of co-active clusters, for different values of the representative δmean(E) perturbation, mean occurrence across 5 sessions). **b)** Perturbations induce consistent changes in the average cluster activation timescale τ (mean$\pm$S.D. across 10 simulated sessions) and in the single neuron intrinsic gain ($\Delta Gain$=Gain(pert.)-Gain(unpert.), estimated as in panel **e** bottom panel: color-coded markers represent different perturbations, linear regression, $R^2 = 0.96$). **c)** Schematic of the effect of perturbations on network dynamics in a two-cluster network, captured by a double-well potential (top panel). Potential wells represent two attractors where either cluster is active (A and B; left: unperturbed), separated by a barrier with height Δ. Mean field theory links perturbation effects on the barrier height (top right, lower barrier) to changes in the intrinsic neuronal gain (bottom right, lower gain). **d)** Mean field theory predictions linking the height of the barrier Δ separating the attractors to the intrinsic cell gain, for all perturbations (linear regression of Δ vs gain). **e)** Inferring the single cell intrinsic transfer function from spiking activity during ongoing periods. By recasting the distribution of spike counts (left) into quantiles (top center), one can match those values to corresponding quantiles of a gaussian input current distribution (bottom center) to obtain the current-to-rate function (right). **d)**: A single-cell transfer function (bottom, empirical data in blue; sigmoidal fit in brown) can be estimated by matching a neuron's firing rate distribution during ongoing periods (top) to a gaussian distribution of input currents (center, quantile plots; red stars denotes matched median values).

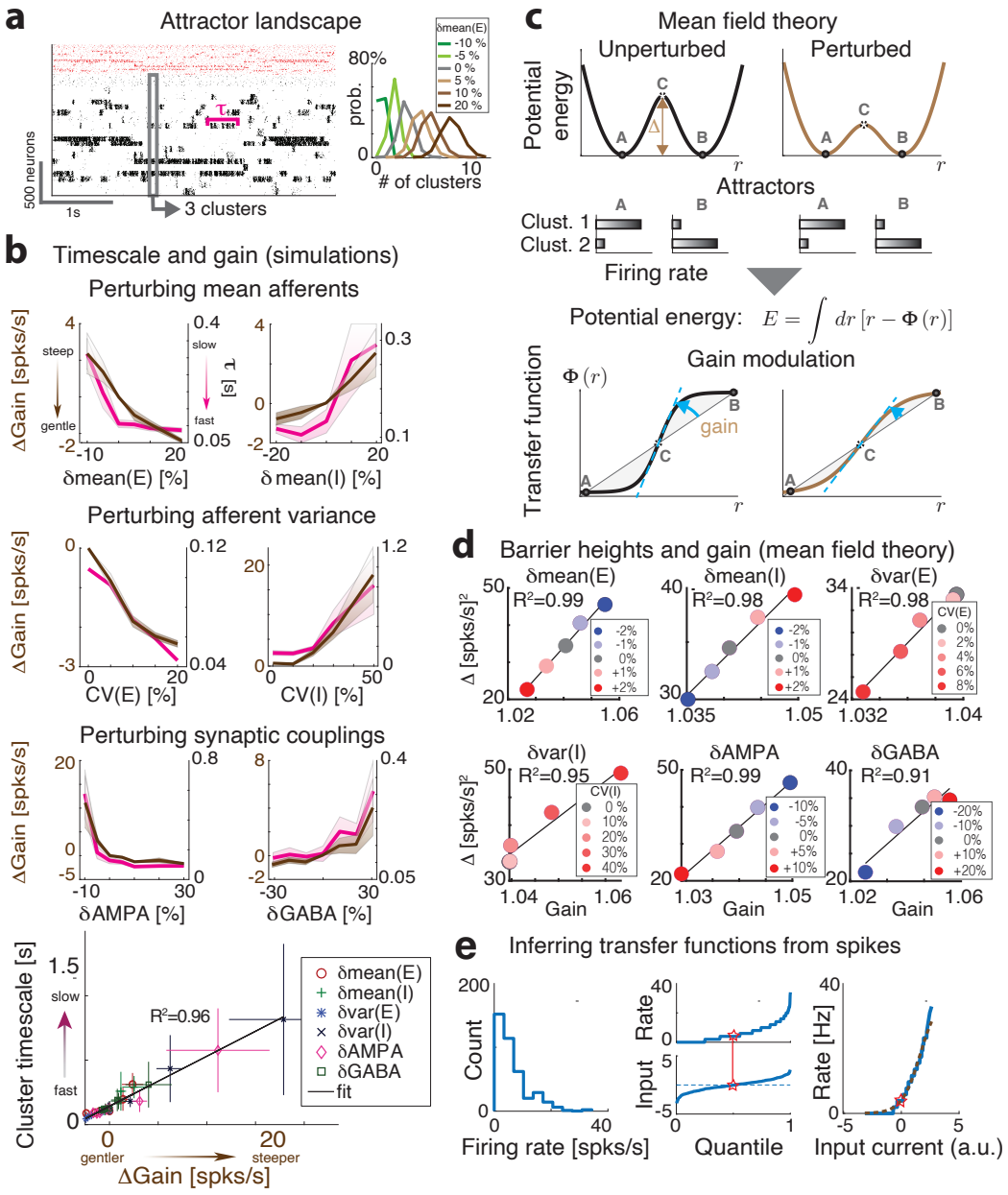

Figure 3.3.

clusters using simulations. While barrier heights and the network's attractor landscape can be exactly calculated in the simplified two-cluster network, this task is infeasible in large networks with a large number of clusters where the number of attractors is exponential in the number of clusters. On the other hand, changes

in barrier heights Δ are equivalent to changes in gain, and the latter can be easily estimated from spiking activity (Fig. 3.3b-d). We thus tested whether the relation between gain and timescale held in the high-dimensional case of a network with many clusters. We estimated single-cell transfer functions from their spiking activity during ongoing periods, in the absence of sensory stimuli but in the presence of different perturbations (Fig. 3.3e, Lim et al. (2015); Recanatesi et al. (2020)). We found that network perturbations strongly modulated single-cell gain in the absence of stimuli, verifying mean field theory predictions in all cases (Fig. 3.3e). In particular, we confirmed the direct relationship between gain modulation and cluster timescale modulation: perturbations that decreased (increased) the gain also decreased (increased) cluster timescale (Fig. 3.3b, $R^2 = 0.96$). For all perturbations, gain modulations explained the observed changes in cluster timescale.

3.4.3 Controlling information processing speed with perturbations

We found that changes in cortical state during ongoing activity, driven by external perturbations, control the circuit's dynamical timescale. The neural mechanism mediating the effects of external perturbations is gain modulation, which controls the timescale of the network switching dynamics. How do such state changes affect the network information processing speed?

To investigate the effect of state perturbations on the network's information-processing, we compared stimulus-evoked activity by presenting stimuli in an unperturbed and a perturbed condition. In unperturbed trials (Fig. 3.4a), we presented one of four sensory stimuli, modeled as depolarizing currents targeting a subset of stimulus-selective E neurons with linearly ramping time course. Stimulus selectivities were mixed and random, all clusters having equal probability of being stimulus-selective. In perturbed trials (Fig. 3.4b), in addition to the same sensory

stimuli, we included a state perturbation, which was turned on before the stimulus and was active until the end of stimulus presentation. We investigated and classified the effect of several state changes implemented by perturbations affecting either the mean or variance of cell-type specific afferents to E or I populations, and the synaptic couplings. State perturbations were identical in all trials of the perturbed condition for each type; namely, they did not convey any information about the stimuli.

We assessed how much information about the stimuli was encoded in the population spike trains at each moment using a multiclass classifier (with four class labels corresponding to the four presented stimuli, Fig. 3.4c). In the unperturbed condition, the time course of the cross-validated decoding accuracy, averaged across stimuli, was significantly above chance after $0.21 + / - 0.02$ seconds (mean±s.e.m. across 10 simulated networks, black curve in Fig. 3.4c) and reached perfect accuracy after a second. In the perturbed condition, stimulus identity was decoded at chance level in the period after the onset of the state perturbation but before stimulus presentation (Fig. 3.4c), consistent with the fact that the state perturbation did not convey information about the stimuli. We found that state perturbations significantly modulated the network information processing speed. We quantified this modulation as the average latency to reach a decoding accuracy between 40% and 80% (Fig. 3.4c, yellow area), and found that state perturbations differentially affected processing speed.

State perturbations had opposite effects depending on which cell-type specific populations they targeted. Increasing δmean(E) monotonically improved network performance (Fig. 3.4d, left panel): in particular, positive perturbations induced an anticipation of stimulus-coding (shorter latency), while negative ones led to longer latency and slower coding. The opposite effect was achieved when increasing

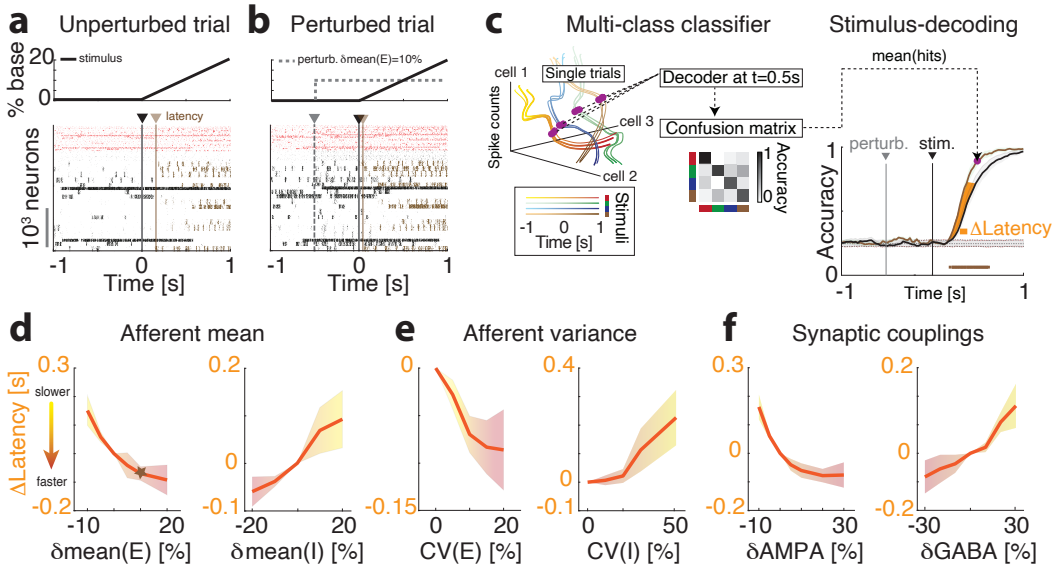

Figure 3.4. Perturbations control stimulus-processing speed in the clustered network. **a-b**) Representative trials in the unperturbed (**a**) and perturbed (**b**) conditions; the representative perturbation is an increase in the spatial variance δvar(E) across E neurons. After a ramping stimulus is presented at $t = 0$ (black vertical line on raster plot; top panel, stimulus time course), stimulus-selective E-clusters (brown tick marks represent their spiking activity) are activated at a certain latency (brown vertical line). In the perturbed condition (**b**), a perturbation is turned on before stimulus onset (gray-dashed vertical line). The activation latency of stimulus-selective clusters is shorter in the perturbed compared to the unperturbed condition. **c**) Left: schematic of stimulus-decoding analysis. A multi-class classifier is trained to discriminate between the four stimuli from single-trial population activity vectors in a given time bin (curves represent the time course of population activity in single trials, color-coded for 4 stimuli; the purple circle highlights a given time bin along the trajectories), yielding a cross-validated confusion matrix for the decoding accuracy at that bin (central panel). Right: Average time course of the stimulus-decoding accuracy in the unperturbed (black) and perturbed (brown) conditions (horizontal brown: significant difference between conditions, $p < 0.05$ with multiple bin correction). **d-f**: Difference in stimulus decoding latency in the perturbed minus the unperturbed conditions (average difference between decoding accuracy time courses in the [40%,80%] range, yellow interval of **c**; mean±S.D. across 10 networks) for the six state-changing perturbations (see Methods and main text for details; the brown star represents the perturbation in **b-c**).

δmean(I), which slowed down processing speed (Fig. 3.4d, right panel). State perturbations that changed the spatial variance of the afferent currents had counterintuitive effects (Fig. 3.4e). We measured the strength of these perturbations via their coefficient of variability $CV(\alpha) = \sigma_\alpha/\mu_\alpha$, for $\alpha = E, I$, where σ and μ are the

standard deviation and mean of the across-neuron distribution of afferent currents. Perturbations δvar(E) that increased $CV(E)$ led to faster processing speed. The opposite effect was achieved with perturbations δvar(I) inducing a spatial variance across afferents to I neurons, which slowed down stimulus-processing speed (Fig. 3.4e). Perturbations δAMPA which increased the glutamatergic synaptic weights improved performance proportionally to the perturbation. The opposite effect was achieved by perturbations δGABA that increased the GABAergic synaptic weights, which monotonically decreased network processing speed (Fig. 3.4f). We thus concluded that afferent current perturbations differentially modulated the speed at which network activity encoded information about incoming sensory inputs. Such modulations exhibited a rich dynamical repertoire (Table 3).

3.4.4 Gain modulation regulates the network information processing speed

Our mean field framework demonstrates a direct relationship between the effects of perturbations on the network information processing speed and its effects on the cluster timescale (Fig. 3.3). In our simplified network with two clusters, stimulus presentation induces an asymmetry in the well depths, where the attractor B corresponding to the activated stimulus-selective cluster has a deeper well, compared to the attractor A where the stimulus-selective cluster is inactive. Upon stimulus presentation, the network ongoing state will be biased to transition towards the stimulus-selective attractor B with a transition rate determined by the barrier height separating A to B. Because external perturbations regulate the height of such barrier via gain modulation, they control in turn the latency of activation of the stimulus-selective cluster. We thus aimed at testing the prediction of our theory: that the perturbations modulate stimulus coding latency in the same way as they modulate

cluster timescales during ongoing periods; and, as a consequence, that these changes in stimulus coding latency can be predicted by intrinsic gain modulation. Specifically, our theory predicts that perturbation driving a decrease (increase) in intrinsic gain during ongoing periods will induce a faster (slower) encoding of the stimulus.

We thus proceeded to test the relationship between perturbations effects on cluster timescales, gain modulation, and information processing speed. In the representative trial where the same stimulus was presented in the absence (Fig. 3.4a) or in the presence (Fig. 3.4b) of the perturbation δmean(E)= 10%, we found that stimulus-selective clusters (highlighted in brown) had a faster activation latency in response to the stimulus in the perturbed condition compared to the unperturbed one. A systematic analysis confirmed this mechanism showing that, for all perturbations, the activation latency of stimulus-selective clusters was a strong predictor of the change in decoding latency (Fig. 3.5a right panel, $R^2 = 0.93$). Moreover, we found that the perturbation-induced changes of the cluster timescale τ during ongoing periods predicted the effect of the perturbation on stimulus-processing latency during evoked periods (Fig. 3.5b,d). Specifically, perturbations inducing faster τ during ongoing periods, in turn accelerated stimulus coding; and vice versa for perturbations inducing slower τ.

We then tested whether perturbation-induced gain modulations during ongoing periods explained the changes in stimulus-processing speed during evoked periods, and found that the theoretical prediction was borne out in the simulations (Fig. 3.5c,e). Let us summarize the conclusion of our theoretical analyses. Motivated by mean field theory linking gain modulation to changes in transition rates between attractors, we found that gain modulation controls the cluster timescale during ongoing periods, and, in turn, regulates the onset latency of stimulus-selective clusters

	Latency	Activity	Response	τ	Gain
δmean(E)$\nearrow$	$\searrow$	E[$\nearrow$], I[$\nearrow$]	E[$\nearrow$], I[$\nearrow$]	$\searrow$	$\searrow$
δmean(I)$\nearrow$	$\nearrow$	E[$\searrow$], I[$\searrow$]	E[$\searrow$], I[mixed]	$\nearrow$	$\nearrow$
δvar(E)$\nearrow$	$\searrow$	E[$\nearrow$], I[$\nearrow$]	E[mixed], I[mixed]	$\searrow$	$\searrow$
δvar(I)$\nearrow$	$\nearrow$	E[$\searrow$], I[=]	E[$\searrow$], I[$\searrow$]	$\nearrow$	$\nearrow$
δAMPA$\nearrow$	$\searrow$	E[$\nearrow$], I[$\nearrow$]	E[$\nearrow$], I[$\nearrow$]	$\searrow$	$\searrow$
δGABA$\nearrow$	$\nearrow$	E[$\searrow$], I[$\searrow$]	E[$\searrow$], I[mixed]	$\nearrow$	$\nearrow$
Locomotion	$\searrow$	E[$\nearrow$], I[$\nearrow$]	E[mixed], I[mixed]	$\searrow$	$\searrow$

Table 3. Classification of state-changing perturbations. Effect of on neural activity of an increasing ($\nearrow$) state-changing perturbation: latency of stimulus decoding ('Latency', Fig. 3.2e); average firing rate modulation ('Activity') and response to perturbations ('Response', proportion of cells with significant responses) of E and I cells in the absence of stimuli (Fig. 3.3); cluster activation timescale ('τ', Fig. 3.4b); single-cell intrinsic gain modulation at rest ('Gain', Fig. 3.5e). $\nearrow, \searrow, =$ represent increase, decrease, and no change, respectively. 'Mixed' responses refer to similar proportions of excited and inhibited cells. The effect of locomotion is consistent with a perturbation increasing δvar(E).

upon stimulus presentation. Changes in onset latency of stimulus-selective clusters explained changes in stimulus-coding latency. We thus linked gain modulation to changes in stimulus-processing speed (Fig. 3.5, Table 3).

3.4.5 Physiological responses to perturbations

Our results show that cortical processing speed can be accelerated or slowed down via external perturbations. We found that different types of perturbations may induce similar effects on processing speed: a dynamical acceleration may be obtained by either increasing the mean or the variance of the external input to E neurons, or either decreasing the mean or the variance of external inputs to I neurons. A dynamical deceleration may be obtained by the opposite perturbations. In order to devise an experimental test of our theory to dissect the specific effects of each type of perturbation, we then examined the single-cell responses to perturbations. By combining single-cell responses with dynamical effects, we will be able to isolate the effects of each perturbation.

We characterized single-cell responses to perturbations during ongoing periods, in the absence of sensory stimuli (Fig. 3.6). For this perturbation-only condition, we introduced perturbation time course modeled as a double exponential, to estimate single-cell responses to a physiologically plausible perturbation profile in the absence of sensory stimuli. We found that perturbations differentially affected neuronal responses in a cell-type specific way. Perturbations changed the average population firing rates, and led to complex patterns of response across E and I populations (Fig. 3.6). Specifically, perturbations increasing δmean(E) induced higher firing rates and induced proportionally excited responses in both E and I populations. On the other hand, perturbations that increased δmean(I) led to a decrease in both E and I average firing rates. This paradoxical effect (Tsodyks et al., 1997) revealed that the network operates in the inhibition stabilized regime (Fig. 3.2c). When increasing the inhibitory current beyond δmean(I)=50%, the network reached a reversal point where the E population activity became silent and the I population rebounded, starting to increase their firing rates again (Fig. 3.2c).

Perturbations increasing the variance δvar(E) and δvar(I) led to surprising effects (Fig. 3.6). Increasing δvar(E) induced higher firing rates in both E and I populations, despite leaving the mean afferents unchanged; moreover, it led to mixed responses at the single cell level, with a prevalence of excited responses in both E and I populations. We will see below that this set of responses is consistent with locomotion-induced effects in the visual cortical hierarchy. Increasing δvar(I) left firing rates of I populations unchanged but led to a decrease of E population firing rates. This perturbation also induced mixed responses at the single cell level, with a prevalence of inhibited responses in both populations. Finally, perturbations δAMPA and δGABA

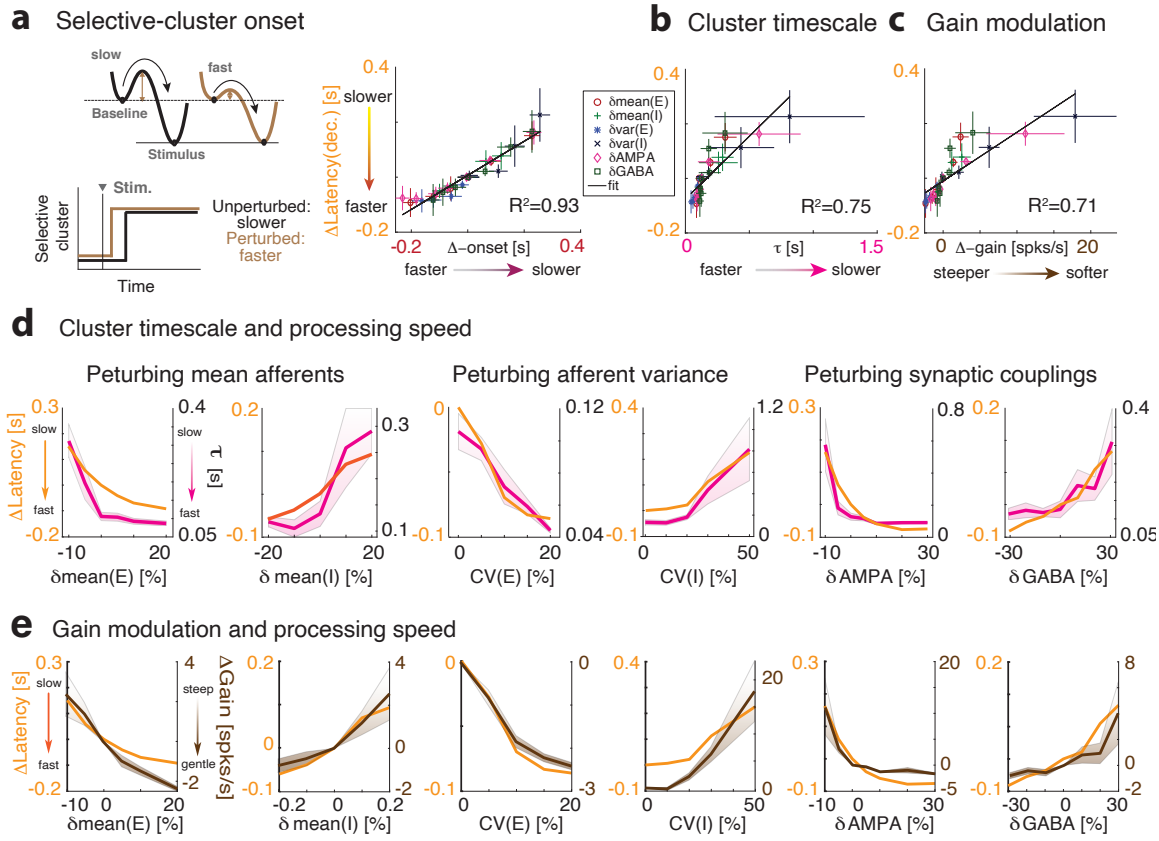

Figure 3.5. Linking gain modulation to changes in processing speed. **a)** Left: Schematic of the effect of an accelerating perturbation on stimulus encoding during evoked activity (same notations as in 3.3c). A shrinking barrier from non-selective to selective attractors drives a faster activation of stimulus-selective cluster after stimulus presentation. Right: Changes in stimulus processing speed (y-axis: latency of stimulus decoding from 3.4d-f) are predicted by changes in activation latency of stimulus selective clusters (x-axis: mean±S.D. across 10 simulated sessions; linear regression, $R^2 = 0.93$); **b, d**) by changes in cluster timescale (same values as 3.3b; $R^2 = 0.93$); **c, e**) and by changes in single-cell intrinsic gain (same values as 3.3b; $R^2 = 0.71$).

led to responses similar to those found when driving the mean E- or I-afferents, respectively.

Our theory thus suggests that it is possible to identify a specific perturbation by combining all its effects, including gain modulation, changes in stimulus-processing speed and single-cell physiological responses (Table 3) .

3.4.6 Changes in single-cell responses cannot explain the effects of perturbations on evoked activity

Although we found that gain modulation captures the effects of perturbations on network activity, we investigate whether alternative explanations were also possible, in terms of traditional measures of stimulus responsiveness and selectivity.

We found that perturbations strongly affected the peak of single-cell responses to stimuli compared to baseline (ΔPSTH, Fig. 3.6c), as well as single-cell selectivity to stimuli with significant changes in their d' (Fig. 3.6d). We then tested whether perturbation-induced changes in stimulus responses or selectivity could explain the observed changes in stimulus-processing speed. We first hypothesized that, if the response increase induced by the perturbation were larger for stimulus-selective compared to nonselective neurons (i.e., if ΔPSTH(sel)>ΔPSTH(nonsel)), then a perturbation increasing stimulus-responses could lead to faster stimulus-processing speed. Likewise, we hypothesized that faster stimulus-processing speed may be induced by perturbations improving single-cell selectivity (d') to stimuli. Surprisingly, we found a complex relation between changes in single-cell responsiveness and selectivity to stimuli, induced by the perturbations, and modulation of stimulus-processing speed (Fig. 3.6c-d). For perturbations targeting I populations (δmean(I), δvar(I), and δGABA) changes in responsiveness and selectivity were consistent with changes in processing speed ($R^2 = 0.92, 0.70$ for responsiveness and selectivity, respectively). However, for perturbations targeting E populations (δmean(E), δvar(E), and δAMPA) changes in responsiveness and selectivity were not consistent with changes in processing speed ($R^2 = 0.05, 0.02$ for responsiveness and selectivity, respectively). Strikingly, in the case of the perturbation δvar(E), processing speed increased with larger perturbations even though responses and selectivity

increasingly degraded. In the case of the perturbation δmean(E) and δAMPA, network performance likewise increased but single-cell metrics where non-monotonic in the value of the perturbation (Fig. 3.6c-d). Because changes in single-cell stimulus properties were only consistent with changes in processing speed for some perturbations (δmean(I), δvar(I), and δGABA), but inconsistent for other perturbations, we thus conclude that they could not represent an alternative mechanism underlying the observed effects of perturbations.

3.4.7 Locomotion decreases single-cell gain and accelerates visual processing speed

Our theory predicts a link between gain modulations measured during ongoing periods and changes in stimulus-processing speed during evoked periods. We sought to experimentally test this prediction in freely running mice using electrophysiological recordings from the visual hierarchy including the primary visual cortex (V1) and 4 higher cortical visual areas LM, AL, PM, AM (open-source neuropixels dataset available from the Allen Institute, Siegle et al., 2019). We interpreted periods where the animal was resting as akin to the "unperturbed" condition in our model, and periods where the animal was running as the "perturbed" condition (Fig. 3.7a in the data). We thus set out to test our theory in the following three steps: i) in each area, we estimated the effect of locomotion on single-cell responses and on intrinsic gain during ongoing periods, ii) based on these changes, we built a biologically plausible model of cortical circuits processing visual stimuli and predicted whether locomotion would accelerate or slow down visually-evoked responses; iii) we tested the prediction in each area with a decoding analysis of visually-evoked population activity.

During periods of ongoing activity (in the absence of visual stimuli), we found that locomotion induced an overall increase in firing rate across all visual cortical

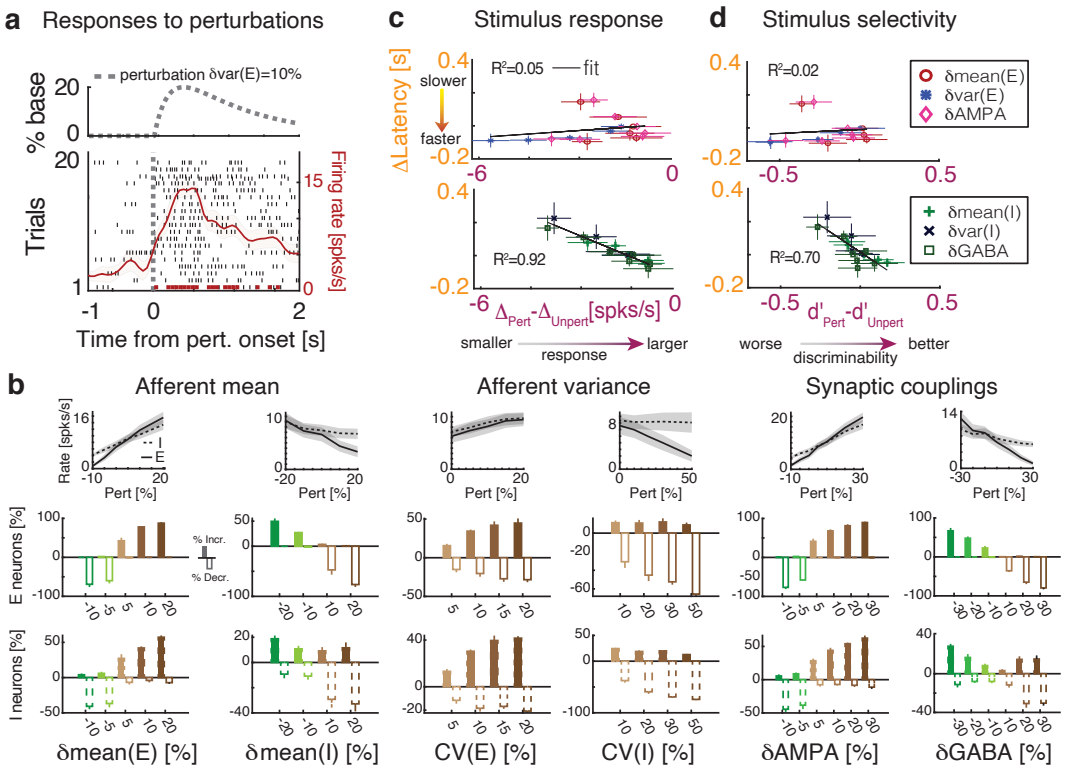

Figure 3.6. Single-cell responses to perturbations. **a)** Representative single cell response to the perturbation δvar(E)=10% in the absence of stimuli (top: dashed line, time course of perturbation, occurring at $t = 0$; bottom: dashed line, perturbation onset; red curve, response PSTH, mean±S.D. across 20 trials; horizontal red bar: significant response, t-test, $p < 0.05$ with multiple bin correction. **b)** Top: Average firing rate change across E (full) and I (dashed) populations in response to each state-changing perturbations (mean±S.D. across 10 simulated networks). Histograms: Average fractions of E (top row) and I (bottom row) neurons whose firing rate significantly increase (positive bars) or decrease (negative bars) in responses to the perturbations (single-cell significant response was based on a t-test of the baseline vs. perturbation evoked activity, $p < 0.05$). **c-d**) Single-cell changes in firing rate response to stimuli (Δ =peak response-baseline in each perturbed or unperturbed condition) as well as changes in stimulus selectivity [d'(perturbed trials)-d'(unperturbed trials)] due to the perturbations are overall uncorrelated to changes in stimulus-decoding latency (mean±s.e.m. across 5 networks).

areas (Fig. 3.7b left), in agreement with previous studies (Niell and Stryker, 2010; Stringer et al., 2019a; Dipoppa et al., 2018). Although, we found that locomotion led to complex responses inducing mixed excited and inhibited responses across neurons

(Fig. 3.7b right), as previously reported Dipoppa et al. (2018). We then estimated the single-cell transfer functions from spiking activity during ongoing periods both when the animal was at rest and in motion (Fig. 3.7c). We found that locomotion strongly modulated the single-cell gain in the absence of stimuli in all visual cortical areas (Fig. 3.7d). Specifically, we found that locomotion on average decreased the single-cell gain.

Our theory predicts that, in all visual cortical areas, the locomotion-induced increase in firing rates, the mixed excited and inhibited neural responses and the decrease in intrinsic gain are consistent with a state-changing perturbation mediated by an increase in the variance of the input currents to E neurons (δvar(E), Table 3). According to our theory, the decrease in gain leads to an acceleration of stimulus-processing speed in all visual cortical areas.

We aimed at refining the model predictions on the locomotion effects on V1 and the visual cortical hierarchy by introducing a biologically plausible stimulation protocol in our spiking network. Following experimental evidence on anatomical connectivity in the visual pathway (Zhuang et al., 2013; Miska et al., 2018; Kloc and Maffei, 2014; Khan et al., 2018), we then modeled incoming visual stimuli as a transient increase in the input currents to both E and I neurons (Fig. 3.7e). We then modeled the effect of locomotion as an external perturbation inducing an increase in the variance of the inputs to E neurons δvar(E), capturing the observed empirical effects of locomotion on ongoing periods in terms of gain decrease and mixed single-cell responses. In this model of visual processing, we found that locomotion accelerated visual processing speed during evoked period by 21 ± 9ms on average (mean$\pm$S.D. across 10 sessions, Fig. 3.7f). We thus set out to test this prediction in the empirical data.

Previous studies have observed an improvement in peak decoding performance during locomotion (Dadarlat and Stryker, 2017), but changes in decoding latency have not been investigated. To probe the speed and accuracy of visual responses in perturbed and unperturbed conditions, we performed a cross-validated classification analysis to assess the amount of information regarding the orientation of drifting

---

Figure 3.7 (next page). Locomotion effects on visual processing are mediated by gain modulation. **a)** Representative raster plots from five cortical visual areas (color-coded) with population spiking activity during passive presentation of drifting gratings (dashed red line represents stimulus onset) during periods of running (right, running speed in top panels) rest (left).**b)** Left: Firing rate modulation induced by running per area (colors) and in the model (gray, CV(E)=5%), averaged across all periods of ongoing activity. Right: Fraction of neurons by area (colors) and in the model (gray, CV(E)=5%) with significantly excited (positive bars) and inhibited (negative bars) responses to bouts of running (rank-sum test, $* = p < 0.005$). **c)**: A representative single-cell distribution of firing rates for rest (blue) and running (red) conditions. The overlaid distributions of firing rates are obtained by passing a standard normal distribution through the sigmoidal transfer function fit shown in the inset for rest (full gray line) and running (dashed gray line). The gain for each behavioral condition (orange lines) was estimated as the slope of the sigmoidal transfer function fit at the inflection point (see Methods). **d)**: Single-cell gain modulation (Δgain=gain(running)-gain(rest)) by area (colors) and in the model (gray, CV(E)=5%) across all neurons during ongoing periods (bars show 95% confidence interval; rank-sum test $* = p < 0.005$). **e)** Schematic of the network architecture, similar to Fig. 3.2a, but with the sensory stimuli targeting subsets of both E and I clusters, during both unperturbed and perturbed (δvar(E)) conditions. **f)** Average time course of the stimulus-decoding accuracy in the unperturbed (black) and perturbed (brown) conditions for the new stimulus input (notations as in Fig. 3.4c). **g)** Time course of the mean stimulus-decoding accuracy across orientations during running and rest using neurons from V1 as predictors shows the anticipation of stimulus coding in the running condition (single sessions and session average, thin and thick lines, respectively; see Methods). **h)** Decoding latency (first bin above chance decoding regions in e) slows down along the anatomical hierarchy (x-axis: anatomical hierarchy score from (Siegle et al., 2019)). Dotted (dashed) line with diamond ("x") symbols show the latency during rest (running). **i)** Difference in processing speed between running and resting (average latency of decoding accuracy between 40% and 80%, yellow area in panel e) reveals running-induced coding acceleration in all areas (colors) and in the model (gray, CV(E)=5%). t-test, $* = p < 0.01$.

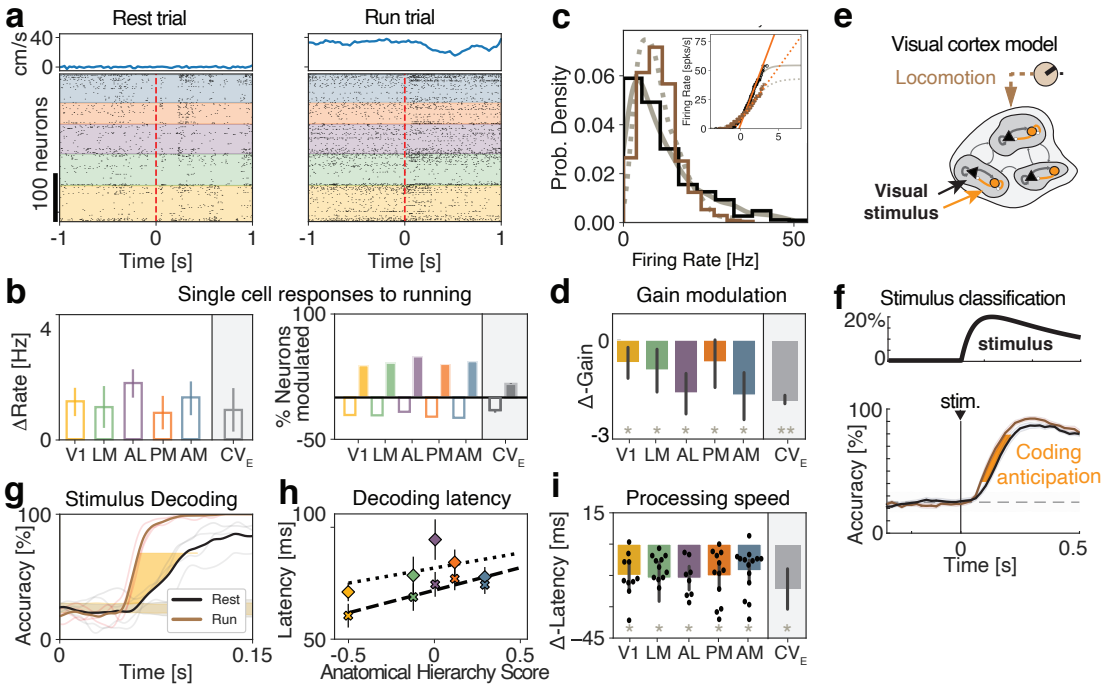

Figure 3.7.

grating stimuli present in population spiking activity along the visual cortical hierarchy. Crucially, because decoding accuracy depends on sample size, we equalized number of trials between resting and running conditions. We found that trials in which the animal was running revealed both an increase in peak decoding accuracy and an anticipation of stimulus coding (shorter latency) as compared to trials where the animal was stationary (Fig. 3.7g), consistently across the whole visual hierarchy (Fig. 3.7i). Furthermore, the time to reach significant decoding for each cortical area followed the anatomical hierarchy score in both unperturbed and perturbed conditions, consistent with the idea that information about the visual stimulus travels up a visual hierarchy in a feed-forward fashion (Fig. 3.7h, Siegle et al., 2019).

Given that locomotion induced an increase in firing rates in all cortical areas (Fig. 3.7b), we then examined the extent to which the observed effects of locomotion (increased peak accuracy and anticipation) were merely due to the increase in firing

rates. We thus matched the distribution of firing rates between running and resting (see methods and Fig. 3.8). We found that after rate matching the change in peak decoding accuracy decreased significantly (Fig. 3.8d). Crucially, the anticipation of stimulus processing speed induced by locomotion was still present in the rate-matched condition (Fig. 3.8c), confirming that it was independent of changes in firing rates. The same effect was preserved in the rate-matched model simulations as well (not shown). We thus concluded that the anticipation of visual processing speed induced by locomotion is consistent with a mechanism whereby locomotion decreases single-cell gain via an increase in the afferent variance δvar(E) as predicted by our theory (Table 3).

3.5 Discussion

Cortical circuits flexibly adapt their information processing capabilities to changes in environmental demands and internal state. Empirical evidence suggests that these state-dependent modulations may occur already in the sensory cortex where they may be induced by top-down pathways or neuromodulation. Here, we presented a mechanistic theory explaining how stimulus-processing speed can be regulated in a state-dependent manner via gain modulation, induced by transient changes in the afferent currents or in the strength of synaptic transmission.

Our theory entails a recurrent spiking network where excitatory and inhibitory neurons are arranged in clusters, generating metastable activity in the form of transient activation of subsets of clusters. We showed that gain modulation controls the timescale of metastable activity and thus the network's information-processing speed and reaction times upon stimulus presentation. Specifically, our theory predicted that perturbations that decrease (increase) the intrinsic single-cell gain during ongoing periods accelerate (slow down) the latency of stimulus responses.

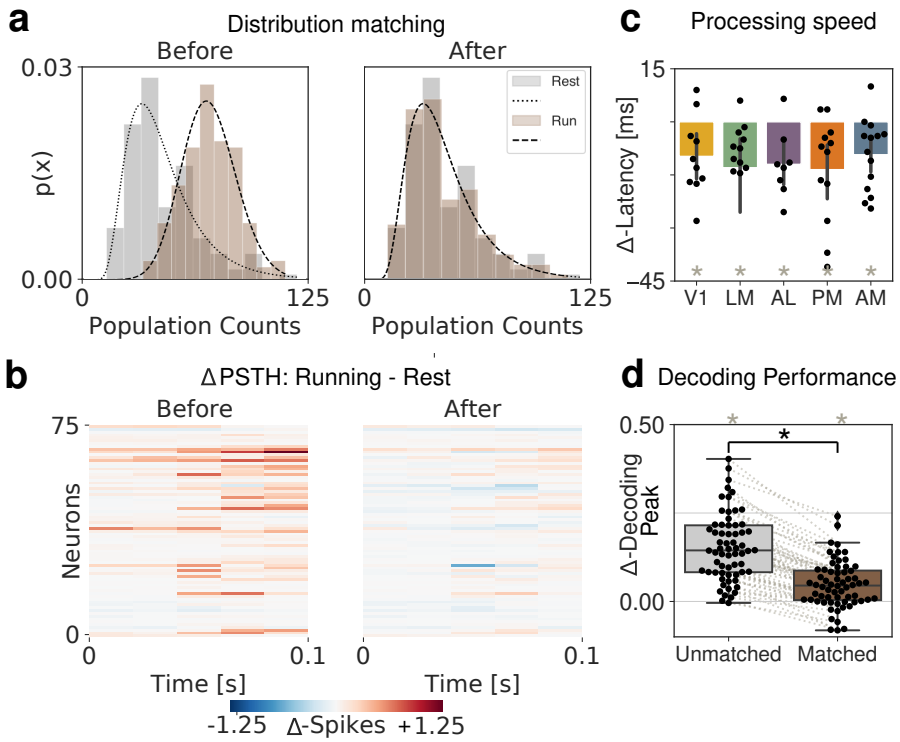

Figure 3.8. Anticipation of stimulus decoding persists even after matching the distribution of firing rates across behavioral conditions, but reduces the change in peak decoding. **a)** Firing rate distributions for both rest and running before (left) and after (right) randomly removing spikes from the running condition. Black lines show log-normal fits of distributions. **b)** Δ PSTH between behavioral conditions before and after distribution matching shows effects of match across each neuron's firing rate. **c)** Summary of changes in processing speed due to locomotion by area after distribution matching. (t-test, p$<$ 0.01) **d** The difference in peak decoding between behavioral conditions is reduced after matching the distributions (rank-sum test, gray $* = p < 0.005$). The change in Δ-Decoding peak between non-matched and matched datasets was significant (rank-sum test, black $* = p < 0.005$).

We tested this prediction by examining the effect of locomotion on visual processing in freely running mice. We found that locomotion reduced the intrinsic single-cell gain during ongoing periods, thus accelerating stimulus-coding speed across the visual cortical hierarchy. Our theory suggests that the observed effects of locomotion are consistent with a perturbation that increases the spatial variance of the afferent currents to the local excitatory population. These results establish a

new theory of state-dependent adaptation of cortical responses via gain modulation, unifying the effect of different pathways under a shared computational mechanism.

3.5.1 Metastable activity in cortical circuits

The crucial dynamical feature of our model is its metastable activity, whereby single-trial ensemble spike trains unfold through sequences of metastable states. State are long-lasting, with abrupt transitions between consecutive states. Metastable activity has been ubiquitously observed in a variety of cortical and subcortical areas, across species and tasks (Abeles et al., 1995; Jones et al., 2007; Engel et al., 2016; Ponce-Alvarez et al., 2012; Maboudi et al., 2018; Rich and Wallis, 2016; Sadacca et al., 2016; Taghia et al., 2018; Deco et al., 2019). Metastable activity can be used to predict behavior and was implicated as a neural substrate of cognitive function, such as attention (Engel et al., 2016), expectation (Mazzucato et al., 2019), and decision making (Rich and Wallis, 2016; Taghia et al., 2018; Recanatesi et al., 2020). Metastable activity was observed also during ongoing periods, in the absence of sensory stimulation, suggesting that it may be an intrinsic dynamical regime of cortical circuits (Mazzucato et al., 2015; Engel et al., 2016). Here, we showed how cortical circuits can flexibly adjust their performance and information-processing speed via modulations of their metastable dynamics.

Metastable activity may naturally arise in circuits where multiple stable states, or attractors, are destabilized by external perturbations (Miller and Katz, 2010) or intrinsically generated variability (Deco and Hugues, 2012; Litwin-Kumar and Doiron, 2012; Mazzucato et al., 2015; Schaub et al., 2015; Mazzucato et al., 2019; Rostami et al., 2020; Recanatesi et al., 2020). Biologically plausible models of metastable dynamics have been proposed in terms of recurrent spiking networks where neurons are arranged in clusters, reflecting the empirically observed assemblies of functionally

correlated neurons (Kiani et al., 2015; Song et al., 2005; Perin et al., 2011; Lee et al., 2016). Clustered network models of metastable dynamics provide a parsimonious explanation of several physiological observations such as stimulus-induced reductions of trial-to-trial variability (Deco and Hugues, 2012; Litwin-Kumar and Doiron, 2014; Mazzucato et al., 2015; Rostami et al., 2020; Churchland et al., 2010), of firing rate multistability (Mazzucato et al., 2015), and of neural dimensionality (Mazzucato et al., 2016). Compared to previous models of metastable dynamics, our results extend the biological plausibility of clustered networks in several aspects. The introduction of pairs of E and I clusters induces a tight balance where the E and I contributions to the postsynaptic currents of each neuron closely track each other with opposite signs (not shown), as observed experimentally in cortical circuits (Okun and Lampl, 2008). Moreover, we showed that these networks operate in the inhibition stabilized regime (Fig. 3.2c), which is believed to be the operational regime of cortical circuits (Ozeki et al., 2009; Sanzeni et al., 2019; Moore et al., 2018). We showed that a heterogeneous distribution of cluster sizes naturally give rise to lognormal distribution of firing rates (Fig. 3.2b inset), as observed in cortical circuits (Shafi et al., 2007; Hromádka et al., 2008; O'Connor et al., 2010; Roxin et al., 2011). We then generalize the results in Mazzucato et al. (2019) to establish gain modulation as the general mechanism controlling that state-dependent changes in processing speed in recurrent circuits with metastable dynamics. This class of models thus provide a biologically plausible, mechanistic link between connectivity, dynamics, and information-processing.

3.5.2 Linking metastable activity to flexible cognitive function via gain modulation

Recent studies have shown that cortical circuits may implement a variety of flexible cognitive computations by modulating the timescale of their intrinsic

metastable dynamics (Mattia et al., 2013; Mazzucato et al., 2019; Engel et al., 2016; Rich and Wallis, 2016; Deco et al., 2019). Our results establish a comprehensive framework to investigate the extent of this hypothesis. We propose that gain modulation is the neural mechanism underlying flexible state-dependent cortical computation. Specifically, we showed that gain modulation controls the timescale of metastable dynamics, which, in turn, determines the network's information-processing speed.

Our theoretical framework to link gain modulation to changes in potential barrier heights is based on the effective mean field theory following (Mascaro and Amit, 1999; Mattia et al., 2013; Mazzucato et al., 2019), which we used to reduce a multidimensional system to obtain an effective potential describing a single population. Although this approach is exact in the case of networks with symmetric connectivity, it represents only an approximation to the full network dynamics in the case of networks with asymmetric couplings such as the ones considered in this study Rodríguez-Sánchez et al. (2020). It would be interesting to extend our results to an exact framework by estimating the network Lyapunov function (Yan et al., 2013).

3.5.3 Alternative models of gain modulation

Previous studies have suggested gain modulation as a mechanism to sharpen single-cell tuning curves without affecting selectivity (Cardin et al., 2008; Haider and McCormick, 2009), potentially mediating alertness (Cano et al., 2006) or attention (McAdams and Maunsell, 1999; Treue and Trujillo, 1999; Rabinowitz et al., 2015). In those studies, gain modulation was defined as change in the single-neuron response function to stimuli of increasing contrast. Here, we have taken a different approach and defined gain as the slope of the intrinsic neuronal current-to-rate function during ongoing periods (i.e., in the absence of stimuli, see also Chance et al., 2002; Haider

and McCormick, 2009; Mazzucato et al., 2019), as opposed to the contrast response function. We have classified mechanisms of gain modulation which act by changing the mean or spatial variance *across neurons* of the cell-type specific afferent currents to the local cortical circuit, where we modeled afferent currents as constant biases; or by changing the recurrent couplings. The rationale for our choice was to investigate the effects on internally generated variability in a network whose dynamics were entirely deterministic. Alternatively, one could model external currents as time-dependent inputs with fast noise, such as Poisson processes or colored noise. In that case, changes in background noise due to barrages of synaptic inputs are capable of inducing gain modulation as well (Chance et al., 2002; Haider and McCormick, 2009). Previous work compared these different kinds of perturbations (Poisson noise or afferent spatial variance) in the case of the perturbation δvar(E) (Mazzucato et al., 2019), showing they may lead to similar outcomes.

3.5.4 Physiological mechanisms of gain modulation

Several different physiological pathways can modulate the gain of the intrinsic neuronal transfer function, including neuromodulation, top-down and cortico-cortical interactions. Gain modulation can also be induced artificially by means of optogenetic or pharmacological manipulations. The perturbations investigated in our model may be related to different pathways and implicated in various types of cognitive function.

3.5.4.1 Neuromodulation

Neuromodulatory pathways strongly affect sensory processing in cortical circuits by changing cell-type specific afferent currents to the circuit, in some cases controlling their dynamical regime (McGinley et al., 2015b). Our theory may be applicable to explain the effects of cholinergic and serotonergic activation on sensory cortex.

Cholinergic pathways, modulating ionic currents in pyramidal cells (McCormick, 1992), can control cortical states and mediate the effects of arousal and locomotion. Artificial stimulation of cholinergic pathways was shown to improve sensory coding in visual (Goard and Dan, 2009; Pinto et al., 2013) and barrel cortex (Eggermann et al., 2014). Cholinergic stimulation alone in the absence of sensory stimuli was shown to induce mixed responses with different neural populations increasing or decreasing their spiking activity (Goard and Dan, 2009). Our theory shows that these combined experimental observations (coding improvement and mixed firing rate changes) are consistent with a mechanism whereby cholinergic activation induces an increase in δvar(E) afferents to sensory cortex, inducing an acceleration of sensory processing (Fig. 3.7).

Activation of serotonergic pathways by stimulation of dorsal raphe serotonergic neurons or local iontophoresis was shown to transiently degrade stimulus coding in sensory cortex, decreasing responses to mechanosensory stimuli (Dugué et al., 2014) and increasing the latency of the first spike evoked by auditory stimuli (Hurley et al., 2002). Serotonergic stimulation was shown to decrease firing rates in the olfactory cortex (Lottem et al., 2016), inferior colliculus (Hurley et al., 2002), and primary visual cortex (Michaiel et al., 2019; Seillier et al., 2017). Our theory shows that these experimental observations (coding degradation and decreased firing rates) are consistent with two alternative mechanisms: either an increase in the afferent currents to I populations (i.e., δmean(I)> 0) implementing the paradoxical inhibition effect (Tsodyks et al., 1997); or a decrease in the afferents to E populations (i.e., δmean(E)$<$ 0). Future experiments could test between these two alternatives.

3.5.4.2 Top-down projections

A prominent feature of sensory cortex is the integration of feedforward and cortico-cortical feedback pathways at each stage of sensory processing (Felleman and Van, 1991). In particular, top-down projections from higher cortical areas to sensory cortex are known to modulate the speed and accuracy of sensory processing (Mazzucato et al., 2019). Our theory may explain the effects of activating several cortico-cortical pathways.

Activation of feedback axons from motor cortex (M1) to somatosensory cortex (S1) was shown to increase activity in S1 during whisking (Petreanu et al., 2012) and led to faster and more accurate responses to whisker stimulation (Zagha et al., 2013). Suppression of the same pathway induced slower S1 responses to whisking in awake mice. Our theory shows that the effect of these cortico-cortical perturbations is consistent with an increase in the mean afferent currents to E populations in S1 (i.e., the δmean(E) perturbation), leading to higher firing rates and faster processing speed.

Expectation and arousal are known to strongly modulate neural activity in sensory cortices (Salkoff et al., 2020). Expected stimuli are processed faster and more accurately than unexpected stimuli both in auditory (Jaramillo and Zador, 2010) and gustatory cortex (Samuelsen et al., 2012). Experimental evidence shows that the anticipation of sensory processing induced by expectation is mediated by top-down projections from the amygdala to the gustatory cortex (Samuelsen et al., 2012), whose activation elicits mixed excited and inhibited responses in both pyramidal and inhibitory cells in the gustatory cortex (Samuelsen et al., 2012; Vincis and Fontanini, 2016). Our model shows that, while an acceleration of stimulus processing speed may in principle be mediated by different state-changing perturbations, only the δvar(E)

perturbation is consistent with the empirically observed mixed responses. Indeed, our theory suggests that these top-down projections may operate by inducing an increase in the spatial variance of the afferent currents to the E population (δvar(E) perturbation, extending previous results in Mazzucato et al. (2019) to networks including inhibitory clusters).

In attentional tasks, distractors slow down reaction times (Grueninger and Pribram, 1969; Treisman and Gelade, 1980), a behavioral effect that may be mediated by changes in the speed and accuracy of sensory processing in cortical circuits (Desimone and Duncan, 1995). The presence of distracting stimuli within a neurons receptive field suppresses its responses to the preferred stimulus (Knierim and Van Essen, 1992). The underlying mechanism may recruit lateral inhibition onto the local cortical circuit (Reynolds and Heeger, 2009; Gilbert and Li, 2013). Our theory shows that this mechanism is consistent with a modulation of the afferents to local I populations, mediated by either an increase in δmean(I) or δvar(I). It would be interesting to discriminate between these two perturbations with future experiments.

3.5.4.3 Optogenetic and pharmacological manipulations

Our theory may shed light on the effects of manipulation experiments. Optogenetic activation (inactivation) of specific E or I cells (Arenkiel et al., 2007; Li et al., 2019) has been modeled as an increase (decrease) of the afferent currents to those cells (Ebsch and Rosenbaum, 2018; Mahrach et al., 2020; Sanzeni et al., 2019). However, protein expression may not be complete across all cells of the targeted population, and even in the case of complete expression across the targeted population, different cells may be more or less sensitive to laser stimulation. Thus the effect of optogenetic stimulation on the targeted population may then be more accurately modeled by a concurrent change in both mean and variance of the targeted

cell-type specific afferents (e.g., δmean(E) and δvar(E) for E populations; δmean(I) and δvar(I) for I populations). Recent studies showed that, while a homogeneous stimulation of all I cell types simultaneously can be captured by a model of E-I recurrently coupled neurons (as in our model), partial activation of specific inhibitory cell-types may induced more complex responses (Mahrach et al., 2020; Sanzeni et al., 2019; Li et al., 2019; Otchy et al., 2015; Phillips and Hasenstaub, 2016). We plan to revisit this issue in the future.

Our theory may also be applicable to the effects of pharmacological manipulations of different synaptic receptors. In particular, the effects of combined local injection of AMPA/kainate and NMDA receptor antagonists (agonists) may be recapitulated by a decrease (increase) in δAMPA, which correspondingly perturb the value of J_{IE}, J_{EE} couplings. Similarly, the effects of local injection of GABA receptor antagonists (agonists) may be recapitulated by a decrease (increase) in δGABA, which correspondingly perturb the value of J_{EI}, J_{II} couplings.

3.5.5 Locomotion and gain modulation

Locomotion has been shown to modulate visually evoked activity (Niell and Stryker, 2010) and is sufficient in driving activity in mouse V1 (Leinweber et al., 2017; Saleem et al., 2013). Our results were consistent with previous studies in showing that locomotion affects the activity of neurons in the visual cortical hierarchy during both ongoing and stimulus-evoked activity. We found that locomotion in the absence of sensory stimuli induces an average increase in firing rates. At the single-cell level we reported a complex mix of excited and inhibited responses in both E and I cells, also consistent with previous results (Fu et al., 2014; Dipoppa et al., 2018). Crucially, we uncovered that locomotion decreased the single-cell gain during ongoing activity across the board in the visual cortical hierarchy (Fig. 3.7d). Our theory predicted that

the observed decrease in gain would lead to an acceleration of visual processing during locomotion in cortex. This prediction was confirmed in the data (Fig. 3.7i). The acceleration of processing speed observed in cortex did not depend on the locomotion-induced changes in firing rates and was still present even after matching the firing rate distributions between running and rest conditions (Fig. 3.8). Our model of the perturbation effects induced by locomotion (increased firing rates with mixed excited and inhibited responses, and faster visual processing) suggests that the effect of locomotion may be mediated by a increase in the spatial variance of the afferent current to the E populations (δvar(E) perturbation) (Ayaz et al., 2013; Niell and Stryker, 2010; Fu et al., 2014). Concretely, gain modulation may be implemented via the combined effect of activating neuromodulatory pathways such as cholinergic (Fu et al., 2014) and noradrenergic (Polack et al., 2013) inputs.

3.6 Bridge to Chapter IV

In this chapter, we investigated the effects of perturbations on the processing capacity of neural populations - both artificial and biological. We found that periods when the animal is running led to a decrease in single-cell intrinsic gain, which we connect to the visual processing speed of drifting gratings. In chapter IV, we instead look at how mice process natural scene stimuli, in a similar experimental set-up. While the hypothesis here pertained to how the behavioral state of the animal effects sensory processing speeds, in chapter IV we investigate the hypothesis that natural images stimuli presented in a repeated, predictable sequence are encoded more strongly than stimuli presented out of sequence, unexpectedly. The theoretical modelling and predictions made in this chapter - particularly those pertaining to contextual modulations changing the network dynamics through gain modulation - directly apply to understanding the study in chapter IV.

CHAPTER IV

DIFFERENTIAL ENCODING OF TEMPORAL CONTEXT AND EXPECTATION UNDER REPRESENTATIONAL DRIFT ACROSS THE VISUAL HIERARCHY

4.1 Author contributions

Unpublished material with Hannah Choi, Nicholas Cain, Rylan Larsen, Jerome Lecoq, Marina Garrett, Luca Mazzucato. HC and MG conceived the study. NC, RL, JL helped conceptualize experiment. DW and LM designed and performed the decoding analysis. DW, LM, HC, and MG wrote the manuscript.

4.2 Introduction

Neural populations in the visual cortical hierarchy encode specific features of visual stimuli, such as orientation, spatial frequency, and direction of movement (Hubel and Wiesel, 1959, 1962; Siegle et al., 2019). Recent studies have also shown more diverse encoding capacities of the visual cortex. For example, activity in visual cortex exhibits strong modulation by changes in behavioral state (Niell and Stryker, 2010; Stringer et al., 2018; Musall et al., 2019; Salkoff et al., 2020), arousal (McGinley et al., 2015b), and attention (Ito and Gilbert, 1999). Trial and reward history have also been shown to influence visual responses (McMahon and Olson, 2007; Meyer et al., 2014; Nikolić et al., 2009; Shuler and Bear, 2006), indicating that sensory coding is influenced not only by current state but also prior experience and expectation about the future. Neurons in primary visual cortex (V1) learn short spatiotemporal sequences of stimuli upon repeated presentation (Gavornik and Bear, 2014), and enhance their activity for unexpected oddball images, as well as at the start of a novel sequence (Homann et al., 2022; Kim et al., 2019b). How temporal context

and expectation influence representations along the visual hierarchy, and beyond V1, remains unclear.

Retrosplenial cortex (RSP) is an association area that is reciprocally interconnected with V1 and is thought to be involved in memory, spatial navigation, and representing learned visual environments (Vann et al., 2009; Alexander and Nitz, 2015). Inputs from RSP to V1 show ramping activity that increases over the course of task learning (Makino and Komiyama, 2015), and may provide information about the timing of expected events to the visual cortex. The posteromedial higher visual area (PM) sits between V1 and RSP and is highly interconnected with both areas, raising the possibility that PM relays information about natural scene statistics to RSP and information about learned expectations back to V1.

Here we examine how visual responses in hierarchically connected areas V1, PM, and RSP, are modulated by temporal context and expectation. We recorded neural activity as mice viewed repeated natural image sequences, image sequence violations, and image pairs outside the repeated sequence order. We found that image decoding and selectivity were enhanced when images were presented in pairs compared to randomized order. Further, we found that information about temporal context during image sequences improved encoding of image identity, and that this effect was stronger in V1 and PM compared to RSP. While V1 and PM could robustly distinguish expected from unexpected images, population activity in RSP did not differentiate between the expected image in a sequence and a random oddball image that took its place. This suggests that RSP contains a prediction of upcoming information based on past history, while V1 and PM signal deviations from expected sequences. All of these stimulus encodings occur under the influence of representational drift on the timescale of minutes, consistent with previous reports (Deitch et al., 2021; Aitken

et al., 2021). Despite this drift, the population responses in V1 and PM, but not in RSP, generalized across drift epochs by preserving the representational geometry of the responses.

Together, these results provide evidence that temporal context and expectation are differentially represented along the visual hierarchy and beyond V1, with V1 and PM encoding image transitions and image sequence violations, and RSP encoding the expectation of what is yet to come. The predictive coding framework proposes that connections between hierarchically organized areas operate to construct a model of the environment by comparing sensory inputs with prior experience and expectations to continually update the representation of the environment. The pattern of sensory coding we observed across V1, PM, RSP, is consistent with this theory.

4.3 Results

We set out to investigate how neural responses to natural images depend on temporal context and expectation, and whether the effects of temporal context differ across the visual cortical hierarchy. We designed a stimulus protocol in which four natural images were presented in different temporal contexts (250ms stimulus with no interleaving gray screen): either in random order ('randomized control'), or in a four image sequence (ABCD, denoted 'sequence' hereafter), or in randomized pairs of images, recapitulating the transitions between images ('transition control,' with AB, BC, CD, DA, CX, XA pairs randomly interleaved). In this sequence stimulus context, ten rare, "oddball" images randomly replaced the fourth image of the set to form an unexpected sequence (ABCX; Fig. 4.1b). Contrary to previous studies investigating the effects of stimulus history on responses in visual cortex (Kim et al., 2019a), natural images were used in place of gabor patches. Each session featured

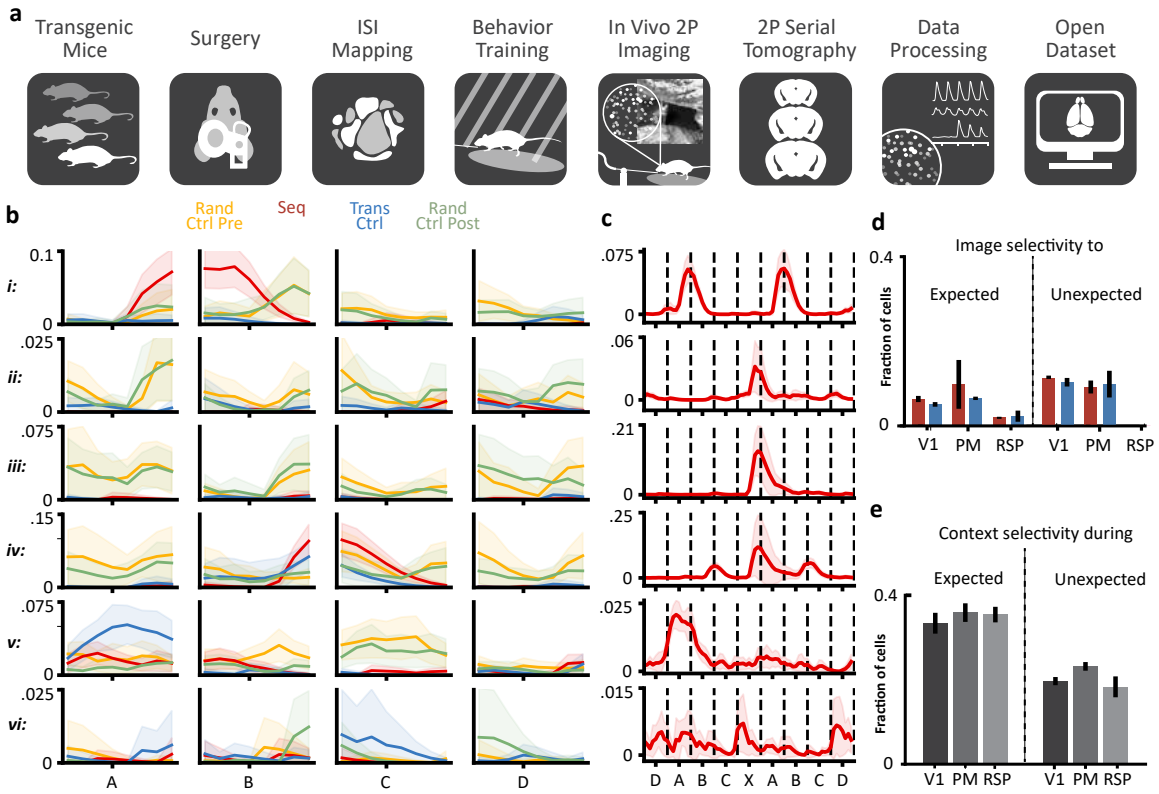

Figure 4.1. Presenting images in a variety of stimulus contexts determines single cell activity along the visual hierarchy **a)** Experimental pipeline from developing transgenic mice, to recording population responses in the visual cortex, to finally an open dataset available to the public.**b)** Mice passively view natural images in different stimulus contexts while neural activity is recorded from V1, PM, or RSP. Images are either presented in random order (blue and green) or in common, expected sequences with the occasional oddball interleaved (red), or in transition control (yellow) which only preserves pairwise transitions. **c)** Single cell PSTHs to expected images in each stimulus context. Rows i and ii: V1 neurons. Rows iii and iv: PM neurons. Rows v and vi: RSP neurons. **d)** Single cell PSTHs to unexpected images. **e)** One-way anova results. Left: A small fraction of cells in V1, PM, and RSP are selective to expected images, but only in the sequence and transition control contexts (Randomized control context not shown). Right: Selectivity to unexpected images emerges in the sequence and transition control contexts for cells in V1 and PM. **f)** Using responses to expected images (ABCD) and unexpected images (X), we can assess whether a cell is selective to the stimulus context in which the image was presented. Error bars indicate standard deviation across depths.

four blocks, where randomized control occurred both as the first and the last block in each session.

Experiments using this stimulus protocol were conducted through the Openscope program at the Allen Institute using a standardized pipeline for in vivo 2-photon calcium imaging (Fig. 4.1a). Populations of excitatory neurons were measured across multiple depths (range: 125 - 450 μm), with a total of 2299 neurons in V1, 2071 neurons in PM, and 1628 neurons in RSP. Of these neurons, we observed a variety of single cell responses (Fig. 4.1c and d). Interestingly, we see hints of context dependence to the responses to expected images in many cells. In V1 and PM, we find cells that respond strongly to the presentation of an unexpected image X (Fig. 4.1d rows ii-iv), whereas in RSP we find cells that respond similarly to X as they would the common, expected image D (Fig. 4.1d row vi). Looking at the aggregate, a small fraction were selective to expected images and unexpected images as determined by one-way anova, but only in the sequence and transition control contexts (Fig. 4.1e). Neurons in RSP were the least selective to expected images among the areas and showed no selectivity to unexpected images in any context. However, we did find a larger fraction of neurons across all 3 areas that were selective to the context in which the image was presented in (Fig. 4.1f). To understand how expectation and stimulus context is represented at a population level, we took a decoding approach, described in the sections below.

4.3.1 Transitions between images key to encoding natural scenes

First we set out to investigate how natural scenes are encoded across the visual hierarchy. We constructed classifiers to decode image identity using population responses within each stimulus context. We first focused on decoding the expected main sequence images (ABCD; Fig. 4.2a). We found significant decoding of natural image identity during the sequence and transition control stimulus contexts for all depths of V1 (Fig. 4.2b) and PM, while in RSP, only supragranular (Fig. 4.2c), but

not infragranular layers (not shown) significantly encoded the identity of expected images. The decoding performance in the sequence and transition control contexts decreased along the visual processing hierarchy (Fig. 4.2d). Surprisingly, we were not able to decode expected image identity in either of the randomized control blocks. Thus, we hypothesize that the *transitions* between consecutive images, rather than the identity of the image itself, is key to encoding expected images. Indeed, for each image presented to the animals, the preceding image is conserved in both the sequence and transition control contexts. While differences in experimental design and recording methodology are explored in the discussion, we note here that control analyses on other datasets does not lead us to interpret these findings differently.

Next, we asked whether this context-specific encoding of natural images for expected stimuli across the visual hierarchy translated to the encoding of unexpected oddball images. Again, we constructed classifiers to decode oddball identity using population responses within each stimulus context (Fig. 4.2e). V1 (Fig. 4.2f) and PM, but not RSP (Fig. 4.2g), displayed significant encoding of oddball image identity in both the sequence and transition control stimulus contexts. Again, information about image identity was not present in the randomized control contexts. Contrary to decoding expected main sequence images from each other, the preceding image for all oddball presentations is image C in the sequence. Thus, while the relative improvement from randomized control to transition control can be ascribed to the transitions between images, the subsequent improvement in decoding performance in the sequence must be something additional related to disruptions in expectation. We explore this in section 2.3 and in Figure 4.3.

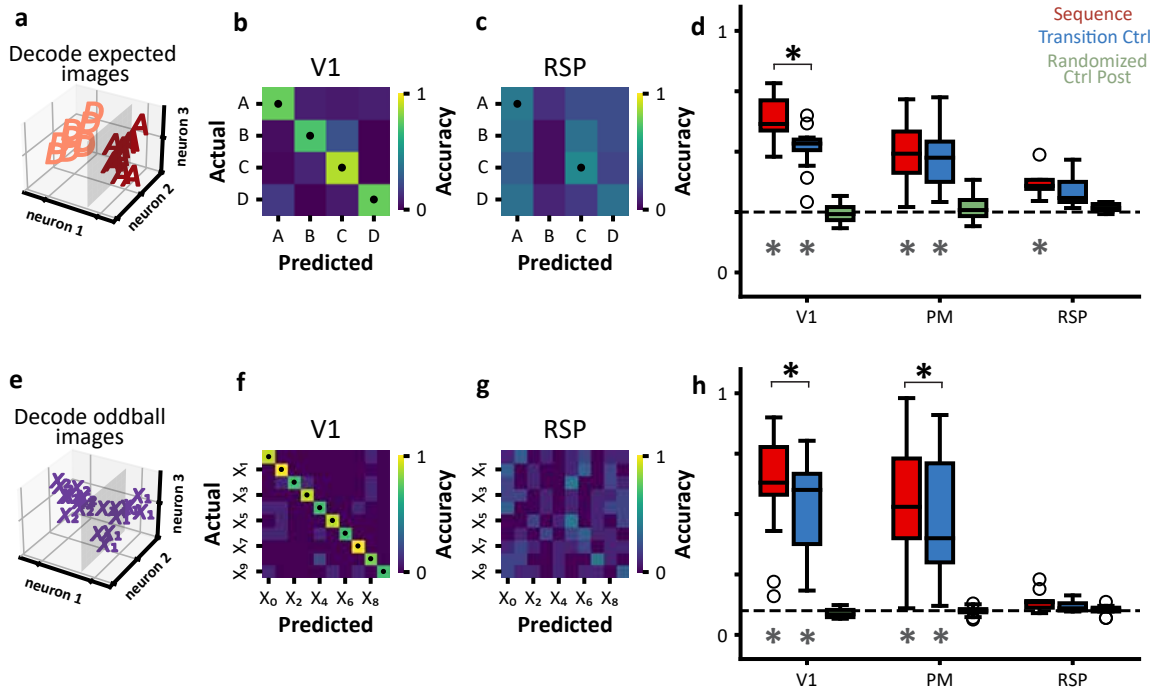

Figure 4.2. Decoding of natural images across different stimulus contexts. **a)** Schematic showing linearly separable population responses to natural images presented in the main-sequence. **b)** Example confusion matrix showing signficant decoding of main-sequence images using V1 population responses. **c)** Example confusion matrix showing signficant decoding of main-sequence images using RSP population responses. **d)** Significant decoding of expected images along the visual hierarchy in the sequence and transition control stimulus contexts. Superficial RSP has significant decoding, while neurons in deep RSP (not shown) remain at chance level. Accuracy in V1 is significantly larger in the sequence context than the transition control (Wilcoxon rank-sum test, $p < 0.05$). **e - h)** Same as a - c except for the 10 unexpected images.

4.3.2 Temporal context of sequence improves encoding of natural scenes

By presenting natural images in a variety of stimulus contexts, we were able to discern whether the embedding of natural images into 4-image sequence improved the encoding of stimulus representations. We found that the decoding performance of expected images within the sequence context was significantly higher than that of the transition control for V1, suggesting the presence of temporal effects beyond merely the transition between image pairs. In PM and RSP, we found no significant difference between contexts. On the contrary, unexpected images were more strongly

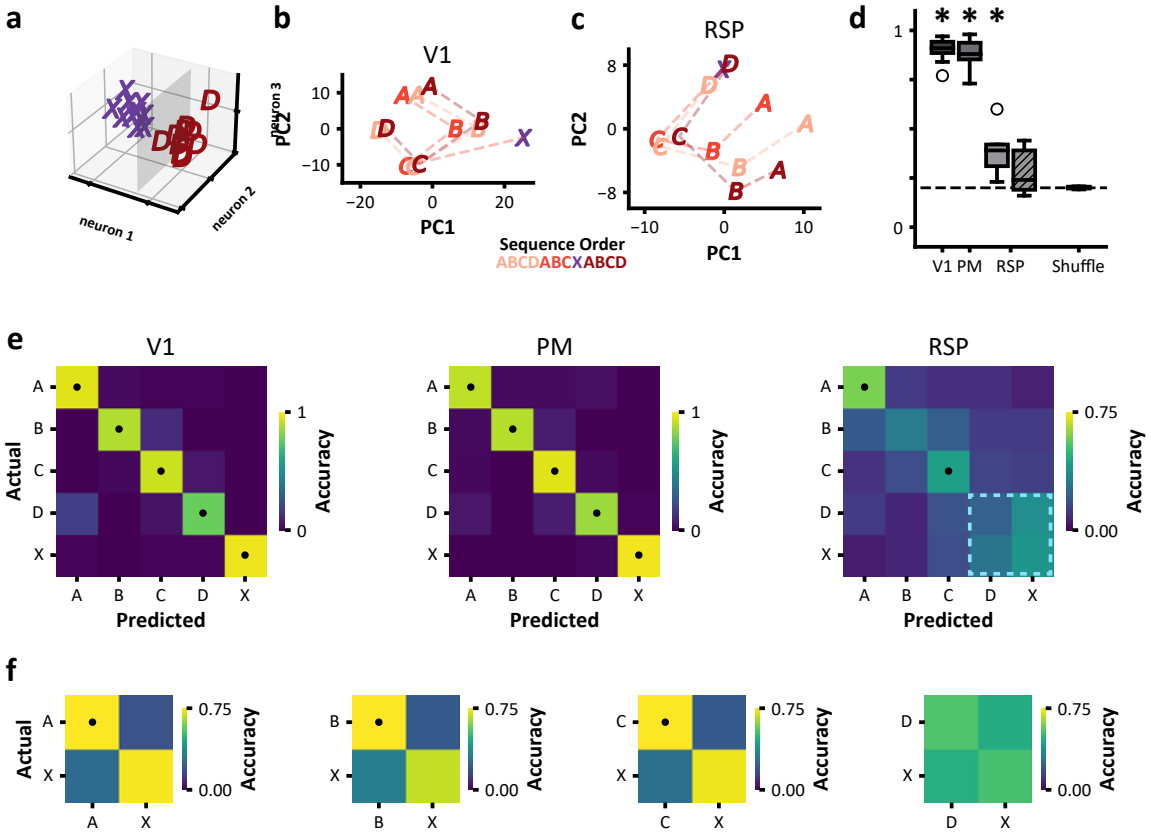

Figure 4.3. Decoding responses to expected and unexpected natural images reveals possible predictive coding mechanism in RSP. **a)** Schematic showing the task of classifier. **b)** PCA of V1 PSTHs for the sequences immediately preceding and following an unexpected event (X). Oddball representation in PCA is distinct from expected image D. **c)** PCA of RSP PSTHs for the sequences immediately preceding and following an unexpected event (X). Oddball representation in PCA space overlaps that of the expected image D. **d)** Significant decoding of main-sequence and oddball images along the visual hierarchy in the sequence. RSP separated into superficial and deep (hash pattern) layers. **e)** Example confusion matrices from V1, PM, and superficial RSP. **f)** Confusion matrices of binary classifiers between each main sequence image with the oddball images from RSP.

encoded in the sequence context compared to the transition control in both V1 and PM, highlighting again the differences in the encoding of expected and unexpected natural images.

4.3.3 Unexpected events disrupt encoding of natural scenes in RSP

In V1 and PM, single-cell image selectivity was larger for oddball responses X compared to main sequence images ABCD (Fig. 4.1d), suggesting the presence of a surprise signal; RSP cells did not exhibit significant selectivity for oddball identity. Indeed, our single cell results were confirmed by a population classification analysis. We found significant decoding of oddball images in the sequence and transition control stimulus contexts, but only in V1 and PM, not RSP, suggesting that RSP may help reinforce familiar visual environments and not novel ones (Fig. 4.2f). These results indicate that there are distinct population responses to oddball images in V1 and PM, but not RSP. We hypothesized that the lack of oddball decoding in RSP during the sequence context could be due to predictive coding effects, whereas during the presentation of ABCX, the RSP population evokes a representation of the missing D image in place of the unexpected X. After performing dimensionality reduction of visually evoked responses using Principal Component Analysis (PCA), we found that the V1 population forms a unique representation of the oddball stimulus X, that is distinct from that of the expected image D (Fig. 4.3b; Similar representations exist in PM, which is not shown). However, in RSP, the oddball representation is entirely overlapping with that of image D (Fig 4.3c), suggesting that RSP might encode for the missing image D when presented the oddball. If this was the case, we expected that the false positive rate of a decoder trained to classify X from main sequence images ABCD would be larger for D than for the other ABC images.

To investigate the origin of this difference within the visual processing hierarchy, we combined the decoding of both main sequence and oddball images within the sequence stimulus context. In order to enhance classification accuracy, we concatenated all recorded neurons into a single pseudo-population vector, which we

used as our classifier predictor. We constructed a classifier using population responses to trials that directly precede an unexpected oddball image (ABC**DABCX**; Fig. 4.3a) . The qualitative picture from the PCA analysis was confirmed by our decoding results. V1 and PM were able to discriminate between expected and unexpected images (Fig. 4.3d). In RSP, we found a more complicated picture, whereby the false positive accuracy of misclassifying oddball X as the missing D was larger than misclassification with ABC (Fig. 4.3e right). Breaking up this multi-class classifier into four binary classification tasks, we find responses to images A, B, and C can be linearly separated from responses to the oddball images, but not when we compare responses to image D with X (Fig. 4.3f). We concluded that RSP confounded the unexpected oddball image X with the missing image D. Therefore, we hypothesize that RSP instantiates a predictive coding mechanism by sending information about expected visual signals to lower levels of the visual hierarchy.

4.3.4 Contextual information represented in population activity

The differential encoding of natural images across stimulus contexts naturally led us to ask whether there is contextual information on top of the representation of natural image identity or transition identity. We first tested whether single-cell responses could discriminate whether the same image was presented in different contexts ('randomized ctrl pre','main sequence','transition ctrl','randomized ctrl post'). We found that single cells in all areas exhibited very pronounced selectivity for the temporal context, above and beyond the selectivity for image identity (Fig. 4.1d and e). Remarkably, whereas RSP cells did not encode image identity, they strongly encoded the temporal context each image was presented in.

To further quantify this, linear classifiers were constructed based on population responses to the same image in each of the four different stimulus contexts (Fig.

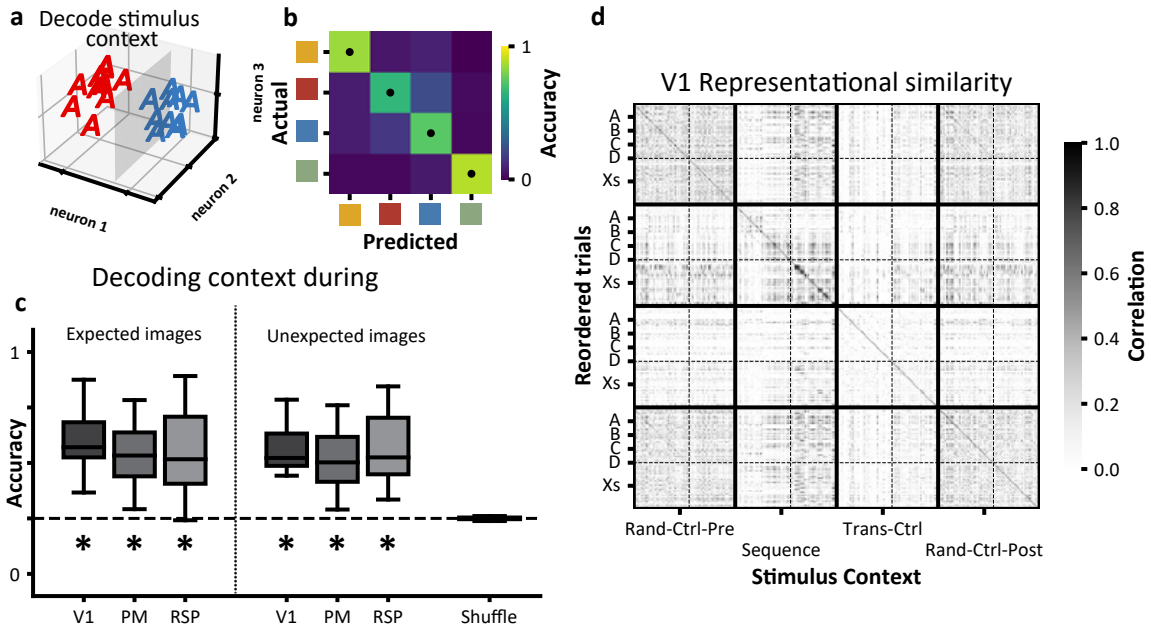

Figure 4.4. Decoding of stimulus contexts using the population responses to the same natural image. **a**) Schematic showing linearly separable population responses to the same natural image presented in different stimulus contexts. **b**) Example confusion matrix showing signficant decoding of stimulus context using V1 population responses. Blue: Randomized Control Pre. Red: Sequence. Yellow: Transition Control. Green: Randomized Control Post. **c**) Decoding results across visual areas separated by the images used for decoding. **d**) Representational similarity matrix for V1 averaged across animals shows correlation structure within stimulus contexts, but not between contexts.

4.4a). Using population responses to each of the main-sequence and oddball images, we were able to discriminate the stimulus context in which the image was presented in. This contextual information was present across layers and depths recorded and was encoded equally well during both main sequence (expected) vs oddball (unexpected) images. Interestingly, despite similar results in decoding natural images (Fig. 4.2), sequence and transition control stimulus contexts were separable from each other (Fig. 4.4c). This suggests that the representation formed in the sequence context depends more than just the pairwise transitions between image presentations. This is reflected in the example representational similarity matrix for V1, where there are clear off-

diagonal correlation structure within the ABCD trials of the sequence context, but not for the transition control context (Fig. 4.4d).

4.3.5 Representational drift and within-sequence generalization

What is the origin of the strong encoding for temporal context? Because randomized control pre and post conditions were linearly separable from each other, we hypothesized that temporal context may be partially due to representational drift occurring over the course of the recording session (Aitken et al., 2021). While definitions of representational drift usually pertain to changes in representations over the course of days and weeks (Aitken et al., 2021; Schoonover et al., 2021; Marks and Goard, 2021), we investigated whether drift was evident over one 33 minute session (Deitch et al., 2021). We focused on representations of expected images and tested information about the passing of time within the sequence context. We observed significant decoding of the epoch in which an image was presented, which we interpret as evidence of representational drift (Fig. 4.5c). Time encoding was equally present in all areas.

Next, we investigated the evolution of the representational geometry of expected natural images under drift. Two alternative scenarios may arise: In the first scenario, the drift may occur in directions orthogonal to the stimulus-encoding ones, leading to a stable representation of image identity which can generalize across different epochs. In the second scenario, the drift direction may overlap with the stimulus-coding axis, thus leading to epoch-specific representations which cannot generalize across epochs. We compared these two scenarios in PCA space, where we found that the representational geometry of the main sequence, defined as the relative position of responses to the ABCD images, is preserved across epochs in V1 and PM, but not in RSP (Fig. 4.6c,f). This result is evident in the representational similarity plots of V1

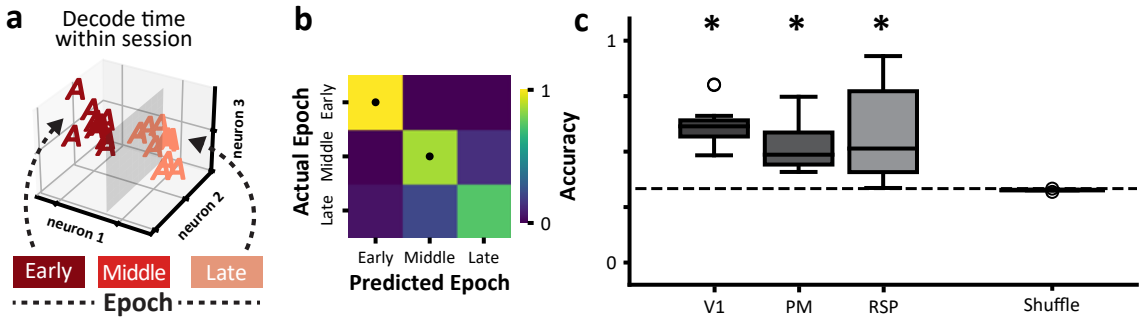

Figure 4.5. Representational drift revealed from decoding the epoch in which an image was presented in. **a**) Schematic showing population responses to the same image at different epochs within the sequence stimulus context. **b**) Example confusion matrix showing signficant decoding of time using V1 population responses. **v**) Decoding results across visual areas

vs RSP, where large correlations on the off-diagonal blocks in V1 imply generalization across epochs, while the weak correlation in off-diagonal blocks in RSP suggests lack of generalization across epochs (Fig. 4.6d,g).

To quantify the difference between two scenarios we trained a linear classifier on trials from one epoch and test its generalization performance on test trials from a different epoch. We found that linearly separable stimulus representations in V1 and PM, but not for RSP, generalized across epochs (Fig. 4.6h). We call this within-sequence generalization performance, after Bernardi et al. (2020) work on abstraction in monkey hippocampus and prefontal cortex. These results show that visual representations in V1 and PM, but not RSP, maintain a consistent linear readout available to downstream neurons, despite representational drift occurring on the timescale of minutes. Previous studies have reported this persistence in decoding performance in V1 (Aitken et al., 2021).

4.4 Discussion

The representation of naturalistic visual stimuli in cortex depends on the context in which the image was presented. We found that natural images were

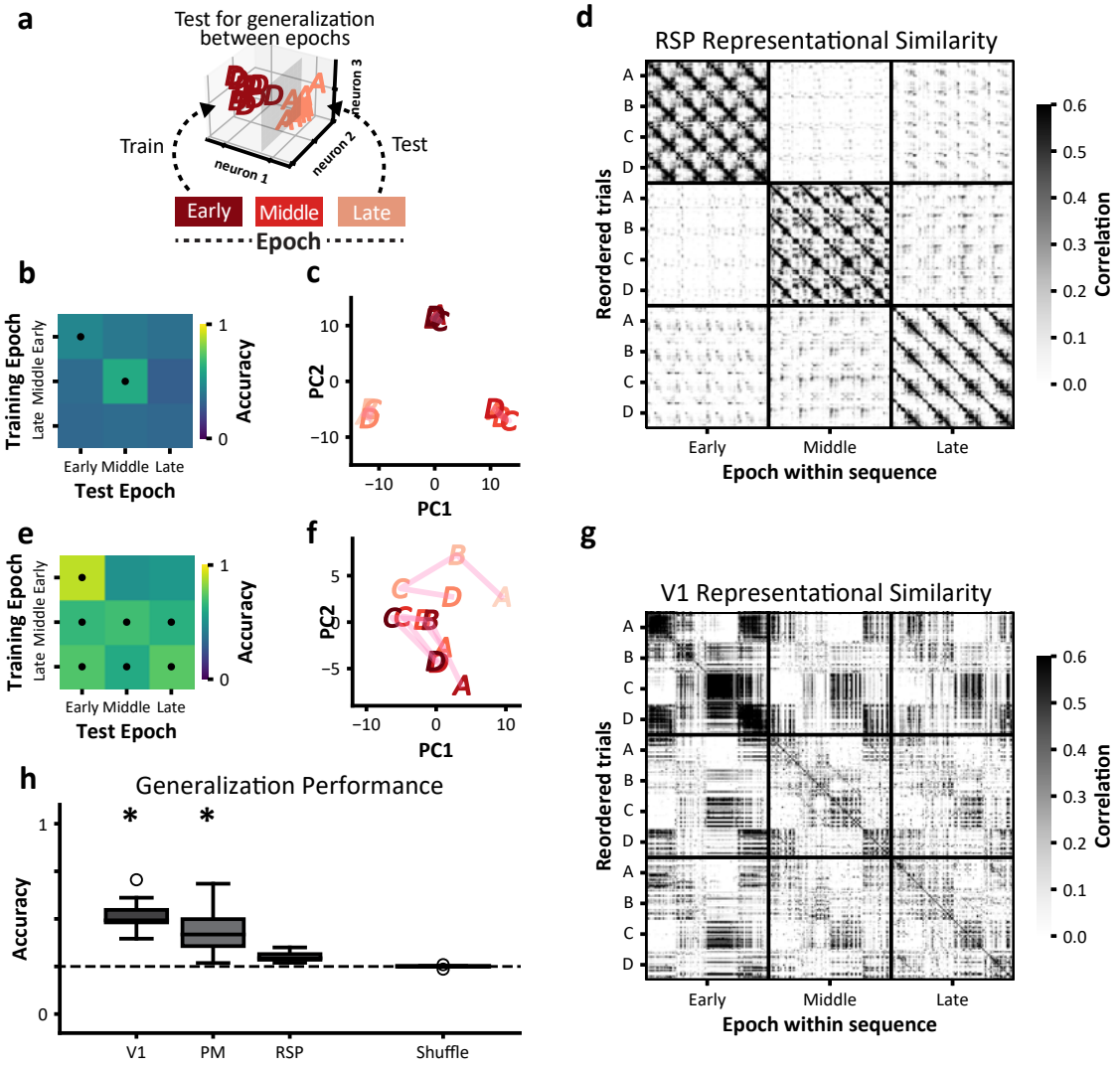

Figure 4.6. Generalization performance under representational drift. **a)** Schematic showing population responses to main sequence images in different epochs of the session, with training and testing occuring using different epochs of data. **b)** Confusion matrix from RSP shows that only within-epoch decoding is possible (main diagonal). **c)** PCA of main sequence PSTHs from RSP show epochs cluster apart. **d)** Representational similarity from RSP reveals a strong correlation of within-epoch population responses, with little to no correlation between-epoch blocks. **e)** Confusion matrix from V1 shows that both within-epoch decoding (main diagonal) and between-epoch decoding (off-diagonal) is possible. **f)** PCA of main sequence PSTHs from V1 shows conservation of geometry between epochs. **g)** Representational similarity from V1 reveals reveals a strong correlation of within-epoch population responses, with comparable correlations for between-epoch blocks. **h)** Generalization decoding results over single sessions.

encoded only when they were presented in a four set sequence or pairwise sequence (Fig. 4.2). Thus, we conclude that the transition between successive images is an essential part of the encoding of naturalistic stimuli in 2-photon datasets. While this may seem unsurprising on the surface, we found no studies in the literature that report the context dependence of naturalistic stimuli encoding. Past studies at the Allen Institute itself have only reported responses to naturalistic images that are presented in random order (de Vries et al., 2020; Siegle et al., 2021), while other studies have reported responses to naturalistic images that are presented with an interleaving gray screen (Kowalewski et al., 2021), an altogether different stimulus paradigm which changes stimulus response dynamics. Previous studies suggested that only the preceding image matters for distinguishing between sequential image presentations (Nikolić et al., 2009). However, we found a significant increase in decoding performance of expected images in the four-image sequence vs. the two-image transition control contexts in V1, highlighting the potential role of longer sequences in enhancing encoding of visual stimuli.

We found that the stimulus context in which an image was presented in (randomized control pre and post vs sequence vs transition control) was encoded more strongly than image identity in the neural populations of V1, PM, and even RSP (Fig. 4.1d, Fig. 4.4), the latter having similar context-dependent decoding accuracy as the other areas, despite much lower decoding performance of expected images chance-level decoding performance to unexpected images. One interpretation of this finding is that of representational drift (Deitch et al., 2021), which we observed in our data as well. The difference in visual representations across stimulus contexts may include other effects beyond representational drift. In particular, different contexts differ in how many images a sequence comprised (1 image in the randomized control, 2 images in

the transition control, and 4 in the main sequence). History effects extending beyond pairwise transitions may thus explain the discriminability of sequence vs transition control. This was observed in sequences of randomized gabor patches (Kim et al., 2019a), but not for extended sequences of natural images.

Our results on representational drift are consistent with previous studies in V1 and PM (Deitch et al., 2021; Marks and Goard, 2021; Aitken et al., 2021), and extend them to RSP. Our measure of representational drift - decoding - allows us to simultaneously observe whether drift is occurring, but it also allows us to test for generalization in the population response. This idea is fundamentally linked to the problem of representational drift: How do neural populations faithfully encode for sensory representations, memories, etc while being subject to persistent drift? We found that representations of expected natural images in V1 and PM generalize despite drift, but while RSP does not.

We also investigated the differences in decoding performances of expected vs unexpected natural scenes along the visual hierarchy. First, we determined that representations of individual unexpected oddball images could be linearly separated from each other in both the sequence and transition control contexts for areas V1 and PM (Fig. 4.2h). Moreover, single cell selectivity for oddball images was stronger than for the main sequence images ABCD. Furthermore, unexpected stimuli were more strongly encoded in the sequence context than the transition control context for both V1 and PM. We interpret this finding within the paradigm of predictive coding, following recent experimental evidence Gillon et al. (2021). In the sequence context, violations of the expected image D cause prediction error signals that are not present in the transition control context, which shuffles the expected sequence of images into

a random set of pairwise images. No expectation can be formed in the transition control beyond the pairwise transition.

Combining expected and unexpected stimuli paradigms into one classification task (Fig. 4.3), we found that V1 and PM had no problem differentiating expected (main-sequence) vs unexpected (oddball) stimuli. In RSP, population responses to the unexpected oddball images are confounded with that of the expected image D (Fig. 4.3c). One interpretation of this finding is that RSP encodes for the expectation of the next image, rather than what was actually shown. Indeed, we found RSP does not encode for distinct oddball representations at all (Fig. 4.2g). One possible limitation of this interpretation is that RSP may simply have a low signal to noise ratio when it comes to encoding naturalistic stimuli. However, we do report above chance decoding performance of expected images in RSP; thus, we believe expectation may play a role in this area.

4.5 Methods

4.5.1 Experimental data

All experiments and procedures were performed in accordance with protocols approved by the Allen Institute Animal Care and Use Committee. The dataset used in this paper was collected as part of the Allen Institute for Brain Science's OpenScope initiative. Data were collected and processed using the Allen Brain Observatory data collection and processing pipelines (de Vries et al., 2020). Here we include a brief description on experimental procedure. The full details on the data collection and process are described in (de Vries et al., 2020).

Transgenic mice expressing GCaMP6f in excitatory cells were used (Slc17a7-IRES2-Cre x CaMk2-tTA x Ai93(GCaMP6f). Only sessions that pass quality control criteria described in (de Vries et al., 2020) were included in our analyses, resulting

in total N = 14 mice, with 16 sessions from V1 (2299 neurons), 23 sessions from PM (2071 neurons), and 12 sessions from RSP (1628 neurons). Two-photon calcium imaging was performed a Nikon A1R MP+, with imaging depths ranging 125 - 450μm to capture neuronal activities across cortical layers.

Mice were injected with a retrograde tracer (AAVRetro.CAG.mRuby3) in either V1 or RSP, however this data was not used in our study. No differences were observed in labeled vs unlabeled cells for any of the quantifications in this manuscript. Information on the identity of retrogradely labeled cells is available at: XXXX [put .csv of cell IDs and labels on GitHub].

Each mouse experienced 4 imaging sessions, each with the same stimulus protocol. Two sessions were recorded in one cortical area (ex: V1, PM, or RSP) at 2 cortical depths, typically around 175μm (approximately layer 2/3) and 375μm (approximately layer 5), and two sessions were recorded at similar depths in a different cortical area. Areas were chosen such that the injection site for retrograde labeling was never imaged. For the full details on on animal surgery, habituation, quality control, data collection, and post-collection data processing, please see (de Vries et al., 2020).

The dataset along with the code for analyses included in this paper is available on GitHub, and the full dataset is publicly available in Neurodata Without Borders (NWB) format in the DANDI Archive

4.5.2 Stimulus protocol

Visual stimuli consisted of a subset of the natural images publicly available Allen Brain Observatory dataset (`https://observatory.brain-map.org/visualcoding/`;(de Vries et al., 2020)). The images were presented in grayscale, contrast normalized, matched to have equal mean luminance, and resized to 1,174 × 918 pixels. Four natural images (Brain Observatory image IDs: im013, im026,

im068, im078) were used to form a familiar sequence of four images (ABCD), and ten additional images served as unexpected oddball images (im06, im017, im022, im051, im071, im089, im103, im110, im111, im112).

Stimuli were presented in 4 distinct blocks over the course of a 1 hour imaging session (64 minutes). Individual stimuli were presented for 250ms with no intervening gray period (i.e. one image after the other) in all blocks. Blocks were separated by 60 seconds of gray screen in which spontaneous activity could be measured. A schematic of the stimulus design is shown in Supplemental Figure XXX.

In the Randomized Control blocks, the 14 images (4 sequence images and 10 oddball images) were presented in random order. Each image was shown for 250ms with no intervening gray screen. Each image was presented 30 times. The randomized control stimulus block was presented once at the beginning of the session and once at the end of the session and lasted 0.25s x 30 repeats = 105 seconds each time.

In the Sequence block, the series of expected sequence images ABCD was repeated 20 times per cycle, with an oddball image randomly taking the place of image D after 10-19 repeats of the sequence in that cycle. The number of sequence repeats between oddball image occurrences was random and not predictable. Each oddball image was presented a total of 10 times throughout the sequence block, for a total of 100 cycles (10 images x 10 cycles each). Each main sequence image ABCD was presented 20 times per cycle x 100 cycles = 2000 times. The entire block lasted 33.33 minutes.

In the Transition Control block, the image transitions shown in the Sequence block were preserved as pairs (AB, BC, CD, DA, CX, XA, with X = 10 oddball images, giving a total of 24 pairs), but the global sequence was not preserved. Each image transition pair was treated as a distinct stimulus (lasting 0.5 seconds) and

presented in random order. Each image pair was shown 30 times, for a total of 24 pairs x 30 repeats = 720 pair presentations. The Transition Control block was 6 minutes in duration (720 pair presentations x 0.5 seconds per pair = 360 seconds).

Two additional stimuli not used in this study were also shown. The occlusion stimulus, consisted of the 10 oddball images with 6 differing levels of spatial occlusion. Each occlusion image was presented for 0.5 seconds, with 0.5 seconds of gray screen between stimuli. Each occlusion image was presented 10 times, for a total of 10 images x 6 occlusion levels x 10 repeats = 600 individual stimulus presentations, for a total of 600 seconds = 6 minutes. A 30 second natural movie clip (Brain Observatory stimulus set Natural Movie 1) was repeated 10 times at the end of the session.

4.5.3 Decoding analysis

We decoded stimulus context, image identity, and time within a session from single-trial population response vectors. To construct these response vectors, we performed deconvolution (Jewell et al., 2019; de Vries et al., 2020) on the delta fluorescence traces and summed the deconvolved events within a circumscribed response window of 50ms to 250ms relative to stimulus onset. Results presented within this paper remain unchanged if instead we perform the mean over this response window. Specifically, we trained multi-class linear support vector machine (SVM) classifiers to decode the above categorical variables from the single-trial population response vectors. Stratified 5-fold cross-validation was used for decoding expected images (Fig. 4.2d), context(Fig. 4.4), and "time" (Fig. 4.5), ensuring equal number of trials for each class. When performing classification with oddball images (Fig. 4.2h and Fig. 4.3), 10-fold stratified cross-validation was performed, as each oddball was only presented 10 times during the sequence stimulus context.

To calculate the statistical significance of decoding accuracies, we performed an iterative shuffle procedure on each fold of the cross-validation. In each shuffle, the training labels which the classifer was trained to decode were shuffled randomly across trials of the training set, and the classifer's accuracy was evaluated on the unshuffled test data-set. This shuffle was performed 100 times to create a shuffle distribution of decoding accuracies for each fold of the cross-validation. From these distributions we calculated the z-score of decoding accuracy for each class in each cross-validation fold. These z-scores were then averaged across the folds of cross-validation and used to calculate the overall p-value of the decoding accuracy obtained on the original data.

4.5.4 Validation

To examine issues of stimulus design or recording methodology, we performed the same decoding analysis on three independent datasets where natural images were presented in random order. First we compared decoding performances evaluated on data from the publicly available Allen Brain Observatory Visual Coding ("ophys") and Neuropixels ("ephys") datasets (de Vries et al., 2020; Siegle et al., 2019). Both experiments consisted of 118 natural image stimuli from the same dataset as ours with identical presentation protocol. We confirmed significant decoding accuracy across all 118 images using electrophysiological population responses reliably across experiments (Fig. 4.7g). On the other hand, using population responses extracted from 2-photon imaging data severely reduced the reliability and magnitude of decoding performance to chance level (Fig. 4.7h). Restricting ourselves to the 14 images presented in our experiment (4 main-sequence and 10 oddball images) increased decoding performance to slightly above chance level for 2-photon data. In another study where natural images were presented for 1s, with an interleaving gray screen of 2s, V1 populations

recorded using 2-photon imaging showed significant decoding (Kowalewski et al., 2021).

From this comparative set of analyses, we concluded that the lack of significant decoding of natural images during the randomized blocks may be due to weaker and less reliable visually-evoked responses than those collected previous studies, due both to the nature of our 2p recordings (less reliable compared to electrophysiological ones) and to particular protocol used (flashed presentation consecutive images, compared to long image presentations interleaved by black screen. Our results further suggest that, in our 2p setup with flashed images, the information about visual stimuli encoded in neural responses is likely related to the identity of *transitions* between consecutive images, rather than the identity of single images. Taken together, these results highlight the particularities of stimulus design and calcium dynamics as it pertains to natural stimulus encoding.

4.5.5 Single-cell selectivity

To assess single cell selectivity to particular natural images and stimulus context, we performed simple one-way anovas using the open source python software pingouin (Vallat, 2018). For each stimulus context, we determined if a cell was selective to one of the four main-sequence images or to one of the 10 oddball images by using the population response vectors for those trials in which the images were presented (Fig. 4.1d left, middle). We ensured each group had an equal number of trials. For each image, we determined if a cell was selective to the context in which it was presented in by using population response vectors for the trials in which the same image was presented in different contexts (Fig. 4.1d right). P-values were calculated from the F-distribution and were considered significant (selective) if below a threshold of 0.05.

4.6 Supplemental Figures

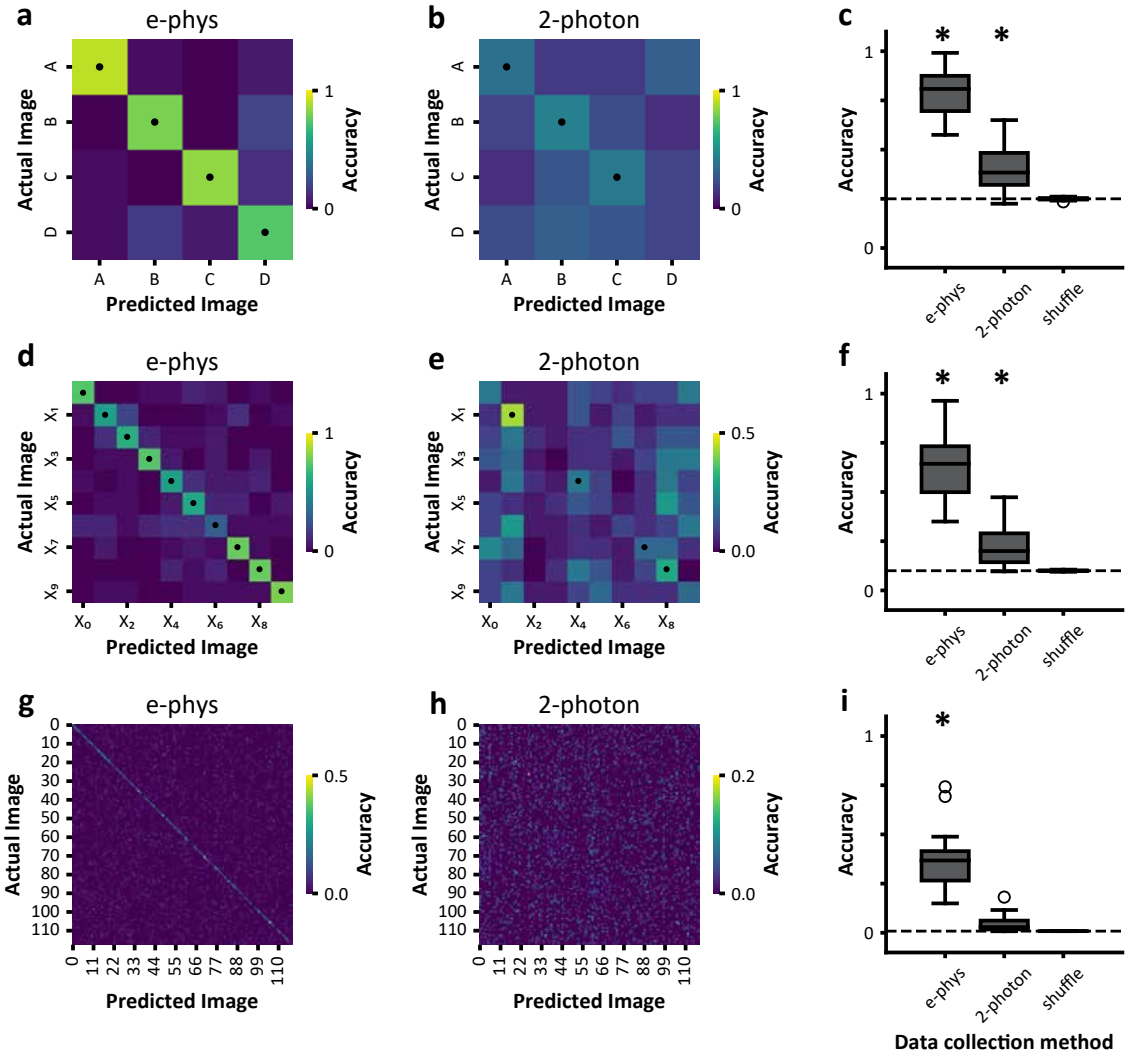

Figure 4.7. Validation of decoding results with field standard electrophysiological and two-photon functional datasets. **a)** Example confusion matrix from one session of e-phys data trained to classify the four main-sequence images within our study. **b)** Example confusion matrix from one session of two-photon data trained to classify the four main-sequence images within our study. **c)** Summary boxplot over 32 sessions of e-phys and 39 sessions of 2-photon data. **d - f)** Same as a-c, but for the 10 oddball images. **g - i)** Same as a-c, but for all 118 images within the Berkeley Segmentation Dataset

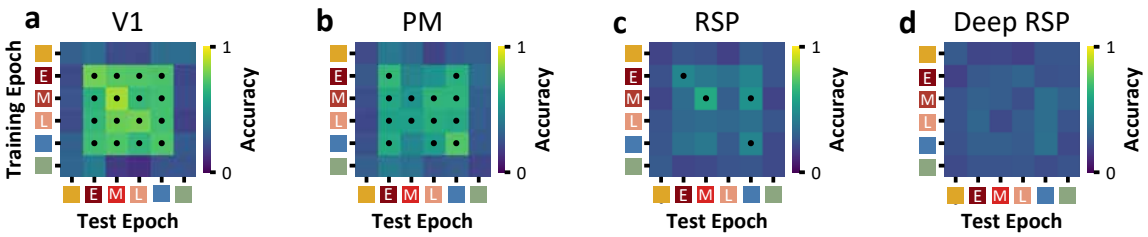

Figure 4.8. Generalization of main sequence image representations extends to transition control context. **a-d)** Example confusion matrices showing the generalization performance across epochs. Each element summarizes the main decoding results from that training/test set pair.

CHAPTER V

REFERENCES CITED

M. Abeles, H. Bergman, I. Gat, I. Meilijson, E. Seidemann, N. Tishby, and E. Vaadia. Cortical activity flips among quasi-stationary states. *Proc Natl Acad Sci USA*, 92: 8616–8620, 1995.

Ehud Ahissar and Eldad Assa. Perception as a closed-loop convergence process. *eLife*, 5(May 2016):1–26, 2016. ISSN 2050084X. doi: 10.7554/eLife.12830.

Kyle Aitken, Marina Garrett, Shawn Olsen, and Stefan Mihalas. The Geometry of Representational Drift in Natural and Artificial Neural Networks. *bioRxiv*, page 2021.12.13.472494, 2021. URL `https://www.biorxiv.org/content/10.1101/2021.12.13.472494v2`.

Josef Albers. Tautonym, B, 1944.

Andrew S. Alexander and Douglas A. Nitz. Retrosplenial cortex maps the conjunction of internal and external spaces. *Nature Neuroscience*, 18(8):1143–1151, 2015. ISSN 15461726. doi: 10.1038/nn.4058.

Efrén Álvarez-Salvado, Angela M Licata, Erin G. Connor, Margaret K. McHugh, Benjamin M.N. Mn King, Nicholas Stavropoulos, John P. Crimaldi, Katherine I. Nagel, Jonathan D Victor, John P. Crimaldi, and Katherine I. Nagel. Elementary sensory-motor transformations underlying olfactory navigation in walking fruit-flies. *bioRxiv*, 7:e04577, 2018. ISSN 2050-084X. doi: 10.1101/307660.

D. J. Amit and N. Brunel. Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb Cortex*, 7(3):237–52, 1997. URL `http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9143444`.

R.A. Andersen and V.B. Mountcastle. The influence of the angle of gaze upon the excitability of the light-sensitive neurons of the posterior parietal cortex. *Journal of Neuroscience*, 3(3):532–548, 1983. ISSN 02706474. doi: 10.1523/jneurosci.03-03-00532.1983.

Benjamin R Arenkiel, Joao Peca, Ian G Davison, Catia Feliciano, Karl Deisseroth, George J Augustine, Michael D Ehlers, and Guoping Feng. In vivo light-induced activation of neural circuitry in transgenic mice expressing channelrhodopsin-2. *Neuron*, 54(2):205–218, 2007.

Aslı Ayaz, Aman B Saleem, Marieke L Schölvinck, and Matteo Carandini. Locomotion controls spatial integration in mouse visual cortex. *Current Biology*, 23(10):890–894, 2013.

K.L. Baker, M. Dickinson, T.M. Findley, D.H. Gire, M. Louis, M.P. Suver, J.V. Verhagen, K.I. Nagel, and Matthew Smear. Algorithms for Olfactory Search across Species. *The Journal of Neuroscience*, 38(44):9383–9389, 2018. ISSN 15292401. doi: 10.1523/JNEUROSCI.1668-18.2018.

Cornelia I Bargmann. Comparative chemosensation from receptors to ecology. *Nature*, 444(7117):295–301, 2006. doi: https://doi.org/10.1038/nature05402.

Howard C Berg. Motile behavior of bacteria. *Physics Today*, 53(1):24, 2000. ISSN 00319228. doi: 10.1063/1.882934.

Gordon J Berman, William Bialek, and Joshua W Shaevitz. Predictability and hierarchy in drosophila behavior. *Proceedings of the National Academy of Sciences*, 113(42):11943–11948, 2016.

Silvia Bernardi, Marcus K. Benna, Mattia Rigotti, Jérôme Munuera, Stefano Fusi, and C. Daniel Salzman. The Geometry of Abstraction in the Hippocampus and Prefrontal Cortex. *Cell*, 183(4):954–967.e21, 2020. ISSN 10974172. doi: 10.1016/j.cell.2020.09.031.

Jennifer Beshel, Nancy Kopell, and Leslie M Kay. Olfactory Bulb Gamma Oscillations Are Enhanced with Task Demands. *Journal of Neuroscience*, 27(31):8358–8365, 2007. doi: 10.1523/JNEUROSCI.1199-07.2007.

U. Bhattacharyya and U. Singh Bhalla. Robust and Rapid Air-Borne Odor Tracking without Casting. *eNeuro*, 2(6):1–26, 2015. ISSN 2373-2822. doi: 10.1523/ENEURO.0102-15.2015.

Shuangyu Bi and Victor Sourjik. Stimulus sensing and signal processing in bacterial chemotaxis. *Current opinion in microbiology*, 45:22–29, 2018. doi: 10.1016/j.mib.2018.02.002.

Kevin A. Bolding and Kevin M. Franks. Complementary codes for odor identity and intensity in olfactory cortex. *eLife*, page e22630, 2017. ISSN 2050084X. doi: 10.7554/eLife.22630.

Monica Cano, Tatiana Bezdudnaya, Harvey A Swadlow, and Jose-Manuel Alonso. Brain state and contrast sensitivity in the awake visual thalamus. *Nature neuroscience*, 9(10):1240–1242, 2006.

Jessica A Cardin, Larry A Palmer, and Diego Contreras. Cellular mechanisms underlying stimulus-dependent gain modulation in primary visual cortex neurons in vivo. *Neuron*, 59(1):150–160, 2008.

Kenneth C. Catania. Stereo and serial sniffing guide navigation to an odour source in a mammal. *Nature Communications*, 4:1441–1448, 2013. ISSN 20411723. doi: 10.1038/ncomms2444.

Frances S Chance, Larry F Abbott, and Alex D Reyes. Gain modulation from background synaptic input. *Neuron*, 35(4):773–782, 2002.

Tsai-Wen Chen, Nuo Li, Kayvon Daie, and Karel Svoboda. A map of anticipatory activity in mouse motor cortex. *Neuron*, 94(4):866–879, 2017.

Nathan Childress. bluewhitered, 2020.

M. M. Churchland, B. M. Yu, J. P. Cunningham, L. P. Sugrue, M. R. Cohen, G. S. Corrado, W. T. Newsome, A. M. Clark, P. Hosseini, B. B. Scott, D. C. Bradley, M. A. Smith, A. Kohn, J. A. Movshon, K. M. Armstrong, T. Moore, S. W. Chang, L. H. Snyder, S. G. Lisberger, N. J. Priebe, I. M. Finn, D. Ferster, S. I. Ryu, G. Santhanam, M. Sahani, and K. V. Shenoy. Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat Neurosci*, 13(3):369–78, 2010.

Marlene R Cohen and John HR Maunsell. Attention improves performance primarily by reducing interneuronal correlations. *Nature neuroscience*, 12(12):1594, 2009.

J E Cometto-Muniz and M H Abraham. Odor Detection by Humans of Lineal Aliphatic Aldehydes and Helional as Gauged by Dose-Response Functions. *Chemical Senses*, 35(4):289–299, 2010. doi: https://doi.org/10.1093/chemse/bjq018.

John P Crimaldi, Megan B Wiley, and Jeffrey R Koseff. The relationship between mean and instantaneous structure in turbulent passive scalar plumes. *J Turbul*, 3: N14, 2002. ISSN 1468-5248. doi: 10.1088/1468-5248/3/1/014.

Maria C Dadarlat and Michael P Stryker. Locomotion enhances neural encoding of visual stimuli in mouse v1. *Journal of Neuroscience*, 37(14):3764–3775, 2017.

Sandeep Robert Datta, David J. Anderson, Kristin Branson, Pietro Perona, and Andrew Leifer. Computational Neuroethology: A Call to Action. *Neuron*, 104 (1):11–24, 2019. ISSN 08966273. doi: 10.1016/j.neuron.2019.09.038. URL `https://linkinghub.elsevier.com/retrieve/pii/S0896627319308414`.

Saskia E.J. de Vries, Jerome A. Lecoq, Michael A. Buice, Peter A. Groblewski, Gabriel K. Ocker, Michael Oliver, David Feng, Nicholas Cain, Peter Ledochowitsch, Daniel Millman, Kate Roll, Marina Garrett, Tom Keenan, Leonard Kuan, Stefan Mihalas, Shawn Olsen, Carol Thompson, Wayne Wakeman, Jack Waters, Derric Williams, Chris Barber, Nathan Berbesque, Brandon Blanchard, Nicholas Bowles, Shiella D. Caldejon, Linzy Casal, Andrew Cho, Sissy Cross, Chinh Dang, Tim Dolbeare, Melise Edwards, John Galbraith, Nathalie Gaudreault, Terri L. Gilbert, Fiona Griffin, Perry Hargrave, Robert Howard, Lawrence Huang, Sean Jewell, Nika Keller, Ulf Knoblich, Josh D. Larkin, Rachael Larsen, Chris Lau, Eric Lee, Felix Lee, Arielle Leon, Lu Li, Fuhui Long, Jennifer Luviano, Kyla Mace, Thuyanh Nguyen, Jed Perkins, Miranda Robertson, Sam Seid, Eric Shea-Brown, Jianghong Shi, Nathan Sjoquist, Cliff Slaughterbeck, David Sullivan, Ryan Valenza, Casey White, Ali Williford, Daniela M. Witten, Jun Zhuang, Hongkui Zeng, Colin Farrell, Lydia Ng, Amy Bernard, John W. Phillips, R. Clay Reid, and Christof Koch. A large-scale standardized physiological survey reveals functional organization of the mouse visual cortex. *Nature Neuroscience*, 23(1):138–151, 2020. ISSN 15461726. doi: 10.1038/s41593-019-0550-9.

G. Deco and E. Hugues. Neural network mechanisms underlying stimulus driven variability reduction. *PLoS Comput Biol*, 8(3):e1002395, 2012. ISSN 1553-734x. doi: 10.1371/journal.pcbi.1002395. URL `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3315452/pdf/pcbi.1002395.pdf`.

Gustavo Deco, Josephine Cruzat, Joana Cabral, Enzo Tagliazucchi, Helmut Laufs, Nikos K Logothetis, and Morten L Kringelbach. Awakening: Predicting external stimulation to force transitions between different brain states. *Proceedings of the National Academy of Sciences*, 116(36):18088–18097, 2019.

Daniel Deitch, Alon Rubin, and Yaniv Ziv. Representational drift in the mouse visual cortex. *Current Biology*, 31(19):4327–4339.e6, 2021. ISSN 18790445. doi: 10.1016/j.cub.2021.07.062. URL `https://doi.org/10.1016/j.cub.2021.07.062`.

Robert Desimone and John Duncan. Neural mechanisms of selective visual attention. *Annual review of neuroscience*, 18(1):193–222, 1995.

Marta D'iaz-Quesada, Isaac A Youngstrom, Yusuke Tsuno, Kyle R Hansen, Michael N Economo, and Matt Wachowiak. Inhalation Frequency Controls Reformatting of Mitral/Tufted Cell Odor Representations in the Olfactory Bulb. *Journal of Neuroscience*, 38(9):2189–2206, 2018. doi: 10.1523/JNEUROSCI.0714-17.2018.

Mario Dipoppa, Adam Ranson, Michael Krumin, Marius Pachitariu, Matteo Carandini, and Kenneth D Harris. Vision and locomotion shape the interactions between neuron types in mouse visual cortex. *Neuron*, 98(3):602–615, 2018.

R L Doty, W E Brugger, P C Jurs, and M A Orndorff. Intranasal trigeminal stimulation from odorous volatiles: psychometric responses from anosmic and normal humans. *Physiology and Behavior*, 20(2):175–185, 1978. doi: https://doi.org/10.1016/0031-9384(78)90070-7.

Guillaume P Dugué, Magor L Lörincz, Eran Lottem, Enrica Audero, Sara Matias, Patricia A Correia, Clément Léna, and Zachary F Mainen. Optogenetic recruitment of dorsal raphe serotonergic neurons acutely decreases mechanosensory responsivity in behaving mice. *PloS one*, 9(8), 2014.

Jean-rene Duhamel, Carol L Colby, and Michael E Goldberg. The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, 255(5040):90–92, 1992. ISSN 0036-8075. doi: 10.1126/science.1553535.

Christopher Ebsch and Robert Rosenbaum. Imbalanced amplification: A mechanism of amplification and suppression from local imbalance of excitation and inhibition in cortical circuits. *PLoS computational biology*, 14(3):e1006048, 2018.

Emmanuel Eggermann, Yves Kremer, Sylvain Crochet, and Carl CH Petersen. Cholinergic signals in mouse barrel cortex during active whisker sensing. *Cell reports*, 9(5):1654–1660, 2014.

Tatiana A Engel, Nicholas A Steinmetz, Marc A Gieselmann, Alexander Thiele, Tirin Moore, and Kwabena Boahen. Selective modulation of cortical state during spatial attention. *Science*, 354(6316):1140–1144, 2016.

Frédéric Esclassan, Emmanuelle Courtiol, Marc Thévenet, Samuel Garcia, Nathalie Buonviso, and Philippe Litaudon. Faster, Deeper, Better: The Impact of Sniffing Modulation on Bulbar Olfactory Processing. *PLoS ONE*, 7(7):e40927, jul 2012. doi: 10.1371/journal.pone.0040927.

Daniel J Felleman and DC Essen Van. Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex (New York, NY: 1991)*, 1(1):1–47, 1991.

David J. Field. Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, 4(12):2379, 1987. ISSN 1084-7529. doi: 10.1364/josaa.4.002379.

Aris Fiser, David Mahringer, Hassana K Oyibo, Anders V Petersen, Marcus Leinweber, and Georg B Keller. Experience-dependent spatial expectations in mouse visual cortex. 19(12), 2016. doi: 10.1038/nn.4385.

N. Fourcaud and N. Brunel. Dynamics of the firing probability of noisy integrate-and-fire neurons. *Neural Comput*, 14(9):2057–110, 2002. ISSN 0899-7667 (Print) 0899-7667 (Linking). doi: 10.1162/089976602320264015. URL `http://www.ncbi.nlm.nih.gov/pubmed/12184844`.

Yu Fu, Jason M Tucciarone, J Sebastian Espinosa, Nengyin Sheng, Daniel P Darcy, Roger A Nicoll, Z Josh Huang, and Michael P Stryker. A cortical circuit for gain control by behavioral state. *Cell*, 156(6):1139–1152, 2014.

C R Gallistel. *The Organization of Action: a new synthesis*. New York, 1982. doi: https://doi.org/10.4324/9780203780794.

Jeffrey P Gavornik and Mark F Bear. Learned spatiotemporal sequence recognition and prediction in primary visual cortex. *Nature neuroscience*, 17(5):732–737, 2014.

Ambarish S Ghatpande and Johannes Reisert. Olfactory receptor neuron responses coding for rapid odour sampling. *The Journal of Physiology*, 589(9):2261–2273, 2011. doi: 10.1113/jphysiol.2010.203687.

J.J. Gibson. Observations on active touch. *Psychological Review*, 69(6):477–491, 1962. doi: https://doi.org/10.1037/h0046962.

J.J. Gibson. *The senses considered as perceptual systems.* Houghton Mifflin, Oxford, 1966.

Charles D Gilbert and Wu Li. Top-down influences on visual processing. *Nature Reviews Neuroscience*, 14(5):350–363, 2013.

Colleen J Gillon, Jason E Pina, Jérôme A Lecoq, Ruweida Ahmed, Yazan Billeh, Shiella Caldejon, Peter Groblewski, Tim M Henley, Eric Lee, Jennifer Luviano, et al. Learning from unexpected events in the neocortical microcircuit. *BioRxiv*, 2021.

David H Gire, Vikrant Kapoor, Annie Arrighi-Allisan, Agnese Seminara, and Venkatesh N Murthy. Mice Develop Efficient Strategies for Foraging and Navigation Using Complex Natural Stimuli. *Current Biology*, 26(10):1261–1273, 2017. doi: 10.1016/j.cub.2016.03.040.

Michael Goard and Yang Dan. Basal forebrain activation enhances cortical coding of natural scenes. *Nature neuroscience*, 12(11):1444, 2009.

Alex Gomez-Marin and Matthieu Louis. Active sensation during orientation behavior in the Drosophila larva: more sense than luck. *Current Opinion in Neurobiology*, 22(2):208–215, 2012. ISSN 09594388. doi: 10.1016/j.conb.2011.11.008.

Alex Gomez-Marin, Greg J. Stephens, and Matthieu Louis. Active sampling and decision making in Drosophila chemotaxis. *Nature Communications*, 2(441):441, 2011. ISSN 2041-1723. doi: 10.1038/ncomms1455.

Alex Gomez-Marin, Joseph J. Paton, Adam R. Kampff, Rui M. Costa, and Zachary F. Mainen. Big behavioral data: psychology, ethology and the foundations of neuroscience. *Nature Neuroscience*, 17(11):1455–1462, 2014. ISSN 15461726. doi: https://doi.org/10.1038/nn.3812.

D.A. Green. A colour scheme for the display of astronomical intensity images. *Bulletin of the Astronomical Society of India*, 39:289–295, jun 2011.

Walter E Grueninger and Karl H Pribram. Effects of spatial and nonspatial distractors on performance latency of monkeys with frontal lesions. *Journal of comparative and physiological psychology*, 68(2p1):203, 1969.

Zengcai V Guo, Nuo Li, Daniel Huber, Eran Ophir, Diego Gutnisky, Jonathan T Ting, Guoping Feng, and Karel Svoboda. Flow of cortical activity underlying a tactile decision in mice. *Neuron*, 81(1):179–194, 2014.

P Gupta, D Albeanu, and U Bhalla. Olfactory bulb coding of odors, mixtures and sniffs is a linear sum of odor time profiles. *Nature Neuroscience*, 18:272–281, 2015. doi: https://doi.org/10.1038/nn.3913.

Bilal Haider and David A McCormick. Rapid neocortical dynamics: cellular and network mechanisms. *Neuron*, 62(2):171–189, 2009.

Peter Hänggi, Peter Talkner, and Michal Borkovec. Reaction-rate theory: fifty years after kramers. *Reviews of modern physics*, 62(2):251, 1990.

Jan Homann, Sue Ann Koay, Kevin S Chen, David W Tank, and Michael J Berry. Novel stimuli evoke excess activity in the mouse primary visual cortex. *Proceedings of the National Academy of Sciences*, 119(5):e2108882119, 2022.

Tomáš Hromádka, Michael R DeWeese, and Anthony M Zador. Sparse representation of sounds in the unanesthetized auditory cortex. *PLoS biology*, 6(1), 2008.

Chengcheng Huang, Douglas A Ruff, Ryan Pyle, Robert Rosenbaum, Marlene R Cohen, and Brent Doiron. Circuit models of low-dimensional shared variability in cortical networks. *Neuron*, 101(2):337–348, 2019.

David H Hubel and Torsten N Wiesel. Receptive fields of single neurones in the cat's striate cortex. *The Journal of physiology*, 148(3):574, 1959.

David H Hubel and Torsten N Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, 160 (1):106, 1962.

T Hummel, E Iannilli, J Frasnelli, J Boyle, and J Gerber. Central processing of trigeminal activation in humans. In *Annals of the New York Academy of Sciences*, volume 1170, pages 190–195, 2009. ISBN 9781573317382. doi: 10.1111/j.1749-6632.2009.03910.x.

Laura M Hurley, Ann M Thompson, and George D Pollak. Serotonin in the inferior colliculus. *Hearing research*, 168(1-2):1–11, 2002.

Minami Ito and Charles D Gilbert. Attention modulates contextual influences in the primary visual cortex of alert monkeys. *Neuron*, 22(3):593–604, 1999.

Santiago Jaramillo and Anthony Zador. Auditory cortex mediates the perceptual effects of acoustic temporal expectation. *Nature Precedings*, pages 1–1, 2010.

Sean W Jewell, Toby Dylan Hocking, Paul Fearnhead, and Daniela M Witten. Fast nonconvex deconvolution of calcium imaging data. *Biostatistics*, February 2019. doi: 10.1093/biostatistics/kxy083. URL `https://doi.org/10.1093/biostatistics/kxy083`.

Ahmad Jezzini, Luca Mazzucato, Giancarlo La Camera, and Alfredo Fontanini. Processing of hedonic and chemosensory features of taste in medial prefrontal and insular networks. *The Journal of Neuroscience*, 33(48):18966–18978, 2013.

E Jones, T Oliphant, P Peterson, , et al. SciPy: Open source scientific tools for Python, 2001.

L. M. Jones, A. Fontanini, B. F. Sadacca, P. Miller, and D. B. Katz. Natural stimuli evoke dynamic sequences of states in sensory cortical ensembles. *Proc Natl Acad Sci U S A*, 104(47):18772–7, 2007.

Peter W Jones and Nathan N Urban. Mice follow odor trails using stereo olfactory cues and rapid sniff to sniff comparisons. *bioRxiv*, page 293746, 2018. doi: 10.1101/293746.

R Jordan, M Kollo, , and A Schaefer. Sniffing Fast: Paradoxical Effects on Odor Concentration Discrimination at the Levels of Olfactory Bulb Output and Behavior. *eNeuro*, 5(5), 2018. doi: 10.1523/ENEURO.0148-18.2018.

Nikolas Karalis and Anton Sirota. Breathing Coordinates Limbic Network Dynamics Underlying Memory Consolidation. *SSRN Electronic Journal*, 2018. doi: 10.2139/ssrn.3283711.

Leslie M Kay. Theta oscillations and sensorimotor performance. *Proceedings of the National Academy of Sciences of the United States of America*, 102(10):3863–3868, 2005. ISSN 00278424. doi: 10.1073/pnas.0407920102.

Georg B. Keller and Thomas D. Mrsic-Flogel. Predictive Processing: A Canonical Cortical Computation. *Neuron*, 100(2):424–435, 2018. ISSN 10974199. doi: 10.1016/j.neuron.2018.10.003. URL `https://doi.org/10.1016/j.neuron.2018.10.003`.

Georg B Keller, Tobias Bonhoeffer, and Mark Hu. Report Sensorimotor Mismatch Signals in Primary Visual Cortex of the Behaving Mouse. pages 809–815, 2012. doi: 10.1016/j.neuron.2012.03.040.

J S Kennedy and D Marsh. Pheromone-regulated anemotaxis in flying moths. *Science (New York, NY)*, 184(4140):999–1001, 1974. doi: 10.1126/science.184.4140.999.

Adam Kepecs, Naoshige Uchida, and Zachary F Mainen. The sniff as a unit of olfactory processing. *Chemical Senses*, 31(2):167–179, 2006. doi: https://doi.org/10.1093/chemse/bjj016.

Adil G Khan and Sonja B Hofer. ScienceDirect Contextual signals in visual cortex. *Current Opinion in Neurobiology*, 52:131–138. ISSN 0959-4388. doi: 10.1016/j.conb.2018.05.003. URL `https://doi.org/10.1016/j.conb.2018.05.003`.

Adil G Khan, Jasper Poort, Angus Chadwick, Antonin Blot, Maneesh Sahani, Thomas D Mrsic-Flogel, and Sonja B Hofer. Distinct learning-induced changes in stimulus selectivity and interactions of gabaergic interneuron classes in visual cortex. *Nature neuroscience*, 21(6):851–859, 2018.

Adil Ghani Khan, Manaswini Sarangi, and Upinder Singh Bhalla. Rats track odour trails accurately using a multi-layered strategy with near-optimal sampling. *Nature Communications*, 3:703, feb 2012. ISSN 20411723. doi: 10.1038/ncomms1712.

Roozbeh Kiani, Christopher J Cueva, John B Reppas, Diogo Peixoto, Stephen I Ryu, and William T Newsome. Natural grouping of neural responses reveals spatially segregated clusters in prearcuate cortex. *Neuron*, 85(6):1359–1373, 2015.

Hyewon Kim, Jan Homann, David Tank, and Michael Berry. A Long Timescale Stimulus History Effect in the Primary Visual Cortex. *bioRxiv*, 2019a. doi: 10.1101/585539.

Hyewon Kim, Jan Homann, David W Tank, and Michael J Berry. A long timescale stimulus history effect in the primary visual cortex. *BioRxiv*, page 585539, 2019b.

David Kleinfeld, Ehud Ahissar, and Mathew E Diamond. Active sensation: insights from the rodent vibrissa sensorimotor system. *Current Opinion in Neurobiology*, 16(4):435–444, 2006. doi: 10.1016/j.conb.2006.06.009.

David Kleinfeld, Martin Deschênes, Fan Wang, and Jeffrey D. Moore. More than a rhythm of life: Breathing as a binder of orofacial sensation. *Nature Neuroscience*, 17(5):647–651, 2014. ISSN 15461726. doi: 10.1038/nn.3693.

Michelle Kloc and Arianna Maffei. Target-specific properties of thalamocortical synapses onto layer 4 of mouse primary visual cortex. *Journal of Neuroscience*, 34(46):15455–15465, 2014.

James J Knierim and David C Van Essen. Neuronal responses to static texture patterns in area v1 of the alert macaque monkey. *Journal of neurophysiology*, 67 (4):961–980, 1992.

Nina N. Kowalewski, Janne Kauttonen, Patricia L. Stan, Brian B. Jeon, Thomas Fuchs, Steven M. Chase, Tai Sing Lee, and Sandra J. Kuhlman. Development of Natural Scene Representation in Primary Visual Cortex Requires Early Postnatal Experience. *Current Biology*, 31(2):369–380.e5, 2021. ISSN 18790445. doi: 10.1016/j.cub.2020.10.046. URL `https://doi.org/10.1016/j.cub.2020.10.046`.

Anastasia Kurnikova, Jeffrey D. Moore, Song-Mao Mao Liao, Martin Deschênes, and David Kleinfeld. Coordination of Orofacial Motor Actions into Exploratory Behavior by Rat. *Current Biology*, 27(5):688–696, 2017. ISSN 09609822. doi: 10.1016/j.cub.2017.01.013.

Johannes Larsch, Steven W. Flavell, Qiang Liu, Andrew Gordus, Dirk R. Albrecht, and Cornelia I. Bargmann. A Circuit for Gradient Climbing in C. elegans Chemotaxis. *Cell Reports*, 12(11):1748–1760, 2015. ISSN 22111247. doi: 10.1016/j.celrep.2015.08.032.

A Moses Lee, Jennifer L Hoy, Antonello Bonci, Linda Wilbrecht, Michael P Stryker, and Cristopher M Niell. Identification of a brainstem circuit regulating visual cortical state in parallel with locomotion. *Neuron*, 83(2):455–466, 2014.

Wei-Chung Allen Lee, Vincent Bonin, Michael Reed, Brett J Graham, Greg Hood, Katie Glattfelder, and R Clay Reid. Anatomy and function of an excitatory network in the visual cortex. *Nature*, 532(7599):370–374, 2016.

Marcus Leinweber, Daniel R Ward, Jan M Sobczak, Alexander Attinger, and Georg B Keller. A sensorimotor circuit in mouse cortex for visual flow predictions. *Neuron*, 95(6):1420–1432, 2017.

Nuo Li, Susu Chen, Zengcai V Guo, Han Chen, Yan Huo, Hidehiko K Inagaki, Guang Chen, Courtney Davis, David Hansel, Caiying Guo, et al. Spatiotemporal constraints on optogenetic inactivation in cortical circuits. *Elife*, 8, 2019.

Sukbin Lim, Jillian L McKee, Luke Woloszyn, Yali Amit, David J Freedman, David L Sheinberg, and Nicolas Brunel. Inferring learning rules from distributions of firing rates in cortical neurons. *Nature neuroscience*, 18(12):1804, 2015.

A. Litwin-Kumar and B. Doiron. Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nat Neurosci*, 15(11):1498–505, 2012. ISSN 1546-1726 (Electronic) 1097-6256 (Linking). doi: 10.1038/nn.3220. URL `http://www.ncbi.nlm.nih.gov/pubmed/23001062`.

A. Litwin-Kumar and B. Doiron. Formation and maintenance of neuronal assemblies through synaptic plasticity. *Nat Commun*, 5:5319, 2014. ISSN 2041-1723. doi: 10.1038/ncomms6319.

Annie Liu, Andrew E. Papale, James Hengenius, Khusbu Patel, Bard Ermentrout, and Nathan N. Urban. Mouse Navigation Strategies for Odor Source Localization. *Frontiers in Neuroscience*, 14:1–16, 2020. ISSN 1662453X. doi: 10.3389/fnins.2020.00218.

Shawn R. Lockery. The computational worm: Spatial orientation and its neuronal basis in C. elegans. *Current Opinion in Neurobiology*, 21(5):782–790, 2011. ISSN 09594388. doi: 10.1016/j.conb.2011.06.009.

G Lopes, N Bonacchi, J Frazão, J Neto, B V Atallah, S Soares, L Moreira, S Matias, P M Itskov, P A Correia, R E Medina, L Calcaterra, E Dreosti, J J Paton, and A R Kampff. Bonsai: an event-based framework for processing and controlling data streams. *Frontiers in Neuroinformatics*, 9(10):52:1–14, 2015. doi: https://doi.org/10.3389/fninf.2015.00007.

Eran Lottem, Magor L Lörincz, and Zachary F Mainen. Optogenetic activation of dorsal raphe serotonin neurons rapidly inhibits spontaneous but not odor-evoked activity in olfactory cortex. *Journal of Neuroscience*, 36(1):7–18, 2016.

Kourosh Maboudi, Etienne Ackermann, Laurel Watkins de Jong, Brad E Pfeiffer, David Foster, Kamran Diba, and Caleb Kemere. Uncovering temporal structure in hippocampal output patterns. *eLife*, 7:e34467, 2018.

F Macrides, H B Eichenbaum, and W B Forbes. Temporal relationship between sniffing and the limbic theta rhythm during odor discrimination reversal learning. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 2(12):1705–1717, 1982. doi: 10.1523/JNEUROSCI.02-12-01705.

Alexandre Mahrach, Guang Chen, Nuo Li, Carl van Vreeswijk, and David Hansel. Mechanisms underlying the response of mouse cortical networks to optogenetic manipulation. *eLife*, 9, 2020.

Hiroshi Makino and Takaki Komiyama. Learning enhances the relative impact of top-down processing in the visual cortex. 18(8), 2015. doi: 10.1038/nn.4061.

Nathalie Mandairon and Christiane Linster. Odor Perception and Olfactory Bulb Plasticity in Adult Mammals. *Journal of Neurophysiology*, 101(5):2204–2209, 2009. doi: 10.1152/jn.00076.2009.

Jeffrey E Markowitz, Winthrop F Gillis, Celia C Beron, Shay Q Neufeld, Keiramarie Robertson, Neha D Bhagat, Ralph E Peterson, Emalee Peterson, Minsuk Hyun, Scott W Linderman, et al. The striatum organizes 3d behavior via moment-to-moment action selection. *Cell*, 2018.

Tyler D. Marks and Michael J. Goard. Stimulus-dependent representational drift in primary visual cortex. *Nature Communications*, 12(1):1–16, 2021. ISSN 20411723. doi: 10.1038/s41467-021-25436-3. URL http://dx.doi.org/10.1038/s41467-021-25436-3.

M. Mascaro and D. J. Amit. Effective neural response function for collective population states. *Network*, 10(4):351–73, 1999. ISSN 0954-898X (Print) 0954-898X (Linking). URL http://www.ncbi.nlm.nih.gov/pubmed/10695764.

Alexander Mathis, Pranav Mamidanna, Kevin M. Cury, Taiga Abe, Venkatesh N. Murthy, Mackenzie Weygandt Mathis, and Matthias Bethge. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, 21(9):1281–1289, 2018. ISSN 15461726. doi: 10.1038/s41593-018-0209-y.

Mackenzie Weygandt Mathis and Alexander Mathis. Deep learning tools for the measurement of animal behavior in neuroscience. *Current Opinion in Neurobiology*, 60:1–11, 2020. ISSN 18736882. doi: 10.1016/j.conb.2019.10.008.

M. Mattia and M. V. Sanchez-Vives. Exploring the spectrum of dynamical regimes and timescales in spontaneous cortical activity. *Cogn Neurodyn*, 6(3):239–50, 2012. ISSN 1871-4080 (Print). doi: 10.1007/s11571-011-9179-4. URL http://www.ncbi.nlm.nih.gov/pubmed/23730355.

Maurizio Mattia, Pierpaolo Pani, Giovanni Mirabella, Stefania Costa, Paolo Del Giudice, and Stefano Ferraina. Heterogeneous attractor cell assemblies for motor planning in premotor cortex. *Journal of Neuroscience*, 33(27):11155–11168, 2013.

Luca Mazzucato, Alfredo Fontanini, and Giancarlo La Camera. Dynamics of multistable states during ongoing and evoked cortical activity. *The Journal of Neuroscience*, 35(21):8214–8231, 2015.

Luca Mazzucato, Alfredo Fontanini, and Giancarlo La Camera. Stimuli reduce the dimensionality of cortical activity. *Frontiers in systems neuroscience*, 10:11, 2016.

Luca Mazzucato, Giancarlo La Camera, and Alfredo Fontanini. Expectation-induced modulation of metastable activity underlies faster coding of sensory stimuli. *Nature neuroscience*, page 1, 2019.

Carrie J McAdams and John HR Maunsell. Effects of attention on orientation-tuning functions of single neurons in macaque cortical area v4. *Journal of Neuroscience*, 19(1):431–441, 1999.

David A McCormick. Neurotransmitter actions in the thalamus and cerebral cortex and their role in neuromodulation of thalamocortical activity. *Progress in neurobiology*, 39(4):337–388, 1992.

Matthew J McGinley, Stephen V David, and David A McCormick. Cortical membrane potential signature of optimal states for sensory signal detection. *Neuron*, 2015a.

Matthew J McGinley, Martin Vinck, Jacob Reimer, Renata Batista-Brito, Edward Zagha, Cathryn R Cadwell, Andreas S Tolias, Jessica A Cardin, and David A McCormick. Waking state: rapid variations modulate neural and behavioral responses. *Neuron*, 87(6):1143–1161, 2015b.

David BT McMahon and Carl R Olson. Repetition suppression in monkey inferotemporal cortex: relation to behavioral priming. *Journal of neurophysiology*, 97(5):3532–3543, 2007.

Travis Meyer, Suchitra Ramachandran, and Carl R Olson. Statistical learning of serial visual transitions by neurons in monkey inferotemporal cortex. *Journal of Neuroscience*, 34(28):9332–9337, 2014.

Angie M Michaiel, Philip RL Parker, and Cristopher M Niell. A hallucinogenic serotonin-2a receptor agonist reduces visual response gain and alters temporal dynamics in mouse v1. *Cell reports*, 26(13):3475–3483, 2019.

Angie M. Michaiel, Elliott T.T. Abe, and Cristopher M. Niell. Dynamics of gaze control during prey capture in freely moving mice. *eLife*, 9:1–27, 2020. ISSN 2050084X. doi: 10.7554/eLife.57458.

P. Miller and D. B. Katz. Stochastic transitions between neural states in taste processing and decision-making. *J Neurosci*, 30(7):2559–70, 2010. ISSN 0270-6474. doi: 10.1523/jneurosci.3047-09.2010. URL `http://www.jneurosci.org/content/30/7/2559.full.pdf`.

Nathaniel J Miska, Leonidas MA Richter, Brian A Cary, Julijana Gjorgjieva, and Gina G Turrigiano. Sensory experience inversely regulates feedforward and feedback excitation-inhibition ratio in rodent visual cortex. *Elife*, 7:e38846, 2018.

Alexandra K Moore, Aldis P Weible, Timothy S Balmer, Laurence O Trussell, and Michael Wehr. Rapid rebalancing of excitation and inhibition by cortical circuitry. *Neuron*, 97(6):1341–1355, 2018.

Jeffrey D. Moore, Martin Deschênes, Takahiro Furuta, Daniel Huber, Matthew C. Smear, Maxime Demers, and David. Kleinfeld. Hierarchy of orofacial rhythms revealed through whisking and breathing. *Nature*, 497(7448):205–210, 2013. ISSN 00280836. doi: 10.1038/nature12076.

J. Murlis, J. S. Elkinton, and R. T. Carde. Odor plumes and how insects use them. *Annual Review of Entomology*, 37(86):505–532, 1992. ISSN 00664170. doi: 10.1146/annurev.en.37.010192.002445.

Kevin P Murphy. *Machine learning: a probabilistic perspective (adaptive computation and machine learning series)*. 2012. ISBN 0262018020.

Simon Musall, Matthew T. Kaufman, Ashley L. Juavinett, Steven Gluf, and Anne K. Churchland. Single-trial neural dynamics are dominated by richly varied movements. *Nature Neuroscience*, 22(10):1677–1686, 2019. ISSN 15461726. doi: 10.1038/s41593-019-0502-4. URL `http://dx.doi.org/10.1038/s41593-019-0502-4`.

Anders Nelson, David M Schneider, Jun Takatoh, Katsuyasu Sakurai, Fan Wang, and Richard Mooney. A circuit for motor cortical modulation of auditory cortical activity. *Journal of Neuroscience*, 33(36):14342–14353, 2013.

Cristopher M Niell and Michael P Stryker. Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron*, 65(4):472–479, 2010.

Pekka Niemi and Risto Näätänen. Foreperiod and simple reaction time. *Psychological bulletin*, 89(1):133, 1981.

Danko Nikolić, Stefan Häusler, Wolf Singer, and Wolfgang Maass. Distributed fading memory for stimulus properties in the primary visual cortex. *PLoS biology*, 7(12): e1000260, 2009.

Daniel H O'Connor, Simon P Peron, Daniel Huber, and Karel Svoboda. Neural activity in barrel cortex underlying vibrissa-based object localization in mice. *Neuron*, 67(6):1048–1061, 2010.

Michael Okun and Ilan Lampl. Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nature neuroscience*, 11(5):535–537, 2008.

Timothy M Otchy, Steffen BE Wolff, Juliana Y Rhee, Cengiz Pehlevan, Risa Kawai, Alexandre Kempf, Sharon MH Gobes, and Bence P Ölveczky. Acute off-target effects of neural circuit manipulations. *Nature*, 528(7582):358–363, 2015.

Hirofumi Ozeki, Ian M Finn, Evan S Schaffer, Kenneth D Miller, and David Ferster. Inhibitory stabilization of the cortical network underlies visual surround suppression. *Neuron*, 62(4):578–592, 2009.

Ana Parabucki, Alexander Bizer, Genela Morris, Antonio E. Munoz, Avinash D.S. S Bala, Matthew Smear, and Roman Shusterman. Odor concentration change coding in the olfactory bulb. *eNeuro*, 6(1):1–13, 2019. ISSN 23732822. doi: 10.1523/ENEURO.0396-18.2019.

K Parthasarathy and U S Bhalla. Laterality and Symmetry in Rat Olfactory Behavior and in Physiology of Olfactory Input. *Journal of Neuroscience*, 33(13):5750–5760, 2013. doi: https://doi.org/10.1523/JNEUROSCI.

Michael Andrew Patterson, Samuel Lagier, and Alan Carleton. Odor representations in the olfactory bulb evolve after the first breath and persist as an odor afterimage. *Proceedings of the National Academy of Sciences*, 110(35):E3340–E3349, 2013. doi: 10.1073/pnas.1303873110.

Ulises Pereira and Nicolas Brunel. Attractor dynamics in networks with learning rules inferred from in vivo data. *Neuron*, 99(1):227–238, 2018.

Rodrigo Perin, Thomas K Berger, and Henry Markram. A synaptic organizing principle for cortical neuronal groups. *Proceedings of the National Academy of Sciences*, 108(13):5419–5424, 2011.

Leopoldo Petreanu, Diego A Gutnisky, Daniel Huber, Ning-long Xu, Dan H O'Connor, Lin Tian, Loren Looger, and Karel Svoboda. Activity in motor–sensory projections reveals distributed coding in somatosensation. *Nature*, 489(7415):299, 2012.

Elizabeth AK Phillips and Andrea R Hasenstaub. Asymmetric effects of activating and inactivating cortical interneurons. *Elife*, 5:e18383, 2016.

Lucas Pinto, Michael J Goard, Daniel Estandian, Min Xu, Alex C Kwan, Seung-Hee Lee, Thomas C Harrison, Guoping Feng, and Yang Dan. Fast modulation of visual perception by basal forebrain cholinergic neurons. *Nature neuroscience*, 16(12): 1857–1863, 2013.

Pierre-Olivier Polack, Jonathan Friedman, and Peyman Golshani. Cellular mechanisms of brain state–dependent gain modulation in visual cortex. *Nature neuroscience*, 16(9):1331, 2013.

A. Ponce-Alvarez, V. Nacher, R. Luna, A. Riehle, and R. Romo. Dynamics of cortical neuronal ensembles transit from decision making to storage for later report. *J Neurosci*, 32(35):11956–69, 2012.

Jasper Poort, Adil G. Khan, Marius Pachitariu, Abdellatif Nemri, Ivana Orsolic, Julija Krupic, Marius Bauza, Maneesh Sahani, Georg B. Keller, Thomas D. Mrsic-Flogel, and Sonja B. Hofer. Learning Enhances Sensory and Multiple Non-sensory Representations in Primary Visual Cortex. *Neuron*, 86(6):1478–1490, 2015. ISSN 10974199. doi: 10.1016/j.neuron.2015.05.037. URL `http://dx.doi.org/10.1016/j.neuron.2015.05.037`.

Alan B Poritz. Linear predictive hidden Markov models and the speech signal. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 1982. doi: 10.1109/ICASSP.1982.1171633.

Jess Porter, Brent Craven, Rehan M. Khan, Shao-Ju Ju Chang, Irene Kang, Benjamin Judkewitz, Benjamin Judkewicz, Jason Volpe, Gary Settles, and Noam Sobel. Mechanisms of scent-tracking in humans. *Nature Neuroscience*, 10(1):27–29, 2007. ISSN 10976256. doi: 10.1038/nn1819.

James F A Poulet and Berthold Hedwig. The cellular basis of a corollary discharge. *Science*, 311(5760):518–522, 2006. doi: 10.1126/science.1120847.

José Esquivelzeta Rabell, Kadir Mutlu, João Noutel, Pamela Martin del Olmo, and Sebastian Haesler. Spontaneous Rapid Odor Source Localization Behavior Requires Interhemispheric Communication. *Current Biology*, 27(10):1–12, 2017. doi: 10.1016/j.cub.2017.04.027.

Neil C Rabinowitz, Robbe L Goris, Marlene Cohen, and Eero P Simoncelli. Attention stabilizes the shared gain of v4 populations. *Elife*, 4:e08998, 2015.

Raghav Rajan, James P Clement, and Upinder S Bhalla. Rats smell in stereo. *Science (New York, NY)*, 311(5761):666–670, 2006. doi: 10.1126/science.1122096.

S Ranade, B Hangya, and A Kepecs. Multiple Modes of Phase Locking between Sniffing and Whisking during Active Exploration. *The Journal of Neuroscience : the official journal of the Society for Neuroscience*, 33(19):8250–8256, 2013. doi: 10.1523/JNEUROSCI.3874-12.2013.

Stefano Recanatesi, Ulises Pereira, Masayoshi Murakami, Zachary Mainen, and Luca Mazzucato. Metastable attractors explain the variable timing of stable behavioral action sequences. *bioRxiv*, 2020.

Jacob Reimer, Matthew J McGinley, Yang Liu, Charles Rodenkirch, Qi Wang, David A McCormick, and Andreas S Tolias. Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nature communications*, 7(1):1–7, 2016.

John H Reynolds and David J Heeger. The normalization model of attention. *Neuron*, 61(2):168–185, 2009.

Erin L Rich and Jonathan D Wallis. Decoding subjective decisions from orbitofrontal cortex. *Nature neuroscience*, 19(7):973, 2016.

Jeffrey A Riffell, Leif Abrell, and John G Hildebrand. Physical Processes and Real-Time Chemical Measurement of the Insect Olfactory Environment. *Journal of Chemical Ecology*, 34(7):837–853, 2008. doi: https://doi.org/10.1007/s10886-008-9490-7.

Pablo Rodríguez-Sánchez, Egbert H Van Nes, and Marten Scheffer. Climbing escher's stairs: A way to approximate stability landscapes in multidimensional systems. *PLOS Computational Biology*, 16(4):e1007788, 2020.

Vahid Rostami, Thomas Rost, Alexa Riehle, Sacha J van Albada, and Martin P Nawrot. Spiking neural network model of motor cortex with joint excitatory and inhibitory clusters reflects task uncertainty, reaction times, and variability dynamics. *bioRxiv*, 2020.

Alex Roxin, Nicolas Brunel, David Hansel, Gianluigi Mongillo, and Carl van Vreeswijk. On the distribution of firing rates in networks of cortical neurons. *Journal of Neuroscience*, 31(45):16217–16226, 2011.

Brian F Sadacca, Narendra Mukherjee, Tony Vladusich, Jennifer X Li, Donald B Katz, and Paul Miller. The behavioral relevance of cortical neural ensemble responses emerges suddenly. *Journal of Neuroscience*, 36(3):655–669, 2016.

Aman B Saleem, Aslı Ayaz, Kathryn J Jeffery, Kenneth D Harris, and Matteo Carandini. Integration of visual motion and locomotion in mouse visual cortex. *Nature neuroscience*, 16(12):1864–1869, 2013.

David B Salkoff, Edward Zagha, Erin McCarthy, and David A McCormick. Movement and performance explain widespread cortical activity in a visual detection task. *Cerebral Cortex*, 30(1):421–437, 2020.

Chad L Samuelsen, Matthew PH Gardner, and Alfredo Fontanini. Effects of cue-triggered expectation on cortical processing of taste. *Neuron*, 74(2):410–422, 2012.

A Sanzeni, Bradley Akitake, Hannah C Goldbach, Caitlin E Leedy, Nicolas Brunel, and Mark H Histed. Inhibition stabilization is a widespread property of cortical networks. *bioRxiv*, page 656710, 2019.

Michael T Schaub, Yazan N Billeh, Costas A Anastassiou, Christof Koch, and Mauricio Barahona. Emergence of slow-switching assemblies in structured neuronal networks. *PLoS computational biology*, 11(7), 2015.

Carl E. Schoonover, Sarah N. Ohashi, Richard Axel, and Andrew J.P. Fink. Representational drift in primary olfactory cortex. *Nature*, 594(7864):541–546, 2021. ISSN 14764687. doi: 10.1038/s41586-021-03628-7. URL `http://dx.doi.org/10.1038/s41586-021-03628-7`.

Charles E Schroeder, Donald A Wilson, Thomas Radman, Helen Scharfman, and Peter Lakatos. Dynamics of Active Sensing and perceptual selection. *Current Opinion in Neurobiology*, 20(2):172–176, 2010. doi: 10.1016/j.conb.2010.02.010.

Lenka Seillier, Corinna Lorenz, Katsuhisa Kawaguchi, Torben Ott, Andreas Nieder, Paria Pourriahi, and Hendrikje Nienborg. Serotonin decreases the gain of visual responses in awake macaque v1. *Journal of Neuroscience*, 37(47):11390–11405, 2017.

M Shafi, Y Zhou, J Quintana, C Chow, J Fuster, and M Bodner. Variability in neuronal activity in primate cortex during working memory tasks. *Neuroscience*, 146(3):1082–1108, 2007.

Marshall G Shuler and Mark F Bear. Reward timing in the primary visual cortex. *Science*, 311(5767):1606–1609, 2006.

Roman Shusterman, Matthew C. Smear, Alexei A. Koulakov, and Dmitry Rinberg. Precise olfactory responses tile the sniff cycle. *Nature Neuroscience*, 14(8):1039–1044, 2011. ISSN 10976256. doi: 10.1038/nn.2877.

Joshua H Siegle, Xiaoxuan Jia, Séverine Durand, Sam Gale, Corbett Bennett, Nile Graddis, Greggory Heller, Tamina K Ramirez, Hannah Choi, Jennifer A Luviano, et al. A survey of spiking activity reveals a functional hierarchy of mouse corticothalamic visual areas. *bioRxiv*, page 805010, 2019.

Joshua H. Siegle, Peter Ledochowitsch, Xiaoxuan Jia, Daniel J. Millman, Gabriel K. Ocker, Shiella Caldejon, Linzy Casal, Andy Cho, Daniel J. Denman, Séverine Durand, Peter A. Groblewski, Gregg Heller, India Kato, Sara Kivikas, Jérôme Lecoq, Chelsea Nayan, Kiet Ngo, Philip R. Nicovich, Kat North, Tamina K. Ramirez, Jackie Swapp, Xana Waughman, Ali Williford, Shawn R. Olsen, Christof Koch, Michael A. Buice, and Saskia E.J. de Vries. Reconciling functional differences in populations of neurons recorded with two-photon imaging and electrophysiology. *eLife*, 10:1–35, 2021. ISSN 2050084X. doi: 10.7554/eLife.69068.

Marc A Sommer and Robert H Wurtz. Brain Circuits for the Internal Monitoring of Movements. *Annual Review of Neuroscience*, 31:317–338, 2008. ISSN 0147-006X. doi: 10.1146/annurev.neuro.31.060407.125627.

Sen Song, Per Jesper Sjöström, Markus Reigl, Sacha Nelson, and Dmitri B Chklovskii. Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS biology*, 3(3), 2005.

Roger W Sperry. Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of comparative and physiological psychology*, 43(6):482, 1950. doi: https://doi.org/10.1037/h0055479.

Carsen Stringer, Marius Pachitariu, Nicholas Steinmetz, Charu Bai Reddy, Matteo Carandini, and Kenneth D Harris. Spontaneous behaviors drive multidimensional, brain-wide population activity. *BioRxiv*, page 306019, 2018.

Carsen Stringer, Marius Pachitariu, Nicholas Steinmetz, Matteo Carandini, and Kenneth D Harris. High-dimensional geometry of population responses in visual cortex. *Nature*, page 1, 2019a.

Carsen Stringer, Marius Pachitariu, Nicholas Steinmetz, Charu Bai Reddy, Matteo Carandini, and Kenneth D. Harris. Spontaneous behaviors drive multidimensional, brainwide activity. *Science*, 364(6437), 2019b. ISSN 10959203. doi: 10.1126/science.aav7893.

Jalil Taghia, Weidong Cai, Srikanth Ryali, John Kochalka, Jonathan Nicholas, Tianwen Chen, and Vinod Menon. Uncovering hidden brain state dynamics that regulate performance and decision-making during cognition. *Nature communications*, 9(1):2505, 2018.

Adriano B.L. Tort, Jurij Brankačk, and Andreas Draguhn. Respiration-Entrained Brain Rhythms Are Global but Often Overlooked. *Trends in Neurosciences*, 41(4): 186–197, 2018. ISSN 1878108X. doi: 10.1016/j.tins.2018.01.007.

Anne M Treisman and Garry Gelade. A feature-integration theory of attention. *Cognitive psychology*, 12(1):97–136, 1980.

Stefan Treue and Julio C Martinez Trujillo. Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, 399(6736):575–579, 1999.

Misha V Tsodyks, William E Skaggs, Terrence J Sejnowski, and Bruce L McNaughton. Paradoxical effects of external modulation of inhibitory interneurons. *Journal of neuroscience*, 17(11):4382–4388, 1997.

R Vallat. Pingouin: statistics in python. *Journal of Open Source Software*, February 2018. doi: https://doi.org/10.21105/joss.01026. URL `https://pingouin-stats.org/`.

Floris van Breugel and Michael H. Dickinson. Plume-tracking behavior of flying drosophila emerges from a set of distinct sensory-motor reflexes. *Current Biology*, 24(3):274–286, 2014. ISSN 09609822. doi: 10.1016/j.cub.2013.12.023.

D. C. Van Essen. Visual areas of the mammalian cerebral cortex. *Annual review of neuroscience*, 2:227–263, 1979. ISSN 0147006X. doi: 10.1146/annurev.ne.02.030179.001303.

C H Vanderwolf. Hippocampal activity, olfaction, and sniffing: an olfactory input to the dentate gyrus. *Brain Research*, 593(2):197–208, 1992. doi: https://doi.org/10.1016/0006-8993(92)91308-2.

Seralynne D. Vann, John P. Aggleton, and Eleanor A. Maguire. What does the retrosplenial cortex do? *Nature Reviews Neuroscience*, 10(11):792–802, 2009. ISSN 1471003X. doi: 10.1038/nrn2733.

V Justus Verhagen, Daniel W Wesson, Theoden I Netoff, John A White, and Matt Wachowiak. Sniffing controls an adaptive filter of sensory input to the olfactory bulb. *Nature Neuroscience*, 10(5):631–639, 2007. doi: https://doi.org/10.1038/nn1892.

Neil J Vickers and Thomas C Baker. Reiterative responses to single strands of odor promote sustained upwind flight and odor source location by moths. *Proceedings of the National Academy of Sciences of the United States of America*, 91(13):5756–5760, 1994. doi: 10.1073/pnas.91.13.5756.

Roberto Vincis and Alfredo Fontanini. Associative learning changes cross-modal representations in the gustatory cortex. *Elife*, 5:e16420, 2016.

Erich von Holst and Horst Mittelstaedt. Das Reafferenzprinzip. *Naturwissenschaften*, 37(20):464–476, 1950. ISSN 1432-1904. doi: https://doi.org/10.1007/BF00622503.

Matt Wachowiak. All in a Sniff: Olfaction as a Model for Active Sensing. *Neuron*, 71(6):962–973, 2011. ISSN 08966273. doi: 10.1016/j.neuron.2011.08.030. URL `http://dx.doi.org/10.1016/j.neuron.2011.08.030`.

Barbara Webb. Neural mechanisms for prediction: Do insects have forward models? *Trends in Neurosciences*, 27(5):278–282, 2004. ISSN 01662236. doi: 10.1016/j.tins.2004.03.004.

Paul Weiss. Self-differentiation of the basic patterns of coordination. In *Dynamics of Development: Experiments and Inferences*, pages 486–581. Academic Press, 1968. doi: 10.1016/b978-1-4832-2919-5.50027-1.

W I Welker. Analysis of sniffing of the albino rat. *Behaviour, Leiden*, 22(3-4):223–244, 1964. doi: https://doi.org/10.1163/156853964X00030.

Daniel W. Wesson, Justus V. Verhagen, Matt Wachowiak, V J Verhagen, and Matt Wachowiak. Why sniff fast? The relationship between sniff frequency, odor discrimination, and receptor neuron activation in the rat. *Journal of Neurophysiology*, 101(2):1089–1102, 2009. ISSN 00223077. doi: 10.1152/jn.90981.2008.

Alexander B Wiltschko, Matthew J Johnson, Giuliano Iurilli, Ralph E Peterson, Jesse M Katon, Stan L Pashkovski, Victoria E Abraira, Ryan P Adams, and Sandeep Robert Datta. Mapping sub-second structure in mouse behavior. *Neuron*, 88(6):1121–1135, 2015.

Han Yan, Lei Zhao, Liang Hu, Xidi Wang, Erkang Wang, and Jin Wang. Nonequilibrium landscape theory of neural networks. *Proceedings of the National Academy of Sciences*, 110(45):E4185–E4194, 2013.

Y Yanovsky, M Ciatipis, A Draguhn, A B L Tort, and J Branka k. Slow Oscillations in the Mouse Hippocampus Entrained by Nasal Respiration. *Journal of Neuroscience*, 34(17):5949–5964, 2014. doi: 10.1523/JNEUROSCI.5287-13.2014.

Alfred L Yarbus. *Eye Movements and Vision*. Springer, Boston, MA, 1967. doi: 10.1007/978-1-4899-5379-7.

Yossi Yovel, Ben Falk, Cynthia Moss, and Nachum Ulanovsky. Optimal Localization by Pointing Off Axis. *Science*, 327(February):701–705, 2010. doi: DOI:https://doi.org/10.1126/science.1183310.

Edward Zagha, Amanda E Casale, Robert NS Sachdev, Matthew J McGinley, and David A McCormick. Motor cortex feedback influences sensory processing by modulating network state. *Neuron*, 79(3):567–578, 2013.

Edward Zagha, Xinxin Ge, and David A McCormick. Competing neural ensembles in motor cortex gate goal-directed motor output. *Neuron*, 88(3):565–577, 2015.

Christina Zelano, Heidi Jiang, Guangyu Zhou, Nikita Arora, Stephan Schuele, Joshua Rosenow, and Jay A Gottfried. Nasal respiration entrains human limbic oscillations and modulates cognitive function. *Journal of Neuroscience*, 36(49):12448–12467, 2016. ISSN 15292401. doi: 10.1523/JNEUROSCI.2586-16.2016.

Siyu Zhang, Min Xu, Tsukasa Kamigaki, Johnny Phong Hoang Do, Wei-Cheng Chang, Sean Jenvay, Kazunari Miyamichi, Liqun Luo, and Yang Dan. Long-range and local circuits for top-down modulation of visual cortex processing. *Science*, 345(6197):660–665, 2014.

Fu-Wen Zhou, Zuo-Yi Shao, Michael T Shipley, and Adam C Puche. Short-term plasticity in glomerular inhibitory circuits shapes olfactory bulb output. *Journal of Neurophysiology*, 123(3):1120–1132, 2020. doi: 10.1152/jn.00628.2019.

Jun Zhuang, Carl R Stoelzel, Yulia Bereshpolova, Joseph M Huff, Xiaojuan Hei, Jose-Manuel Alonso, and Harvey A Swadlow. Layer 4 in primary visual cortex of the awake rabbit: contrasting properties of simple cells and putative feedforward inhibitory interneurons. *Journal of Neuroscience*, 33(28):11372–11389, 2013.