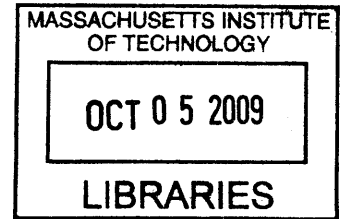# Replay of memories of extended behavior in the rat hippocampus

by

Thomas James Davidson

A.B. History and Science
Harvard College, 1999

SUBMITTED TO THE DEPARTMENT OF BRAIN AND COGNITIVE SCIENCES IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN NEUROSCIENCE
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

SEPTEMBER, 2009

ARCHIVES

Signature of Author: _____
Department of Brain and Cognitive Sciences
August 26, 2009

Certified by: _____
Matthew A. Wilson
Sherman Fairchild Professor of Neuroscience
Thesis Supervisor

Accepted by: _____
Earl K. Miller
Picower Professor of Neuroscience
Chairman, Committee of Graduate Students

Replay of memories of extended behavior in the rat hippocampus

by

Thomas James Davidson

Submitted to the Department of Brain and Cognitive Sciences
on September 1, 2009 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in
Neuroscience

Abstract

The hippocampus is a highly conserved structure in the medial temporal lobe of the brain that is known to be critical for spatial learning in rodents, and spatial and episodic memory in humans.

During pauses in exploration, ensembles of place cells in the rat hippocampus re-express firing sequences corresponding to recent spatial experience. Such 'replay' co-occurs with ripple events: short-lasting (~50-120 ms), high frequency (~200 Hz) oscillations that are associated with increased hippocampal-cortical communication. In previous studies, rats explored small environments, and replay was found to be anchored to the rat's current location, and compressed in time such that replay of the complete environment occurred during a single ripple event.

In this thesis, we develop a probabilistic neural decoding approach that allows us to show that firing sequences corresponding to long runs through a large environment are replayed with high fidelity (in both forward and reverse order). We show that such replay can begin at remote locations on the track, and proceeds at a characteristic virtual speed of ~8 m/s. Replay remains coherent across trains of sharp wave-ripple events.

These results suggest that extended replay is composed of chains of shorter subsequences, which may reflect a strategy for the storage and flexible expression of memories of prolonged experience. We discuss the evidence for the operation of similar mechanisms in humans.

Thesis Supervisor: Matthew Wilson, PhD
Title: Sherman Fairchild Professor of Neurobiology

It is quite true what philosophy says: that life must be understood backwards.
But then one forgets the other principle: that it must be lived forwards.

-Kierkegaard, Journals.

Acknowledgements:

My thanks are due first of all to Matt Wilson for accepting me as his student. I am grateful for his unflagging support since then, for the many hours spent in front of whiteboards or at beer hours, and most of all for the chance to learn the practice of science from such a creative, principled, and dedicated scientist. I am also grateful to my previous mentors Marge Livingstone, Bevil Conway, and Carol Troy, both for giving me the chance to jump in feet-first to the study of neurobiology, and for their continued support and guidance. The entire staff and faculty of the Brain and Cognitive Sciences department have my thanks for their support of graduate education, and, especially in the case of Denise Heintze, for their commitment to the well-being of graduate students.

Fabian Kloosterman and I have worked closely together the last few years, and the results presented in this thesis are only part of the fruits of our collaboration: Fabian's kindness, intelligence, and generosity made working together both a great pleasure and a great learning experience.

I am also grateful to all the other labmates who have made the Wilson lab such a fun and stimulating place to work. In particular, Daoyun Ji, Albert Lee, Linus Sun, Sujith Vijayan, David Foster, my officemate Jun Yamamoto, Jay Levin, and Matt Jones all welcomed me to the lab and generously shared their time, care, advice, space and code. Valerie Ego-Stengel, Eric Jonas, Hector Penagos, Dave Nguyen, and Jeenah Jung, my 'classmates', were invaluable sources of intellectual support and friendship. The many "beloved brave young fellow scientists" (as Miguel might put it) who have come since have only made the lab kinder, smarter, and more sterile (in a good way!). Jared Dunn and Will Caviness deserve special thanks for all their work to make the lab run smoothly, which is no mean feat.

My life over these years has been greatly enriched by many friendships, especially by those of my housemates Aaron Andalman, Ben Scott, and Sarah Jacoby. I am more grateful to Maria Bradley-Moore for her love and patience than I can express here. Lastly, I thank my parents, Joy and Dave, and my brother Simon for loving, supporting, and inspiring me.

Table of Contents

Chapter 1:

Introduction

How are memories of events in our lives encoded, stored, and recalled in the brain? What form do they take, and where in the brain should we look for them? Could a deeper understanding of the way in which remembered experience is represented in the brain shed light on the structure of thought itself? These fundamental questions will provide the motivation for the studies we will describe in this thesis. In particular, we will be interested in one physiological form of memory—the reactivation of structured patterns of activity in the hippocampus—and in the behavioral and cognitive functions this memory might support. We begin with a brief historical account of the work that motivated the current studies.

Memory processing and re-processing: from consolidation to imagination

For more than 100 years, neurologists and neuroscientists have appreciated that memories continue to be processed in the brain long after the initial learning episode (Lechner et al., 1999). Memory for recent events is vulnerable to post-learning interference from additional memory load, from brain injury, and from pharmacological and electrical manipulations, suggesting that memories normally undergo consolidation over time to a more interference-resistant form (reviewed in (McGaugh, 1966)). These observations have helped motivate and constrain the ensuing search for the 'engram'—the physiological instantiation of the memory trace.

An important breakthrough in this search came with Donald Hebb's 1949 proposal of the 'dual-trace' theory of memory. Hebb proposed that a subset of cells in a circuit (a 'cell assembly'), once driven by sensory stimuli, could remain persistently active (the first trace) allowing that pattern of activity to be captured—effectively burned in to the circuit, forming a second trace—by activity-dependent changes in the synaptic connections between the cells in the assembly. (This work drew support and inspiration from Lorente de No's work on closed anatomical loops in cortical circuits (Lorente de Nó, 1938)).

The role of the hippocampal formation

Consolidation theory suggested when the engram might be observed outside of the learning epoch, and Hebb's theory proposed a form for the engram: at the network level, it should resemble the original sensory-evoked activity. Another important development came with Scoville and Milner's 1957 report of severe amnesia in a number of patients who had undergone experimental neurosurgery to treat their epilepsy, bipolar disorder, depression, or psychosis. In

each case, the patients had had portions of their medial temporal lobes removed bilaterally. The authors found that more extensive lesions led to more profound amnesias. The most famous patient from this study (known by his initials: H.M.) experienced both retrograde amnesia extending to several days before the surgery, and a profound permanent anterograde amnesia. That is, he was unable to form new long-term memories for facts or events (though his working memory and his ability to learn new motor tasks was largely spared). Structural magnetic resonance imaging showed that most of H.M.'s hippocampus, amygdaloid complex, subiculum, and the neighboring entorhinal cortex had been removed in the surgery (Corkin et al., 1997).

The intact memory these patients exhibited for events occurring before the surgery suggested a model in which memories might temporarily be stored in the hippocampal formation, and then transferred to long-term memory. It was also apparent that the damaged regions were not required for the recall of remote memory. This finding transformed thinking about memory consolidation: rather than focusing on changes only within an activated circuit, as Hebb had done, the "systems consolidation" view saw consolidation as a cross-brain region phenomenon (albeit one that still depended on synaptic modification within the involved regions). An influential model of associational memory (Marr, 1971) suggested how such a temporary memory store might be instantiated in the hippocampus, with rapid learning taking place in the highly recurrent networks of CA3. Another explanation for the pattern of memory deficits, which would later find currency especially with respect to amygdalar function, is that the lesioned regions *modulate* memory storage processes in other regions, without necessarily ever containing a memory-specific information trace (Cahill and McGaugh, 1996).[1]

The specific pattern of memory deficits seen in patients with hippocampal damage introduced the concept of multiple distinct memory systems in the brain, with H.M.'s spared ability to learn new motor tasks serving as a particularly dramatic example. Squire (1992) introduced the term 'declarative' memory to characterize the memory for facts or events that was apparently supported by the hippocampus and parahippocampal regions. It remains a point of controversy whether the hippocampus proper is required for the acquisition of both memories for facts (semantic memory) and events (episodic memory), with some reports of patients with profound deficits of episodic and spatial memory, who nonetheless retained the ability to learn

---

1 Another explanation for the observed pattern of memory deficits—which would later find currency especially with respect to amygdalar function—is that the lesioned regions *modulate* memory storage processes in other regions, without necessarily ever containing a memory-specific information trace (Cahill and McGaugh, 1996).

new facts and words (Vargha-Khadem et al., 1997). It is widely agreed upon, however, that the hippocampus is required for the acquisition of episodic memories.

New roles for hippocampal reactivation: recall, imagination, self-projection?

The reactivation-consolidation hypothesis presented above is predicated on the notion that the same cortical inputs that are activated during acquisition of a memory are activated during the recall or subsequent 'use' of that memory. This suggests that this same reactivation might be apparent in the hippocampal ensemble during recall.

H.M.'s ability to recall facts and events that occurred before his surgery was famously spared, but there is growing evidence that hippocampal damage in humans may result in additional deficits that were not initially appreciated. In particular, the ability to vividly re-experience episodes seems to be degraded. H.M. himself was reported not to be able to recount any episodes relating to his parents, though he remembered many specific details about their relationship—Corkin reports that his autobiographical memories appeared to have been "semanticized" (Corkin, 2002).

Deficits in imagination and future thinking

Yet another deficit exhibited by hippocampal patients that was not initially reported involved an impaired ability to think about the future or to imagine new scenes. Tulving was the first to report a comorbid deficit of episodic memory and thinking about the future (1985) in a patient with both hippocampal and frontal lobe damage. This patient was unable to imagine events in his personal future, despite having intact imagery abilities. Subsequent work with amnesics and patients with Alzheimer's disease has confirmed that these patients also exhibit deficits when asked to imagine new scenarios. In one study (Hassabis et al., 2007), amnesic patients with bilateral hippocampal damage were asked to imagine themselves "lying on a white sandy beach on a beautiful tropical day". Normal controls produced rich sensory details and reported the feeling of 'being there': "The sand underneath me is unbearably hot; I can hear the sounds of small wavelets lapping on the beach." One amnesic also reported hearing sea noises, and the feel of grains of sand, but when they were asked if they were "actually seeing this in [their] mind's eye", answered "No, the only thing I can see is blue... It's kind of like I'm floating."

Results from amnesics, along with functional imaging studies in uninjured patients, have led to speculation that recall, spatial navigation, and imagination may all be supported by a common cognitive process, namely the capacity for "self-projection", and may all depend on a common network of brain regions that includes the medial temporal lobe (Buckner and Carroll, 2007; Schacter et al., 2007). Buckner and Carroll (2007) have attempted to link this theory to rodent studies of the hippocampus, and have proposed that 'pre-play' of neural activity sequences corresponding to upcoming behavior (see below) could contribute to their ability to evaluate the consequences of their decisions.

Cellular models of memory: electrophysiology of the hippocampus

These studies of human subjects and the resulting theoretical work has been complemented by ongoing electrophysiological studies of the hippocampus. In the following sections we will focus on two major strains in this work: 1) the organization of hippocampal electrical activity (both neural firing and oscillations in the local field potential) into a small number of discrete processing modes; and 2) the finding that the principal cells of the hippocampus are driven reliably by the animal's movement through particular locations. These two themes are of particular importance because they allow the process of memory formation during self-directed behavior to be brought under experimental control, thus leading to the ability to test specific hypotheses about the network-level activity underlying memory processing. In particular, we will argue that the learning and re-expression of sequential neural activity patterns corresponding to sequential spatial behavior is a powerful model for certain essential features of episodic memory.

Hippocampal activity during wakefulness is organized into two modes

During wakefulness, neural activity in the hippocampus of many species, including humans, is organized into two distinct states: 'Theta', characterized by a strong slow (6-12 Hz in the rat) rhythmic firing of the primary cells and many inhibitory interneurons, and an accompanying oscillation in the local field potential; and large irregular activity ('LIA'), characterized by a lack of coherent oscillations, and interrupted by occasional short-lived (~100ms) bursts of activity that recruit a large fraction of the neural population.[2] Both theta and LIA states are are generally observed (and coherently expressed) throughout the hippocampal formation.

---

2 A third state called small irregular activity, or SIA, is sometimes seen briefly at transitions between sleep and wakefulness, but is less well studied and will not be considered further here

In rodents, these states have clear behavioral correlates (Vanderwolf, 1969): theta is robustly associated with both locomotion and attentiveness or arousal, while LIA occurs during periods when the animal is stopped and grooming, eating, or sitting quietly.

The population bursts seen during LIA are of particular relevance to this thesis. These are highly stereotyped, lasting from 50-120 ms, and recruit ~10% of all principal neurons in areas CA1 and CA3 (Ylinen et al., 1995). Each burst recruits a random (to a first approximation, though see below) subset of principal neurons. In contrast, some interneurons fire during every population burst. Each population burst co-occurs with a single large-amplitude deflection in the LFP in area CA1, called a 'sharp wave', which is most prominent in the *stratum radiatum* and likely reflects the synaptic currents induced by a barrage of inputs from area CA3. Superimposed on the sharp wave are several (~5-10) cycles of a smaller-amplitude high frequency (150-250 Hz) oscillation, termed the 'ripple' oscillation, which is most prominent in the principal cell layer. This oscillation in the LFP likely reflects the synchronized firing of units in area CA1, many of which are phase-locked to the ongoing oscillation. The ripple oscillation is coherent across the longitudinal extent of the hippocampus, and is also observed in the deep layers of the entorhinal cortex, which are the target output structures from the hippocampus to the neocortex (Chrobak et al., 2000).

The sharp wave/ripple oscillation/population burst complex seen in the hippocampus will be referred to throughout this thesis as a sharp wave-ripple, or SWR. During quiet rest, SWRs occur without any clear periodicity at an average rate of between 0.1 and 1 times per second. Since the first report of SWRs in the literature, there have been anecdotal reports of SWRs occurring in pairs, triplets, or short 'trains' of up to 7 discrete events, occurring at roughly theta frequency (i.e. ~10Hz) (Buzsáki et al., 1983; O'Keefe and Nadel, 1978; Suzuki and Smith, 1987), however this phenomenon has not been directly studied or quantitatively described. In Chapter 3, we will argue that these trains may have a functional role in supporting memories of extended sequences of behavior.

The principal cells of the rat hippocampus are tuned to the animal's location in its environment

The recording of local field potentials and multi-unit activity in the hippocampus of the behaving animal was soon followed by the recording of the electrical activity of single cells, using newly developed transistor-based headstage amplifiers and flexible microwire electrodes.

Much work focused on the response of single units during various operant conditioning tasks (reviewed in (Vinogradova, 1970)), but an important development came in 1971, when O'Keefe and Dostrovsky reported that the firing rates of single units could also be modulated by the animal's location (O'Keefe and Dostrovsky, 1971). Along with units that fired in response to arousal, "expectation", and other sensory inputs, they reported that 8 of 76 cells recorded in the dorsal hippocampus were responsive to a combination of various sensory inputs and the animal's *location* on a small (24 x 36 cm) platform. For example, one of the reported cells fired only when the rat was touched on one shoulder while in a particular location and facing in a particular direction, and ceased firing when the lights were turned off. The authors claimed that their findings suggested that "the hippocampus provides the rest of the brain with a spatial reference map." The small number of cells reported, the haphazard methods used to discover the preferred stimuli, and the accumulated evidence that hippocampal cells coded for other behavioral parameters meant that this conclusion was initially met with profound skepticism, and was perceived as "not only wrong, but blasphemous" (Ranck and Kubie, 2008).

Subsequent work, however, confirmed this early finding, and showed that many of the principal cells of the hippocampus indeed did fire reliably each time the rat passed through a particular location in its environment. This preferred location came to be known as the cell's "place field." Each cell's place field in a particular environment was found to be stable across a period of months (Thompson and Best, 1990). A given cell may have a place field in any number of tested environments; the spatial relationship between the place fields of pairs of neurons is not preserved, however (Muller and Kubie, 1987). That is, place fields do not reflect a simple topology that is preserved across environments, but rather appear to be distributed randomly and independently in any given environment. The majority (up to 85%) of hippocampal units (as detected during sleep, anesthesia, or antidromic stimulation), are silent during wakefulness, with no place fields, at least in the environments typically used for testing (open fields of up to 1m in diameter or mazes of approximately 2m in length) (Thompson and Best, 1989). Place-related firing is seen almost exclusively during the theta state, when the rat is either running through its environment, or aroused/attentive (McNaughton et al., 1983). During LIA, place cells are mostly quiet; firing during SWR bursts is not restricted to those cells that have place fields in the current location.

The stability of individual cells' place fields, and the seemingly broad distribution of place fields across the environment lent support to O'Keefe and Dostrovsky's early conjecture that the place cell population could accurately encode the animal's location across time. Wilson and McNaughton (1993) directly tested this hypothesis by recording simultaneously from a large number (>100) of neurons in the hippocampus of the behaving animal and demonstrating that they could accurately decode the animal's location during exploration.

Despite the clear modulation of most cells' firing by "place", it is clear that hippocampal units are also tuned to other behavioral parameters (such as head direction and velocity), and to task demands (reviewed in (Eichenbaum, 2008)). In humans, single cells in the hippocampal formation have been identified that that fire in response to the presentation of objects, landmarks, and even particular individuals (reviewed in (Quiroga et al., 2008)). This data has led some researchers to argue that the hippocampal ensemble provides a more general behavioral context signal, rather than simply a spatial map.

For the purposes of the experiments described here, however, this distinction is not critical, since we are interested in whether sequences of ensemble activity are capable of being learned and re-expressed by the hippocampus. So long as the sequence of firing is reliably driven by a stimulus or a behavioral state that we can observe (e.g. the animal's location), we needn't be concerned that the same cells are also capable of coding for other parameters.

Theta and SWRs may subserve the acquisition and recall of sequential memories

Buzsáki proposed that the theta and LIA states may correspond to the two stages of a memory system (1989), with the encoding of new memories taking place during theta, and the SWR bursts observed during LIA and sleep reflecting an internally-generated re-expression of the memory trace. Activity during SWRs would be broadcast to the rest of the brain, there to be consolidated into long-term memory stores. As noted above, the possibility that reactivation and other forms of persistent activity in neural circuits might support learning at the network level. The critical advance made here was the notion that acquisition and reprocessing (specifically consolidation) of memory might occur in association with specific network states.

This theory suggests a number of testable predictions. First, we should expect coherent reactivation of hippocampal firing patterns during SWRs. The fortuitousness of the place tuning of hippocampal neurons can now be fully appreciated. Just as the spatial tuning of each

individual cell in an environment is stable over time, so is the population activity that encodes any given location stable across multiple exposures to the environment. Assuming independence of spatial tuning across environments, we can create entirely new activity patterns and sequences in the hippocampal network simply by allowing the animal to explore a novel environment. The first hints of reactivation came in the form of changes in the firing rates of individual cells (Pavlides and Winson, 1989) after experience in an environment. They found that cells were more active during sleep if the animal had explored that cell's place field during preceding behavior. Another important confirmation came with the finding that reactivation was coherent across multiple simultaneously-recorded place cells (Wilson and McNaughton, 1993). That is, the same combinatorial ensemble code used by the hippocampus to encode locations during behavior was also evident during subsequent sleep, as shown by increased coactivity of pairs of cells that were coactive during behavior. (The details of these and other reactivation studies are presented in the Chapter 2).

The two-stage model predicts not only that coherent activity patterns will be re-expressed during off-line states, but that *sequences* of such patterns will also be re-expressed. Again, the place cell system provides a powerful means of addressing this question experimentally. By placing rats on linear tracks, the experimenter can restrict the sequence of ensemble patterns to a simple linear sequence.[3] This prediction has been borne out—temporally structured firing sequences (known as 'replay') have now been observed in association with SWRs during slow-wave sleep (Lee and Wilson, 2002), and during pauses in exploration (Csicsvari et al., 2007; Diba and Buzsáki, 2007; Foster and Wilson, 2006).

SWR-associated replay has some remarkable and unexpected properties. It is temporally compressed by approximately 20-fold—this may allow for temporal structure as experienced on the behavioral timescale (seconds) to be re-expressed at a timescale appropriate for the expression of synaptic plasticity (10s of milliseconds) (Mehta et al., 2002). Furthermore, sequences observed during awake SWRs can reflect replay of the ensemble sequence in the same order that those places were experienced, or the reverse order (Diba and Buzsáki, 2007; Foster and Wilson, 2006).

---

3   The open field also provides constraints on possible ensemble activity sequences, since the rat must always move between contiguous locations, but the sheer number of possible sequences makes detection of reactivation impractical.

Motivation for the current work: Assessing the behavioral relevance of replay

To date, the study of replay has been largely observational, with no direct demonstration of the behavioral or cognitive role for the phenomenon. The two-stage model of memory predicts that disruptions of SWRs will prevent effective consolidation of the memories being reactivated, but unfortunately these critical experiments have not yet been performed.

We can, however, take a complementary, neuroethological approach to the question of the behavioral relevance of replay. If we assume that the measured neural activity reflects (or is at least correlated with) the mental representations used by the animal in the performance of spatial navigation (Gallistel, 1990), we can ask whether the proposed models of replay could support natural behaviors of which we know the animal to be capable. In the case of the two-stage model of memory formation, this approach immediately reveals a possible limitation of the model, and suggests additional experiments. As we have noted, SWR-associated replay has been found to evolve approximately 20 times faster than behavior, and SWRs are on the order of 100 ms in duration. Given a running speed of 0.5 m/s, this means that the replay seen during a single SWR should recapitulate approximately 1 m of behavior. All previous work studying replay has taken place in open fields of less than 1 meter in diameter, or on linear tracks of 1-4 m. Rats in the wild, however, live communally in complex tunnel systems(Boice, 1977; Calhoun, 1963), and roam over tens, or even hundreds of meters, navigating successfully between multiple nests and feeding sites (Davis et al., 1948; Jackson, 1982). How are memories for long behavioral sequences represented in the brain?

Buzsáki has suggested (1989) that the SWR may represent a fundamental unit, or 'quantum' of information exchange between the hippocampus and neocortex (Buzsáki, 1989). If SWR-associated replay is thought to subserve spatial navigation, we might ask how memories for longer sequences of behavior are mapped onto these events. Are long sequences further compressed in time to fit into a single SWR? Or is the compression factor constant, with longer sequences 'chunked' into shorter subsequences? Would such hypothetical chunking be random, or might it relate to regularities in the animal's behavior in the environment?

We address these questions in the current thesis by studying hippocampal replay in rats as they freely explore a 10 meter-long track. We are motivated by a desire to use this task to 'break' existing models of replay in an interesting way, and also to study this putative spatial navigation system during a more behaviorally relevant model task. Some previous work has examined the

behavior of place cells in larger environments[4]. However, no work has examined off-line reactivation of place cells in larger environments. We further adopt the ethological approach by studying animals early in their exposure to these larger environments (during the first few exposures to the track), while they are still exhibiting irregular exploratory behavior, and while they may therefore still be expected to be learning about a novel spatial environment.

The experiments in this thesis are motivated by existing theories of hippocampal function, but they are primarily observational. That is, they do not aim to directly test the predictions of any particular theory, but rather attempt to characterize the behavior of the hippocampal ensemble during a previously unobserved, ethologically relevant behavior: the exploration of a large environment. As is hopefully apparent from the historical overview just presented here, such open-ended observation, guided by theory, can provide important insights into the system under study.

Summary of the thesis

In Chapter 2, we review existing methods for detecting reactivation in neural ensembles, and develop a new approach to this problem that involves decoding the activity of the ensemble, then looking for patterns in the domain of the decoded stimulus. We present data showing that this method is capable of detecting structured replay during off-line states.

In Chapter 3, we apply the approach developed in chapter 2 to recordings taken from the hippocampus of rats as they freely explore a 10 meter-long track. We make several novel contributions: we show that replay proceeds at a relatively constant 'virtual' velocity; that it can proceed over trajectories as long as the complete environment; that this extended replay spans trains of closely-spaced SWRs; and that replay can begin at locations remote from the animal.

In Chapter 4, we consider various models that might account for the extended replay observed in a large environment, and attempt to place these results into the context of work on the cognitive functions supported by the hippocampus in humans.

---

4   On an 18 meter-long linear track, the size of place fields of cells in area CA3 were found to scale with dorsal-ventral location within the hippocampus (Kjelstrup et al., 2008); in a 1.4 by 1.5 m enclosure, place cells in CA1 were found to exhibit multiple peaks (Fenton et al., 2008).

References:

Boice, R. (1977). Burrows of wild and albino rats: effects of domestication, outdoor raising, age, experience, and maternal state. J. Comp. Physiol. Psychol. *91,* 649-661.

Buckner, R.L., and Carroll, D.C. (2007). Self-projection and the brain. Trends Cogn. Sci. *11,* 49-57.

Buzsáki, G. (1989). Two-stage model of memory trace formation: a role for "noisy" brain states. Neuroscience *31,* 551-570.

Buzsáki, G., Leung, L.W., and Vanderwolf, C.H. (1983). Cellular bases of hippocampal EEG in the behaving rat. Brain Res. Rev. *6,* 139-171.

Cahill, L., and McGaugh, J.L. (1996). Modulation of memory storage. Curr. Opin. Neurobiol. *6,* 237-242.

Calhoun, J.B. (1963). The ecology and sociology of the Norway rat (Bethesda, Md.: U.S. Dept. of Health, Education, and Welfare, Public Health Service; for sale by the Superintendent of Documents, U.S. Govt. Print. Off.).

Chrobak, J.J., Lorincz, A., and Buzsáki, G. (2000). Physiological patterns in the hippocampo-entorhinal cortex system. Hippocampus *10,* 457-465.

Corkin, S. (2002). What's new with the amnesic patient H.M.? Nat. Rev. Neurosci. *3,* 153-160.

Corkin, S., Amaral, D.G., Gonzalez, R.G., Johnson, K.A., and Hyman, B.T. (1997). H. M.'s medial temporal lobe lesion: findings from magnetic resonance imaging. J. Neurosci. *17,* 3964-3979.

Csicsvari, J., O'Neill, J., Allen, K., and Senior, T. (2007). Place-selective firing contributes to the reverse-order reactivation of CA1 pyramidal cells during sharp waves in open-field exploration. Eur. J. Neurosci. *26,* 704-716.

Davis, D.E., Emlen, J.T., and Stokes, A.W. (1948). Studies on Home Range in the Brown Rat. J. Mammal. *29,* 207--225.

Diba, K., and Buzsáki, G. (2007). Forward and reverse hippocampal place-cell sequences during ripples. Nat. Neurosci. *10,* 1241-1242.

Eichenbaum, H. (2008). Hippocampal neuronal activity and memory: should we still be talking about place cells? In Hippocampal place fields : relevance to learning and memory, Mizumori, S. J. ed., (Oxford ; New York: Oxford University Press) pp. 161--174.

Fenton, A.A., Kao, H.Y., Neymotin, S.A., Olypher, A., Vayntrub, Y., Lytton, W.W., and Ludvig, N. (2008). Unmasking the CA1 ensemble place code by exposures to small and large environments: more place cells and multiple, irregularly arranged, and expanded place fields in the larger space. J. Neurosci. *28,* 11250-11262.

Foster, D.J., and Wilson, M.A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. Nature *440,* 680-683.

Gallistel, C.R. (1990). The organization of learning (Cambridge, Mass.: MIT Press).

Hassabis, D., Kumaran, D., Vann, S.D., and Maguire, E.A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. Proc. Natl. Acad. Sci. U. S. A. *104,* 1726-1731.

Hebb, D.O. (1949). The organization of behavior; a neuropsychological theory (New York: Wiley).

Jackson, W.B. (1982). Norway Rat and Allies. In Wild Mammals of North America, Chapman, J. A. and Feldhamer, G. A. eds., (Baltimore, Maryland: Johns Hopkins University Press) pp. 1077-1088.

Kjelstrup, K.B., Solstad, T., Brun, V.H., Hafting, T., Leutgeb, S., Witter, M.P., Moser, E.I., and Moser, M.B. (2008). Finite scale of spatial representation in the hippocampus. Science *321,* 140-143.

Lechner, H.A., Squire, L.R., and Byrne, J.H. (1999). 100 years of consolidation--remembering Muller and Pilzecker. Learn. Mem. *6,* 77-87.

Lee, A.K., and Wilson, M.A. (2002). Memory of sequential experience in the hippocampus during slow wave sleep. Neuron *36,* 1183-1194.

Lorente de Nó, R. (1938). Analysis of the activity of the chains of internuncial neurons. J. Neurophysiol. *1,* 207--244.

Marr, D. (1971). Simple memory: a theory for archicortex. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *262,* 23-81.

McGaugh, J.L. (1966). Time-dependent processes in memory storage. Science *153,* 1351-1358.

McNaughton, B.L., Barnes, C.A., and O'Keefe, J. (1983). The contributions of position, direction, and velocity to single unit activity in the hippocampus of freely-moving rats. Exp. Brain Res. *52,* 41-49.

Mehta, M.R., Lee, A.K., and Wilson, M.A. (2002). Role of experience and oscillations in transforming a rate code into a temporal code. Nature *417,* 741-746.

Muller, R.U., and Kubie, J.L. (1987). The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. J. Neurosci. *7,* 1951-1968.

O'Keefe, J., and Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain Res. *34,* 171-175.

O'Keefe, J., and Nadel, L. (1978). The hippocampus as a cognitive map (Oxford; New York: Clarendon Press; Oxford University Press).

Pavlides, C., and Winson, J. (1989). Influences of hippocampal place cell firing in the awake state on the activity of these cells during subsequent sleep episodes. J. Neurosci. *9,* 2907-2918.

Quiroga, R.Q., Kreiman, G., Koch, C., and Fried, I. (2008). Sparse but not 'grandmother-cell' coding in the medial temporal lobe. Trends Cogn. Sci. *12,* 87-91.

Ranck, J.B., and Kubie, J.L. (2008). Historical Perspective: Place Cells in Ann Arbor and Brooklyn. In Hippocampal place fields : relevance to learning and memory, Mizumori, S. J. ed., (Oxford ; New York: Oxford University Press) pp. xvii--xx.

Schacter, D.L., Addis, D.R., and Buckner, R.L. (2007). Remembering the past to imagine the future: the prospective brain. Nat. Rev. Neurosci. *8,* 657-661.

Scoville, W.B., and Milner, B. (2000). Loss of recent memory after bilateral hippocampal lesions. 1957. J. Neuropsychiatry Clin. Neurosci. *12,* 103-113.

Scoville, W.B., and Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. J. Neurol. Neurosurg. Psychiatry. *20,* 11-21.

Squire, L.R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. Psychol. Rev. *99,* 195-231.

Suzuki, S.S., and Smith, G.K. (1987). Spontaneous EEG spikes in the normal hippocampus. I. Behavioral correlates, laminar profiles and bilateral synchrony. Electroencephalogr. Clin. Neurophysiol. *67,* 348-359.

Thompson, L.T., and Best, P.J. (1990). Long-term stability of the place-field activity of single units recorded from the dorsal hippocampus of freely behaving rats. Brain Res. *509,* 299-308.

Thompson, L.T., and Best, P.J. (1989). Place cells and silent cells in the hippocampus of freely-behaving rats. J. Neurosci. *9,* 2382-2390.

Tulving, E. (1985). Memory and Consciousness. Canadian Psychology *26,* 1--12.

Vanderwolf, C.H. (1969). Hippocampal electrical activity and voluntary movement in the rat. Electroencephalogr. Clin. Neurophysiol. *26,* 407-418.

Vargha-Khadem, F., Gadian, D.G., Watkins, K.E., Connelly, A., Van Paesschen, W., and Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. Science *277,* 376-380.

Vinogradova, O.S. (1970). Registration of information and the limbic system. In Short-term changes in neural activity and behaviour, Horn, Gabriel and Hinde, Robert A. eds., (Cambridge, England: University Press) pp. 95--140.

Wilson, M.A., and McNaughton, B.L. (1993). Dynamics of the hippocampal ensemble code for space. Science *261*, 1055-1058.

Ylinen, A., Bragin, A., Nadasdy, Z., Jando, G., Szabo, I., Sik, A., and Buzsáki, G. (1995). Sharp wave-associated high-frequency oscillation (200 Hz) in the intact hippocampus: network and intracellular mechanisms. J. Neurosci. *15*, 30-46.

Chapter 2:

Detecting neuronal reactivation
in stimulus space

Summary

In this chapter we will present a novel approach to the detection of neuronal reactivation in ensembles of neurons (Davidson et al., 2007; Davidson et al., 2008; Davidson et al., 2009). Reactivation, as discussed in the introductory chapter, is for our purposes defined as the sequential expression of activity patterns in a neural ensemble during an 'off-line' state (such as sleep or quiet rest) that recapitulates the order of activity of that same population of cells during an earlier stimulus-driven, or 'on-line' state. Interest in this phenomenon is driven by the fact that such reactivation represents a physiological memory for learned behavioral sequences. While the approach described is applicable to any population of cells with a reliable response to observable parameters, we will develop the method in the context of the hippocampal place cell system.

Most existing approaches to detecting reactivation consider *the temporal relationships between the spikes emitted by different neurons in an ensemble*, and develop measures to quantify similarity of firing patterns across off-line and on-line states. We refer to this class of methods as operating in 'spike space'. A canonical example of such a measure is the temporal cross-correlation function between the firing of two place cells. Higher-order methods that account for the joint firing of larger numbers of cells have also been developed and are discussed in detail below.

In contrast, the class of methods that we refer to as 'stimulus-space' methods rely on first decoding the neural activity observed during the off-line state, resulting in a time-varying estimate of the encoded stimulus value. In this space, *reactivation appears as a similarity between this estimate of the stimulus and the stimuli actually experienced by the animal*. In the case of the place cell system, for example, we can directly compare the sequence of decoded locations during an off-line state with the animal's preceding behavior.[1]

This approach is appealing for several reasons: it is a powerful dimension-reduction strategy that deals gracefully with large datasets recorded from neural ensembles; it builds on a rich body of work in neural decoding that provides a formal statistical framework for interpreting the activity of the ensemble; and, since reactivation events are described in the same domain as

---

1   Many other forms of structured activity during an off-line state could be detected using variants the methods described here. For instance, in a 2-dimensional arena, a trajectory corresponding to behaviorally-relevant, but never before experienced path through the environment would not meet our definition of 'reactivation', but may be of great interest.

the animal's behavior, it lends itself readily to the formulation and testing of hypotheses about the cognitive roles of reactivation. The specific methods developed below also provide a solution to some practical problems with the application of existing spike-space methods to the analysis of place cell activity in large environments (notably those arising from complex, multimodal tuning curves). Some limitations of the approach are its dependence on strong assumptions about the coherence of the neural representation across on-line and off-line states, and a reduced power to test hypotheses about activity patterns at the cellular level (e.g. the induction of activity-dependent plasticity between co-active cells).

In this chapter, we will begin by discussing some practical considerations in the design of experiments to study neuronal reactivation in the place cell system. This is followed by a review of existing spike space methods as they have been applied to this class of data. We will then introduce and demonstrate stimulus space methods using rat hippocampal data.

Experimental design considerations

The assessment of reactivation requires dividing the experimental data into two non-overlapping sets: an on-line state during which sequences of externally-driven activity are observed; and an off-line state, during which the network re-expresses internally-generated patterns that may or may not be similar to those observed during the on-line state. We will turn to the question of how to assess the similarity of firing patterns across states below; the first task, however, is to design experiments that allow for the expression of sufficiently complex patterns of network activity during the on-line state, and that ensure adequate sampling of the off-line states during which reactivation is hypothesized to occur.

On-line and off-line states have been defined using both behavioral and electrophysiological criteria. Place cells in the behaving rat fire in a place-selective manner most strongly while the rat is engaged in active exploration (McNaughton et al., 1983), so the on-line state is usually defined as periods of time when the rat is moving faster than a certain speed threshold, typically 10-20 cm/s. The prominent theta oscillation in the hippocampal LFP during exploration could also be used to define this state, but periods of running are generally preferred, since they are simpler to calculate reliably across subjects, and they represent the periods during which spatial sampling is taking place (even if they omit some periods of theta during

immobility).

The off-line state has a complementary requirement that the animal not be actively exploring, so that any observed change in the firing of place cells can not be explained by changes in sensory input drive. Much work to date has focused on reactivation during sleep (Lee and Wilson, 2002; Louie and Wilson, 2001; Nadasdy et al., 1999; Pavlides and Winson, 1989; Wilson and McNaughton, 1994), with the stages of sleep scored according to established electrophysiological and electromyographical criteria. In studies of sleep reactivation, the neural ensemble is typically recorded while the animal sleeps in a small space (the 'sleep box') both before and after the animal performs a spatial task in a separate, dedicated arena. Other studies, including those described in this thesis, have studied reactivation during stopping periods while the animal is awake (Csicsvari et al., 2007; Diba and Buzsáki, 2007; Foster and Wilson, 2006; Johnson and Redish, 2007; Karlsson and Frank, 2009).

Neural activity during off-line states can be heterogeneous, raising the possibility that reactivation may be restricted to discrete, identifiable epochs. For example, as described in Chapter 1, the large irregular activity (LIA) state observed while animals pause on the track is characterized by occasional bursts of activity occurring against a quiescent background. Most of the methods discussed below therefore rely on identifying and analyzing only brief 'candidate' periods. In this way, they greatly increase computational tractability, and potentially sensitivity as well (by excluding periods that would otherwise represent a large population of false negatives). Some criteria that have been employed in the identification of candidate events include selecting periods with elevated firing across the entire ensemble of the recorded neurons (i.e. population bursts), or only examining firing surrounding electrophysiological events of interest, such as the sharp wave-ripple (SWR) in hippocampal area CA1.

The nature of the spatial task given to the animal also has significant implications for subsequent reactivation analysis. Of particular importance is the distinction between tasks in 1-dimensional and 2-dimensional environments, i.e. between tracks and open arenas. A typical track experiment involves a rat running back and forth from one end of a linear track to the other for a food reward, whereas in the open field, food reward is typically scattered randomly to encourage foraging. Rats in the wild navigate intricate burrows, and use stereotyped 'runways' to travel between their nests and foraging sites (Boice, 1977; Calhoun, 1963; Jackson, 1982), so

Chapter 2 | 31

both tasks may represent species-typical behaviors. The use of a linear track affords important practical advantages for the study of reactivation. First, when animals run back and forth on a track, they provide even sampling of the space, whereas behavior in the open field can be more variable, with biases towards certain locations (such as the edges of the arena). Second, the sequence in which the cells in the ensemble are activated on each lap of the track is held constant, and can therefore be used as a simple template for reactivation analysis.

The degree of experience that the animals have with the spatial task on which they are being trained is also an important experimental variable that deserves consideration, though these effects are not well-studied. Well-trained animals on a linear track will run swiftly back and forth at a constant speed; such uniform behavior may be particularly useful when attempting to analyze data taken across animals and sessions. On the other hand, such behavior may be fundamentally different from the hesitant, uneven behavior exhibited by rats in their initial exploration of a new environment.

'Spike space' methods for the study of reactivation

To date, most published studies of reactivation in the hippocampus have defined the phenomenon in terms the firing patterns of individual neurons, or in 'spike space'. We will review these in order of the complexity of the patterns detected, beginning with first-order methods that consider only the activity of single neurons, and continuing to higher order (typically template-based) methods that consider activity across a larger number of cells. (The statistical grounding of these methods was reviewed by Brown et al. (2004), though not strictly in the context of the detection of reactivation.)

Single-cell, pairwise, and 'triplet-wise' correlation measures

Pavlides and Winson (1989) were the first to show that place cell activity during behavior could affect that seen during subsequent sleep, using a simple measure of similarity across behavioral states: a within-cell correlation in firing rate. They recorded simultaneously from pairs of place cells with non-overlapping place fields in either an open field or an eight-arm radial arm maze; during each experimental session, the rat was allowed to explore an area containing only one cell's place field. Using a counterbalanced design, they found that the cell

that was driven by behavior had higher firing rates during subsequent slow-wave and REM sleep than did the non-driven cell, a result which they interpreted as evidence for the processing of recently acquired memories during sleep. (A small, but not significant increase was also observed for 'still alert', and 'quiet wakeful' states, during which the animal was stopped but not asleep.) Since it only considered one active place cell per session, their analysis did not support the finding of coherent reactivation across the population, or of any temporal structure.

Wilson and McNaughton (1994) addressed the issue of the coherence of the representation across the ensemble during off-line states through pairwise analyses of place cell firing. They recorded simultaneously from 50-100 cells during exploration of an open field or linear track, as well as during sleep periods both before and after the exploration. They separately considered pairs of cells with either overlapping or non-overlapping place fields during exploration, and found that pairs of cells with overlapping place fields exhibited an increase in sleep co-firing (within 200 ms) after exploration. Pairs of cells with non-overlapping fields did not exhibit such an increase. Importantly, all of the cells studied were active during behavior, so would have been predicted (based on the Pavlides and Winson result presented above) to have exhibited elevated firing. The novel result is that cells coding for overlapping locations during behavior are also reactivated together.

Skaggs and McNaughton (Skaggs et al., 1996) extended these results to the temporal ordering of pairs of place cells during behavior and sleep. Rather than considering the locations of each cell's place fields, they calculated a temporal bias measure from the central 400 ms of the cross-correlation function for each pair. This measure was calculated for each pair in the same way during behavior and sleep, and the size and sign of the bias during behavior (which was assumed to reflect the spatial ordering of place fields) was compared with the size and sign of the same measure calculated during sleep periods. Pairwise temporal bias during behavior was found to correlate with bias during subsequent sleep, but not preceding sleep. The interpretation given to this data was that reactivation of hippocampal ensembles preserved the temporal ordering in which those ensembles were activated during experience.

One implicit assumption underlying this analysis is that temporally structured reactivation proceeds in the same order as experience. Later studies (Diba and Buzsáki, 2007; Foster and Wilson, 2007), including those reported in this thesis, have shown that temporally structured

reactivation can proceed in the reverse order during off-line states. A mixture of equal parts forwards and reverse-ordered reactivation would not be detected by the correlation measure used by Skaggs and McNaughton, since it would obscure any trend in the sign of temporal biases during sleep. Diba and Buzsáki (2007) developed a pairwise measure that is sensitive to reactivation in both orderings. For each pair of cells, they calculate: the distance separating the two cells' place fields; and the typical time lag between their firing during reactivation epochs (measured as the peak of the pairwise cross-correlogram). A correlation between these two values indicates that the firing structure during reactivation follows the structure of spatial experience during behavior. These authors only applied this analysis to determine a 'compression factor' for events that had already been shown to contain significant replay (see below), but such a measure could also be used as a method of detecting reactivation.

Nadasdy et al. (Nadasdy et al., 1999) studied the occurrence of ordered triplets of spikes emitted by groups of 3 neurons. One of the methods employed, based on the joint peri-stimulus time histogram (JPSTH), showed that many triplets re-occurred more frequently than expected by chance, even after controlling for the known pairwise correlations between the members of the triplet, suggesting that higher-order sequences are learned and re-expressed in the hippocampus. One limitation of this study was that the behavioral task involved running on a stationary wheel, rather than actively exploring an environment; the behavior of place cells in this environment is not well understood. Also, as in Skaggs and McNaughton's pairwise method (1996), templates were generated from the short-timescale firing order of cells observed during running, rather than from the ordered spatial configuration of place fields, so the sequence learned by the network can not be said to reflect a behaviorally-determined sequence.

Higher-order methods using spike pattern prototype matching

Temporal structure across pairs or triplets of neurons might reflect even higher-order temporal structure, possibly even the re-expression of memory for extended episodes. Noisy or degraded reactivation, or reactivation on longer timescales may also not be easily detectable by lower-order methods. These considerations have motivated the development of methods that take advantage of data recorded simultaneously from large populations (typically tens) of neurons. The approaches described in this section can all be characterized as relying on the generation of a

template spike sequence corresponding to the ordering that would be expected during a particular behavior, such as the traversal of a linear maze; this prototype sequence is then compared to the observed candidate reactivation events. The various means of constructing a prototype, of comparing observed activity to the prototype, and of assessing the statistical significance of the degree of matching, are described below.

Louie and Wilson (Louie and Wilson, 2001) studied reactivation in REM sleep after a rat ran on a circular track for food reward. The task performed by the animal in this study was designed to elicit a temporally complex pattern of firing while maintaining a simple track topology for the purposes of spatial sampling of place fields: the rat was trained to run around the circle in three-quarter circle increments, always beginning and ending at one of four food wells distributed evenly around the circle. Between 8 and 13 place cells with fields on the track were simultaneously recorded in each session. Continuous running of this task thus resulted in a fixed sequence of four overlapping, but distinct patterns of place cell activity every three full laps of the track. In order to detect reactivation of this complex behavioral sequence, the authors constructed a template corresponding to the actual pattern of firing observed across the recorded cells during the full sequence. Re-expression of the behavioral sequence was tested by calculating a correlation measure between this template and observed activity during REM sleep. Correlations were also calculated for time-dilated and time-contracted versions of the template, with the peak correlation found for a 1.4-fold expansion of the template. Statistical significance was tested by repeating the template-correlation analysis on many randomly shuffled versions of the REM spiking records, giving an estimate of the chance level for the measure.

Previous studies had shown that static features of the encoded environment (e.g. whether place fields were overlapping) could be reflected in later reactivation, but this study provided the first example of reactivation corresponding to a complex behavioral sequence. This study also argues for the utility of higher-order methods in the analysis of noisy neural data—earlier studies had failed to detect correlation effects in REM sleep periods (Kudrimoti et al., 1999).

Louie and Wilson's method relies on detecting modulation of firing rates across long stretches of activity, but might be expected to fail to detect reactivation that reflected the order in which cells fired, but did not recapitulate precise timing relations. While the preservation of timing relations during reactivation is of interest, analyses that consider only relative order of

activation are also of interest, since such sequences would still reflect the basic structure of the animal's spatial experience.

The method of Lee and Wilson (2002; 2004) explicitly examines the relative firing order of cells. First, a firing order template is constructed by ranking cells based on the location of the peak of their place fields along a linear track. Candidate reactivation events are then identified (as bursts of activity) and scored according to the similarity of the observed firing order to the template order. Combinatorial methods are then used to calculate the likelihood of observing a match by chance that is at least as 'good' as the observed match. Since events are scored on two parameters (length of match and number of gaps), there is not a unique function for determining 'goodness', but the authors identify several reasonable scoring functions that emphasize different features of the hypothesized reactivation. One advantage of this method over the template method of Louie and Wilson is that it allows the calculation of precise P-values for individual candidate reactivation events without the use of Monte Carlo methods (under various null hypotheses, such as the assumption that any ordering of cells is equally likely). Ji and Wilson (2007) applied this method to reactivation of both hippocampal and cortical sequences, with some modification: the order of spiking during candidate events was determined by peak firing rate (after smoothing), rather than by the time of the first spike; they also used a modified matching function.

Two later studies, by Foster and Wilson (2006) and Diba and Buzsáki (2007), employed variants of this method. They also created a template based on a ranked list of place cells based on place field location along a linear track. They used Spearman's rank-order correlation between each observed sequence and the original probe sequence as the similarity measure.[2] Notably, both of these studies found significant *negative* correlation between the original cell ranking and the observed order, a phenomenon that Foster and Wilson termed 'reverse replay'. Since in these experiments rats are running repeatedly back and forth on a linear track, a negative correlation might reflect forward-ordered reactivation of a return trip along the just-run section of track. Foster and Wilson tested this by re-analyzing these events after omitting bidirectional place cells, that is, using place cells that fired only when the rat ran through them in one of the two directions (50-65% all place cells in their data). Many more sequences maintained significant reverse-

---

2   As Lee and Wilson (2004) note, this is equivalent to the use of a matching rule with their method that takes into account the magnitude of mismatches during 'gaps'.

rather than forward-ordering in this condition. Diba and Buzsáki performed a similar analysis, and also explicitly tested for correlation with probe sequences corresponding to both the preceding and upcoming directions.

The above methods all rely on the use of a linear track to constrain place cell sequences to a simple set. Csicsvari et al. (2007) developed an elegant method to study reactivation of recent experience in the open field. Rather than constructing a template based on the expected firing order of place cells due to a particular trajectory, they simply compared firing order during a population burst with the recent firing history of place cells. The group of cells firing in a candidate reactivation event were sorted by their most recent activity during running, and rank-order statistics as described above were calculated. The authors found significant negative correlations, suggesting that immediately-preceding trajectories in the open field are replayed in the reverse order. This finding is notable since trajectories during exploration in the open field are likely to be unique; specific replay of these trajectories can therefore be interpreted as evidence for reactivation of memory for specific episodes, rather than for a spatial sequence that has been experienced many times. Since this method does not rely on determining the place-related firing of the cells involved, however, the authors were not able to show the relative contribution of directional and non-directional place cells, and the sequence replay they observe may not be reverse-ordered (e.g. it may reflect forward- or non-ordered replay of a sequence proceeding back over the recently traveled path). Another limitation of this method is that it can only measure reactivation of immediately preceding activity, and is therefore unsuitable for the study of sleep. It also would be unable to detect reactivation corresponding to upcoming sequences (as described above) or of sequences corresponding to remote locations.

'Stimulus space' methods for the detection of reactivation

Since the first recordings from individual place cells, it has been recognized that the stability and apparently random distribution of their firing fields could result in a precise ensemble code for space (O'Keefe and Nadel, 1978). Wilson & McNaughton demonstrated this by recording simultaneously from large numbers (>100) of place cells in the behaving rat, then decoding the animal's location from the time-varying population activity (Wilson and McNaughton, 1993). The place cell system has since become a testing ground for the

development of neural decoding algorithms (Barbieri et al., 2004; Brown et al., 1998; Brown et al., 2004; Zhang et al., 1998); the resulting methods have also been used productively to study aspects of ensemble activity during behavior (Fenton and Muller, 1998; Fenton et al., 2008; Fyhn et al., 2004; Jensen and Lisman, 2000; Skaggs and McNaughton, 1998).

The finding that the hippocampal code for space is also coherently expressed during off-line states (Wilson and McNaughton, 1994) suggested that decoding methods might also be a promising approach to the detection of structured reactivation. This approach was prefigured in a study by Georgopoulos et al. (1989). The authors decoded reach direction in populations of neurons in motor areas of cortex in monkeys, and showed that the decoded movement vector evolved over time during a task in which the subject had to mentally rotate its reach direction before initiating movement. Citing Georgopoulos, Wilson and McNaughton (1993) suggested that the "interpretation of neuronal ensemble activity in a behaving animal … opens the possibility, in principle, of the interpretation of neuronal activity in the absence of explicit behavior, such as during periods of sleep, motor planning, reasoning, or motor consolidation." This potential has also been pointed out in later work: Zhang et al. (1998) observed brief discontinuities in their estimates of position during behavior (sometimes associated with ripple oscillations) and suggested that they might reflect the firing of 'view cells', or a memory consolidation process; Brown et al. (1998) noted that their decoding paradigm "suggests a quantitative way to study the patterns of place cell reactivation during sleep and, hence, to investigate the two-stage hypothesis of information encoding into long-term memory."

While appealing in principle, the application of decoding methods to off-line states presents several challenges. The central issue is that no ground truth is available against which to measure the performance of our decoding attempts. Instead, we must rely on observations of structure in the decoded signal as evidence that we have succeeded in our decoding.[3] This is a delicate foundation on which to build, and as we proceed we will therefore pay careful attention to assumptions or methods that may impose spurious structure in the decoded signal, and carefully control for their contributions.

---

3 Redish and colleagues (Jackson and Redish, 2003; Johnson et al., 2005) have developed "coherencey" measures that capture the self-consistency of activity in an ensemble, which could be useful in characterizing putative reactivation events. The probabilistic methods we develop below also indirectly capture self-consistency in ensemble activity, with inconsistent firing patterns leading to less sharply-tuned probability density estimates, which in turn result in lower 'replay scores' for individual candidate events.

Two groups have published studies using decoding methods to examine the dynamics of hippocampal place ensemble reactivation. Johnson and Redish (2007) recorded from ensembles of 33-72 cells in area CA3, and decoded firing activity in short (20 ms) time bins while animals were stopped at choice points on mazes with one or more 'T' junctions. They described 'sweeps', or trajectories, in decoded position, that traveled down one or the other of the available arms. They interpreted their results in terms of the evaluation of potential future behaviors (Tolman's 'vicarious trial and error'.) Johnson and Redish did not control for the possibility that the sequential structure of the trajectories they observed could have arisen by chance (e.g. by comparing their results to those obtained with shuffled or synthetic data). This omission was compounded by the fact that they did not use an unbiased method of selecting candidate trajectories (these were identified by visual inspecting decoded events and selecting those thought to contain temporally structured trajectories), and by the choice of a decoding algorithm that incorporated a history term which might be expected to impose sequential structure even in the absence of structure in the ensemble activity (see below).

Karlsson and Frank (2009) used a decoding approach to study SWR-associated reactivation of remote environments during exploration. They recorded from areas CA1 and CA3 while rate explored two mazes in the same room in succession, and showed that temporally structured replay of a recently experienced maze while the rat was stopped on a second maze. Their methods were based in part on our previously presented work: they used a similar decoding approach to the one described below, but decoded two environments simultaneously. They employed a simple method for testing the significance of the observed replay: a linear regression on the point estimates obtained for each 20 ms time bin during a candidate event.[4]

In the rest of this chapter, we will first discuss the choice of a decoding approach that is appropriate for application to activity during an off-line state with different neural dynamics from the on-line epoch used for training. We will then report on a method we developed for detecting, and quantifying the observed reactivation events. We will illustrate this discussion with ensemble data recorded from area CA1 during exploration of a 10 meter-long linear track

---

4 Karlsson and Frank used this measure only to identify a significant linear dependence of estimated position on elapsed time for each event. Unlike the trajectory-finding method we develop below (which resembles a 'robust regression'), the simple linear regression they employed was not suitable for characterizing the detailed structure of reactivation events: since the dependent variable (decoded position) is bounded (by the edges of the track), the error about the regression line can be shown to not be normally distributed, and to be correlated across time. Practically speaking, the slope of the regression line will tend to be an underestimate of the true slope.

(this is the 'rat 1' dataset described in Chapter 3).

Choice of decoding algorithms

Since our approach relies on the assumption that the same coherent neural representations are re-expressed by the brain during off-line states as during behavior, we will choose a decoding approach that has been shown to perform well during on-line states. We will then validate the performance of this algorithm during behavior in our hands. Lastly, we will extend its use to the decoding of off-line states.

Probabilistic methods for the decoding of neural populations (Foldiak, 1993; Oram et al., 1998; Sanger, 1996) are robust and efficient. Indeed, it can be shown that they make optimal use of the information available in neural spike trains, given certain reasonable assumptions. A further advantage is that the output of such methods is a probability distribution over the stimulus.[5]

The general approach is to first, in the *encoding* stage, fit a model of how a biological signal drives spiking activity in each of the neurons being recorded. During *decoding*, through the application of Bayes' theorem, an estimate of the original signal is derived from the observed neural spiking. In our application, the encoding stage is all of the periods when the rat is running; the decoding stage is the resting periods between these runs; and the encoded signal is the animal's location. The spiking model that we will use assumes that spikes from each cell are emitted at random intervals at a rate that varies with the animal's location (an inhomogeneous Poisson process).[6]

Another advantage of working within this a formal statistical framework is that additional dependencies and regularities can be modeled, stated formally and included in the decoding process. For instance, incorporating the dependence of spiking on the phase of the ongoing theta oscillation phase has been shown to improve decoding performance (Brown et al., 1998; Jensen

---

5   See Sanger (2003) for a highly accessible primer on the statistical underpinnings of the probabilistic approach; see Brown et al. (2004) for a review of applications and recent advances; and see Zhang et al. (1998) for an account of various linear and probabilistic decoding methods as they have been applied specifically to hippocampal place cells.)

6   This assumption is clearly not completely valid: for instance, cells have refractory periods, firing rates are known to oscillate with hippocampal theta (O'Keefe and Recce, 1993); and spiking also exhibits higher pass-to-pass variance than predicted (Fenton and Muller, 1998). The model has, however been shown empirically to result in good decoding performance, especially when averaging over longer observation times.

and Lisman, 2000). Regularities in the behavior of the rat (and indeed in the physical world) can also be used to improve decoding. For example, if the rat spends more time in certain locations in the environment, this fact can be expressed as a prior likelihood on the animal's location. Similarly, a 'continuity constraint' can be imposed on adjacent position estimates such that the previous estimate is updated as new spikes arrive. In this approach, decoding can be viewed as an ongoing filtering process, with new spiking information used to update an existing estimate, according to a model of the rat's expected behavior (often modeled as a random walk). The use of such a constraint has been shown to significantly improve decoding performance during behavior (Zhang et al., 1998).

For our purposes, however, we must proceed carefully when introducing such additional dependencies. For instance, the theta oscillation is not present during the states we wish to decode, so this parameter will not be useful during decoding. We also do not wish to assume that frequently-visited locations are more or less likely to be represented in a reactivated state. As we noted above, we particularly wish to guard against the possibility that our decoder might impose temporal structure on an our reconstructed signal. The use of a continuity constraint is therefore potentially problematic, since it would temporally 'smooth' the estimated position, and may result in apparent smooth trajectories through stimulus space when there are actual discontinuities in the encoded signal. That is, we can not assume that the dynamics of the decoded signal are the same across the encoding and decoding periods. Rather, we wish to test exactly this hypothesis.

Another choice we faced was how to model the location-dependence of each cells' firing rate. Brown and colleagues have achieved good results fitting various models of place field shape (Barbieri et al., 2004), and have shown that such parameterization is useful for examining the evolution of place field shape and location (Frank et al., 2002). As discussed below, we found that on our 10 meter-long track, many cells had highly complex spatial and directional tuning that were not amenable to fitting with any simple models. We therefore followed Zhang et al. (1998) and relied on an empirical estimate of the stimulus-rate function: the occupancy-normalized tuning curve as calculated from all running epochs during the session. We note that this same approach was taken in other position decoding studies in which cells were found to have with multi-model place fields in the hippocampus (Fenton et al., 2008) and the entorhinal cortex (Fyhn et al., 2004).

Simultaneous decoding of both position and running direction.

It has long been known that some hippocampal place cells can exhibit strong tuning for direction in their preferred place fields, especially on linear tracks (McNaughton et al., 1983; Muller et al., 1994). Such 'unidirectional' single-peaked fields have recently been used to demonstrate that sequential reactivation in the hippocampus can proceed in the reverse order to that observed during experience (Diba and Buzsáki, 2007; Foster and Wilson, 2006; Frank et al., 2002). We found that for cells recorded while a rat explored our 10 meter-long track, direction-specific tuning was the rule, not the exception. Many cells fired strongly and reliably in many different locations on the track, with each field having a distinct directional tuning. This firing could not be explained by a simple conjunction of unimodal place and head direction tuning, since different fields had different headings. An example of a well-isolated cell with a complex firing field is shown in Figure 2.1, and the place-by-direction tuning curves from 30 well-isolated units are shown in Figure 2.2.

This finding encouraged us to attempt to simultaneously decode both the animal's position on the track, and its running direction (i.e. a binary decision about whether it was running towards one end of the linear track or the other, not the animal's head direction), as described below. By comparing the decoded running direction to the evolving direction of a replay event, it should be possible to determine whether the even represents a forward- or reverse-ordered sequential reactivation.

Outline of the algorithm

For the reasons discussed above, we chose to implement a simple Bayesian decoder that does not incorporate any previous estimates or firing history, and that uses an uninformative, uniform prior. We present the approach briefly here; it is presented in more detail as the '1-step Bayesian' algorithm in Zhang et al. (1998). Our modifications include the simultaneous decoding of position and direction.

We are interested in calculating the probability over the animal's linearized position along the track (*pos*) and running direction (*dir*) given a short time window of neural spiking (*spikes*). According to Bayes' rule:

$$Pr(pos, dir | spikes) = \frac{Pr(spikes | pos, dir) \cdot Pr(pos, dir)}{Pr(spikes)} \qquad (1)$$

We assume a uniform prior probability over position and direction, *Pr(pos,dir)*. Since the total probability must sum to 1, we therefore need only compute the likelihood, *Pr(spikes|pos,dir)*, in order to determine *Pr(pos,dir|spikes)*. Assuming the firing rates of all N units are independent, and assuming Poisson firing statistics we have:

$$Pr(spikes | pos, dir) = \prod_{i=1}^{N} Pr(spikes_i | pos, dir) \qquad (2)$$

$$= \prod_{i=1}^{N} \frac{(\tau f_i(pos, dir))^{n_i}}{n_i!} e^{-\tau f_i(pos, dir)} \qquad (3)$$

where τ is the duration of the time window of observation, *f_i(pos,dir)* is the expected firing rate of the i-th unit as a function of position and direction (i.e. the tuning curve of the i-th unit), and n_i is the number of spikes emitted by the i-th cell in the time window. Combining equations (1) and (3) leads to the following result:

$$Pr(pos, dir | spikes) = C \left(\prod_{i=1}^{N} f_i(pos, dir)^{n_i} \right) e^{-\tau \sum_{i=1}^{N} f_i(pos, dir)} \qquad (4)$$

where C is a normalization constant that depends on τ, and the number of spikes emitted by each cell. The output of this estimator is a 2-dimensional probability density function of location and running direction for the given time window. When only an estimate over position is required, we will take the marginal of this distribution. When a point estimate is required, the location of the peak of the 2-dimensional distribution (i.e. the maximum likelihood estimate) is given.

The free parameter in this algorithm are the length of the time window of observation τ. We will consider appropriate values for this parameter as we develop our approach below.

Validating the approach: decoding position during run

Previous work has found that for this class of algorithm, with tens of simultaneously recorded neurons, an observation window in the range of 250-1000 ms gives optimal decoding performance (Brown et al., 1998; Zhang et al., 1998). This agrees with our intuition that a longer integration time will compensate for variability of firing (such as occurs during the theta oscillation), but that we must also have frequent sampling in order to capture the location of a rat that may be running ~50 cm/s.[7]

We chose to use non-overlapping 500 ms observation windows. For the purposes of comparing our results to those in the literature, we begin by reporting the results of decoding position alone (Fig. 2.3). We decoded the activity of 46 simultaneously recorded CA1 principal cells with firing fields on the track (see Chapter 3 for details of the recording and cell selection). A single round of cross-validation was performed: tuning curves were generated from alternating 1 s bins (the training set), and decoding was attempted on the other half of 1 s bins (the test set). Decoding was very successful while the rat was running (Fig. 2.3A-C). Median error along the 10 m track was ~7 cm (i.e. < 1%) for all periods when the rat was running. No previous work has reported error measures for position on a linear track, but this figure is in line with the 5-6 cm error observed when decoding the activity of 32 place cells in a small open field by Barbieri et al. (2004), and appears to be consistent with the error reported by Zhang et al. (1998) for behavior on a figure-8 track.

We next performed simultaneous decoding of both position and running direction. We are not aware of any previous work that has attempted to decode running direction. The complex tuning of place cells to location and direction resulted in distinct ensemble codes for runs in the two directions (Fig. 2.4A-B). A single round of cross-validation showed that we could correctly decode running direction 89% of the time (vs a chance level of 50%), with good performance at most locations on the track (Fig. 2.4C).

Decoding during off-line states

Having developed and validated our decoding approach, we now consider the detection of

---

7   We have also used very short time windows during decoding of behavior to examine the short-timescale spatial structure of theta sequences (see chapter 4).

reactivation during off-line states. The choice of observation time window in this case is poorly constrained, since we do not have access to a ground truth with which to compare our results. Previous work using spike template methods (reviewed above) suggests that there is not a single correct answer: sequential reactivation during REM sleep was found to proceed at roughly the same speed as behavior (Louie and Wilson, 2001), while SWR-associated replay was found to be compressed in time up to 20-fold (Lee and Wilson, 2002). In a sense this parameter of the decoding algorithm is deeply intertwined with the phenomenon under study.

In the exploratory data analysis phase of our experiments, we explored a range of values for this parameter. Reactivation during SWR events took place primarily at a highly compressed timescale. Following the same intuition as we developed in choosing a timescale for run decoding, we chose a time window that was as long as possible (to adequately sample firing rates across the ensemble) while being short enough to capture temporal structure in the decoded reactivation. We found that for SWR-associated population bursts, a time window of 20 ms was optimal. We note that other researchers have also arrived at a similar range of values (15-20 ms) (Johnson and Redish, 2007; Karlsson and Frank, 2009).

We now consider neural reactivation in stimulus space. Figure 2.5 shows several examples of decoded ensemble activity during population bursts while the rat paused on the track. During single-SWR events, we frequently observed coherent, possibly sequentially structured reactivation (Figure 2.5A) similar to what had been reported during behavior and slow-wave sleep. We also observed some events that appeared to be long trains of SWRs lasting several hundreds of milliseconds, with extended sequential structure across time (Figure 2.5B-D).

Exploratory analysis of reactivation in stimulus space

We first asked whether prolonged reactivation events could be due to chance alignment of several single SWR-associated replays. We also wondered whether the apparently constant temporal evolution seen in the examples was typical. To address these questions, we calculated a spatiotemporal autocorrelation function for estimated position during all population bursts (Figure 2.6). The distance to the peak in the cross-correlation was predicted by the lag between the pairs, suggesting a replay of behavioral sequences at a constant 'virtual velocity' of approximately 10 m/s (~20 times faster than typical rat running speed, in agreement with the

results reported for SWS). This analysis suggested a constant-speed model of replay, rather than a random walk model, since a random walk would have result in a diffusion-like spread of activity centered on a distance lag of 0. We therefore fit a constant-speed model to the individual candidate events to characterize their replayed content.. This analysis of single candidate events is important because it allows us to study the possible cognitive roles of reactivation in relation to ongoing behavior.

Since the output of the decoder at each time bin is a probability density function (PDF) over location, we can estimate the likelihood of the animal being in any region along the track by integrating the corresponding region of the PDF. Analogously, we can evaluate the likelihood of any particular constant-speed trajectory by shifting the region of interest by a fixed distance between neighboring PDFs. In order to find the most likely such trajectory (the best fit), we densely sample all the possible trajectories, and select the one with the highest average likelihood. As is apparent from Figure 2.7, this is equivalent to the task of finding lines in images, a computational problem with mature and efficient solutions (Toft, 1996). The detected trajectory is defined by its starting position, and its virtual velocity (see Chapter 3).

This procedure will find a 'most likely' constant-speed trajectory for any population burst, and will assign it a likelihood. But this does not tell us whether the discovered trajectory is statistically significant—that is whether the structure observed is unusual, if we assume a null hypothesis of temporally unstructured neural activity. Methods to establish this will be developed and applied in the next chapter.

We also note that the decoded running direction appears to be consistent across extended replay events. This allows us to characterize individual events as proceeding in either the forward or reverse order, depending on whether the direction of the trajectory agrees with (forward order, see Figure 2C), or disagrees with (reverse order; see Figure 2D) the estimated running direction. In the Chapter 3, we will develop methods of quantifying and analyzing the statistical significance of this phenomenon.

Conclusions and future directions

We have demonstrated that a neural decoding approach can be applied to the detection of neuronal reactivation. The use of a probabilistic framework allows us to examine the content of

individual events in a statistically rigorous manner.

We have also demonstrated for the first time the simultaneous decoding of an animal's position and its running direction. We have also successfully decoded animal's head direction from ensembles of thalamic neurons (data not shown). In principle, we can decode any behavioral parameter that is encoded by the hippocampal ensemble. Hippocampal neurons are known to be tuned to previous and upcoming behavior, and to various task contingencies (Eichenbaum et al., 1999; Eichenbaum, 2008; Frank et al., 2000; Wood et al., 1999; Wood et al., 2000) that are perhaps more properly thought of as cognitive state variables, rather than external stimuli. If such states are also replayed during off-line states, the approach outlined here may provide a powerful tool for analyzing the structure and dynamics of thought processes.

References:

Barbieri, R., Frank, L.M., Nguyen, D.P., Quirk, M.C., Solo, V., Wilson, M.A., and Brown, E.N. (2004). Dynamic analyses of information encoding in neural ensembles. Neural Comput. *16*, 277-307.

Boice, R. (1977). Burrows of wild and albino rats: effects of domestication, outdoor raising, age, experience, and maternal state. J. Comp. Physiol. Psychol. *91*, 649-661.

Brown, E.N., Frank, L.M., Tang, D., Quirk, M.C., and Wilson, M.A. (1998). A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. J. Neurosci. *18*, 7411-7425.

Brown, E.N., Kass, R.E., and Mitra, P.P. (2004). Multiple neural spike train data analysis: state-of-the-art and future challenges. Nat. Neurosci. *7*, 456-461.

Calhoun, J.B. (1963). The ecology and sociology of the Norway rat (Bethesda, Md.: U.S. Dept. of Health, Education, and Welfare, Public Health Service; for sale by the Superintendent of Documents, U.S. Govt. Print. Off.).

Csicsvari, J., O'Neill, J., Allen, K., and Senior, T. (2007). Place-selective firing contributes to the reverse-order reactivation of CA1 pyramidal cells during sharp waves in open-field exploration. Eur. J. Neurosci. *26*, 704-716.

Davidson, T.J., Kloosterman, F., and Wilson, M.A. (2009). Hippocampal replay of extended experience. Neuron *63*, 497-507.

Davidson, T.J., Kloosterman, F., and Wilson, M.A. (February 2008) CoSyNe: Computational and Systems Neuroscience meeting. Salt Lake City, Utah.

Davidson, T.J., Kloosterman, F., and Wilson, M.A. (2007). Stimulus reconstruction reveals extended 'replay' in the rat hippocampus during exploration. CNS: Computation and Neural systems. BMC Neurosci. *8(Suppl. 2)*, P198.

Diba, K., and Buzsáki, G. (2007). Forward and reverse hippocampal place-cell sequences during ripples. Nat. Neurosci. *10*, 1241-1242.

Eichenbaum, H. (2008). Hippocampal neuronal activity and memory: should we still be talking about place cells? In Hippocampal place fields: relevance to learning and memory, Mizumori, S. J. ed., (Oxford; New York: Oxford University Press) pp. 161-174.

Eichenbaum, H., Dudchenko, P., Wood, E., Shapiro, M., and Tanila, H. (1999). The hippocampus, memory, and place cells: is it spatial memory or a memory space? Neuron *23*, 209-226.

Fenton, A.A., Kao, H.Y., Neymotin, S.A., Olypher, A., Vayntrub, Y., Lytton, W.W., and Ludvig, N. (2008). Unmasking the CA1 ensemble place code by exposures to small and large environments: more place cells and multiple, irregularly arranged, and expanded place fields in the larger space. J. Neurosci. *28,* 11250-11262.

Fenton, A.A., and Muller, R.U. (1998). Place cell discharge is extremely variable during individual passes of the rat through the firing field. Proc. Natl. Acad. Sci. U. S. A. *95,* 3182-3187.

Foldiak, P. (1993). The 'ideal homunculus': statistical inference from neural population responses. In Computation and Neural Systems, Eeckman, F. H. and Bower, J. M. eds., Kluwer Academic) pp. 55-60.

Foster, D.J., and Wilson, M.A. (2007). Hippocampal theta sequences. Hippocampus *17,* 1093-1099.

Foster, D.J., and Wilson, M.A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. Nature *440,* 680-683.

Frank, L.M., Brown, E.N., and Wilson, M. (2000). Trajectory encoding in the hippocampus and entorhinal cortex. Neuron *27,* 169-178.

Frank, L.M., Eden, U.T., Solo, V., Wilson, M.A., and Brown, E.N. (2002). Contrasting patterns of receptive field plasticity in the hippocampus and the entorhinal cortex: an adaptive filtering approach. J. Neurosci. *22,* 3817-3830.

Fyhn, M., Molden, S., Witter, M.P., Moser, E.I., and Moser, M.B. (2004). Spatial representation in the entorhinal cortex. Science *305,* 1258-1264.

Georgopoulos, A.P., Lurito, J.T., Petrides, M., Schwartz, A.B., and Massey, J.T. (1989). Mental rotation of the neuronal population vector. Science *243,* 234-236.

Jackson, W.B. (1982). Norway Rat and Allies. In Wild Mammals of North America, Chapman, J. A. and Feldhamer, G. A. eds., (Baltimore, Maryland: Johns Hopkins University Press) pp. 1077-1088.

Jackson, J.C., and Redish, A.D. (2003). Detecting dynamical changes within a simulated neural ensemble using a measure of representational quality. Network *14,* 629-645.

Jensen, O., and Lisman, J.E. (2000). Position reconstruction from an ensemble of hippocampal place cells: contribution of theta phase coding. J. Neurophysiol. *83,* 2602-2609.

Ji, D., and Wilson, M.A. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. Nat. Neurosci. *10,* 100-107.

Johnson, A., and Redish, A.D. (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. J. Neurosci. *27*, 12176-12189.

Johnson, A., Seeland, K., and Redish, A.D. (2005). Reconstruction of the postsubiculum head direction signal from neural ensembles. Hippocampus *15*, 86-96.

Karlsson, M.P., and Frank, L.M. (2009). Awake replay of remote experiences in the hippocampus. Nat. Neurosci. *12*, 913-918.

Kudrimoti, H.S., Barnes, C.A., and McNaughton, B.L. (1999). Reactivation of hippocampal cell assemblies: effects of behavioral state, experience, and EEG dynamics. J. Neurosci. *19*, 4090-4101.

Lee, A.K., and Wilson, M.A. (2004). A combinatorial method for analyzing sequential firing patterns involving an arbitrary number of neurons based on relative time order. J. Neurophysiol. *92*, 2555-2573.

Lee, A.K., and Wilson, M.A. (2002). Memory of sequential experience in the hippocampus during slow wave sleep. Neuron *36*, 1183-1194.

Louie, K., and Wilson, M.A. (2001). Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep. Neuron *29*, 145-156.

McNaughton, B.L., Barnes, C.A., and O'Keefe, J. (1983). The contributions of position, direction, and velocity to single unit activity in the hippocampus of freely-moving rats. Exp. Brain Res. *52*, 41-49.

Muller, R.U., Bostock, E., Taube, J.S., and Kubie, J.L. (1994). On the directional firing properties of hippocampal place cells. J. Neurosci. *14*, 7235-7251.

Nadasdy, Z., Hirase, H., Czurko, A., Csicsvari, J., and Buzsáki, G. (1999). Replay and time compression of recurring spike sequences in the hippocampus. J. Neurosci. *19*, 9497-9507.

O'Keefe, J., and Nadel, L. (1978). The hippocampus as a cognitive map (Oxford; New York: Clarendon Press; Oxford University Press).

O'Keefe, J., and Recce, M.L. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. Hippocampus *3*, 317-330.

Oram, M.W., Foldiak, P., Perrett, D.I., and Sengpiel, F. (1998). The 'Ideal Homunculus': decoding neural population signals. Trends Neurosci. *21*, 259-265.

Pavlides, C., and Winson, J. (1989). Influences of hippocampal place cell firing in the awake state on the activity of these cells during subsequent sleep episodes. J. Neurosci. *9*, 2907-2918.

Sanger, T.D. (2003). Neural population codes. Curr. Opin. Neurobiol. *13*, 238-249.

Sanger, T.D. (1996). Probability density estimation for the interpretation of neural population codes. J. Neurophysiol. *76*, 2790-2793.

Skaggs, W.E., and McNaughton, B.L. (1998). Spatial firing properties of hippocampal CA1 populations in an environment containing two visually identical regions. J. Neurosci. *18*, 8455-8466.

Skaggs, W.E., McNaughton, B.L., Wilson, M.A., and Barnes, C.A. (1996). Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. Hippocampus *6*, 149-172.

Toft, P.A. (1996). The Radon Transform - Theory and Implementation. PhD thesis, Technical University of Denmark. URL: http://petertoft.dk/PhD/

Wilson, M.A., and McNaughton, B.L. (1994). Reactivation of hippocampal ensemble memories during sleep. Science *265*, 676-679.

Wilson, M.A., and McNaughton, B.L. (1993). Dynamics of the hippocampal ensemble code for space. Science *261*, 1055-1058.

Wood, E.R., Dudchenko, P.A., and Eichenbaum, H. (1999). The global record of memory in hippocampal neuronal activity. Nature *397*, 613-616.

Wood, E.R., Dudchenko, P.A., Robitsek, R.J., and Eichenbaum, H. (2000). Hippocampal neurons encode information about different types of memory episodes occurring in the same location. Neuron *27*, 623-633.

Zhang, K., Ginzburg, I., McNaughton, B.L., and Sejnowski, T.J. (1998). Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. J. Neurophysiol. *79*, 1017-1044.
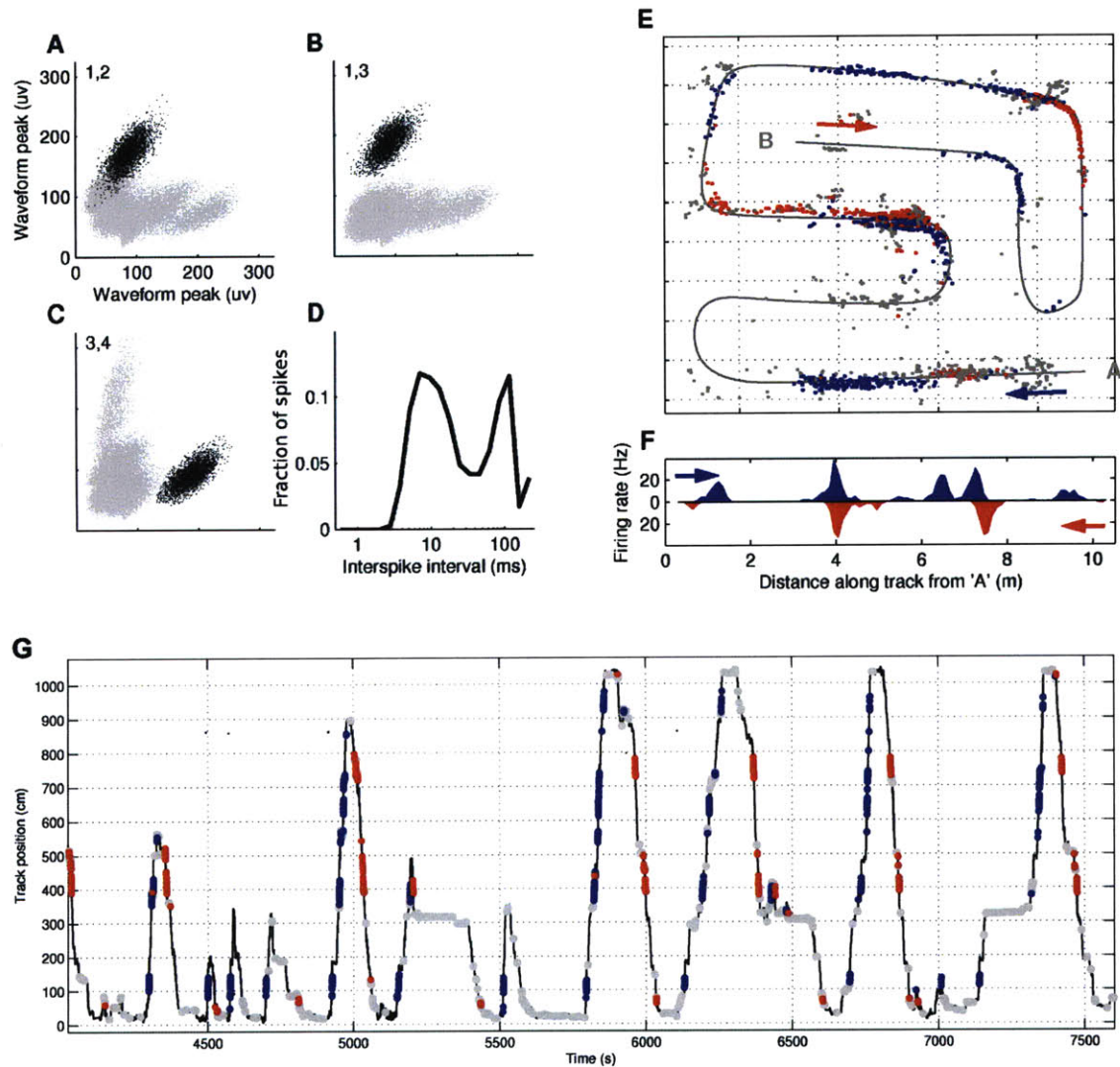
Figure 2.1: Complex tuning of a single cell to location and running direction

A-C, Scatter plots of peak waveform amplitude across pairs of channels for all spikes observed on a single tetrode. Black points represent spikes from a single well-isolated unit. Tetrode channel numbers appear in top left of each panel. **D,** Inter-spike interval histogram during running for the indicated unit, showing a clear refractory period >2 ms as well as modulation at theta frequencies (8-12 Hz; ~100 ms interval). Note log time scale. **E,** Location of rat at all spike times for this unit. Blue points: spikes emitted while rat was running A→B; red points: spikes emitted while rat was running B→A; gray points: spikes emitted when rat was not running. **F,** Joint tuning curve over linearized position and running direction. Blue: place-rate map during A→B runs; Red: place-rate map during B→A runs. G, The same spikes as in (F), but plotted as a function of time and linearized animal position for ~1h of behavior. Note that the complex spatial and directional tuning is stable across time. Black line: animal's position.
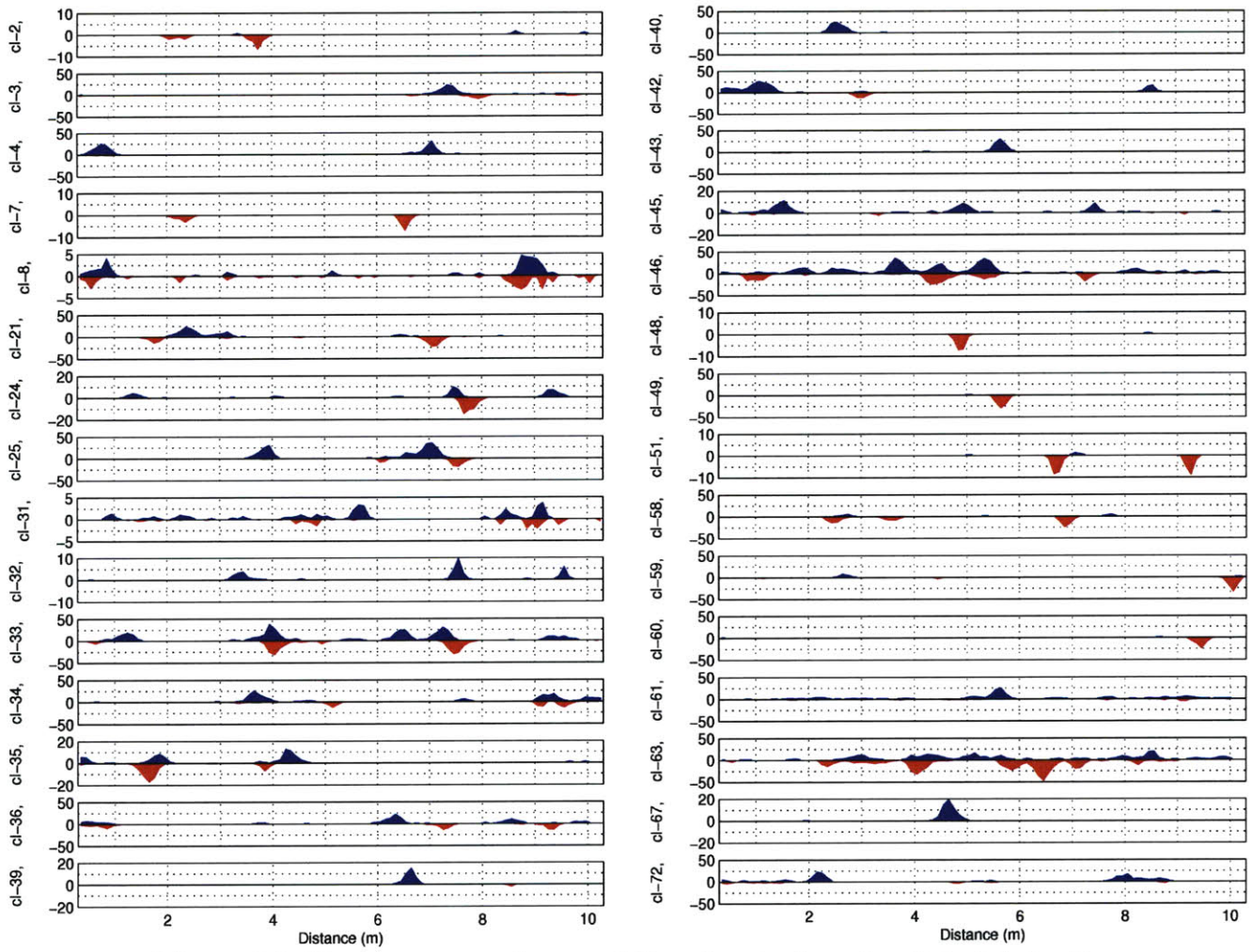
Figure 2.2: Heterogeneous place and direction tuning of well-isolated cells

Place and directional tuning curves as in Figure 2.1(F) for 30 well-isolated units. Note differing firing rate scales on y-axis (Hz). The cell in figure 2.1 is 'cl-33'.
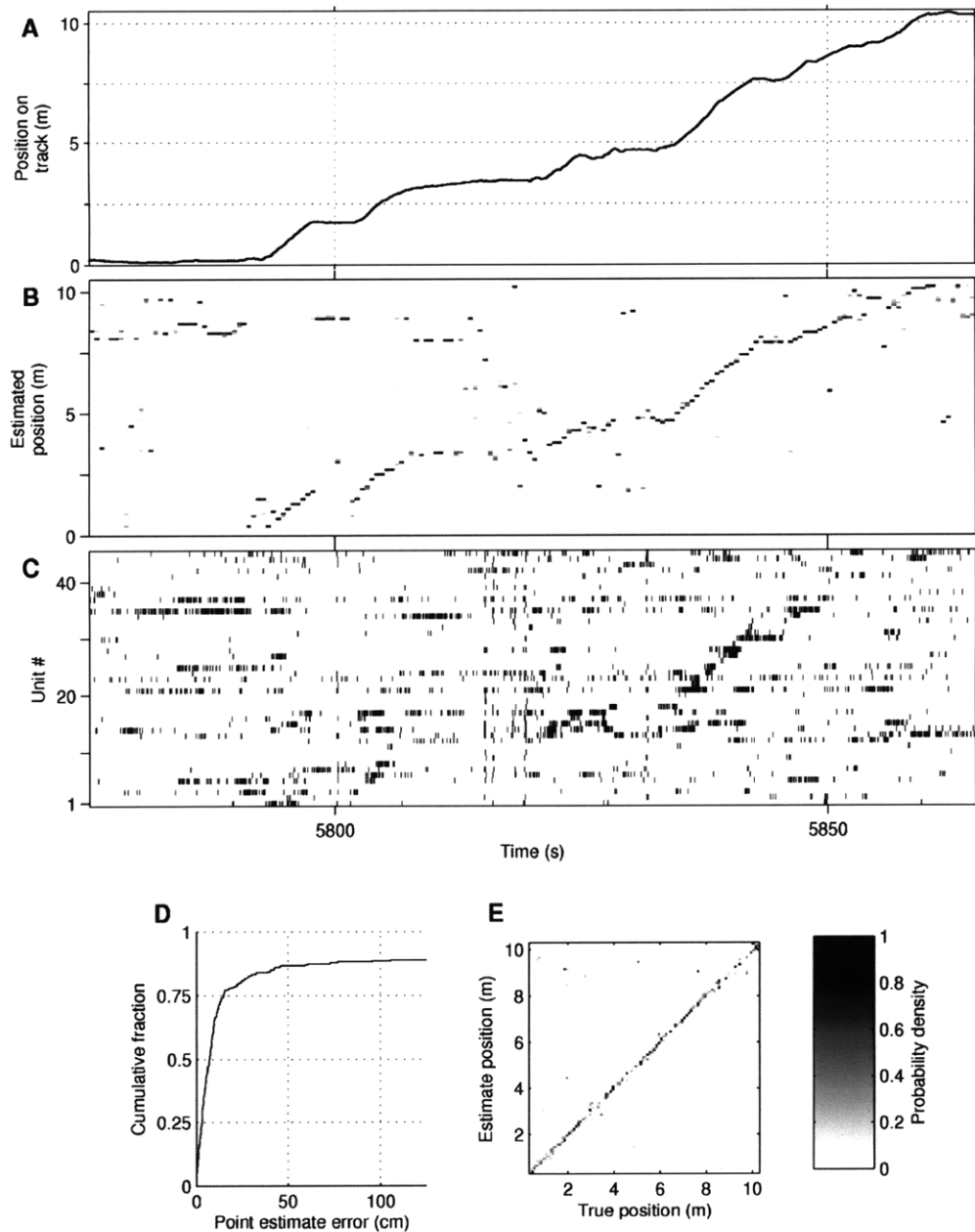
Figure 2.3: Estimation of animal position by ensemble decoding

A-C, Decoding of ensemble activity during 90 seconds of behavior in which the rat runs the full length of the 10 meter track. **A,** Actual position of rat. **B,** Estimated position (500 ms bins); each column is a probability density over position. Note the estimate is more accurate while animal is running. **C,** Raster plots of the 46 place cells used to generate the estimates in (B), sorted by the location of their peak firing. Note organized bursts of activity involving many of the cells near 5820 s. **D-E,** Accuracy of position estimation during running periods (speed > 15 cm/s) in a 1 h session of track exploration. **D,** Cumulative density plot of error between point estimate (i.e. maximum likelihood estimate) of location, and true location. Median error is ~7 cm. **E,** 'Confusion matrix' showing average position estimate (as a column) for each true animal location. Note even coverage across most locations on the track, and the lack of systematic errors (which would appear as power off the diagonal).
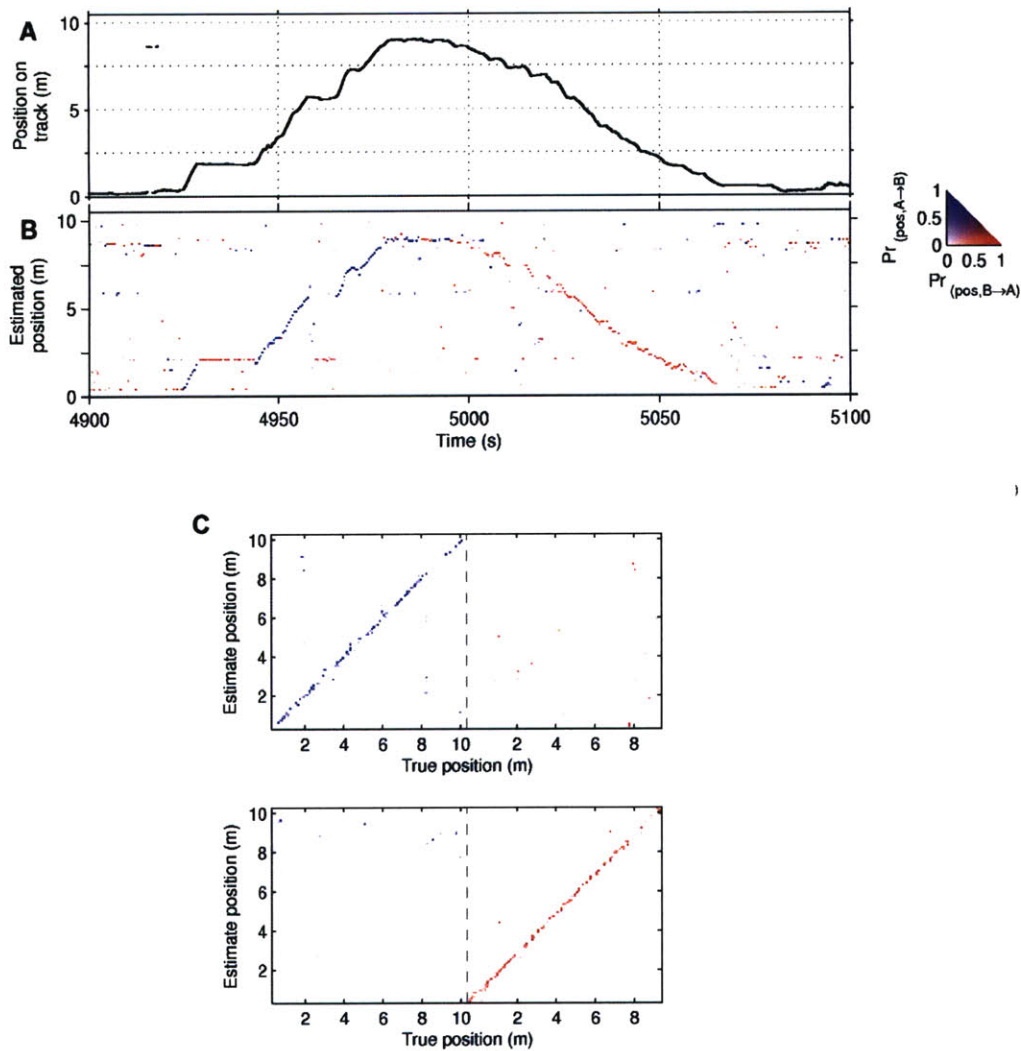
Figure 2.4: Simultaneous estimation of animal location and running direction

A, Animal position during during a run proceeding from 0→9 m, and then returning to 0. **B,** Joint estimation of position and running direction (500 ms bins). Color indicates estimated running direction (see legend to right); direction is usually estimated correctly for both directions. **C,** Confusion matrix showing clear separation of the neural representations for the two running directions. Top: average estimates when rat was running A→B; bottom: when rat was running B→A. Unlike in B, where the estimates for the two directions are merged into the same image, here the two 'slices' of the 2-D place-by-direction estimate are plotted next to each other, separated by a dashed gray line. Confusion of one direction for another at a given location would result in power along the diagonal in the top-right, or bottom-left panels.
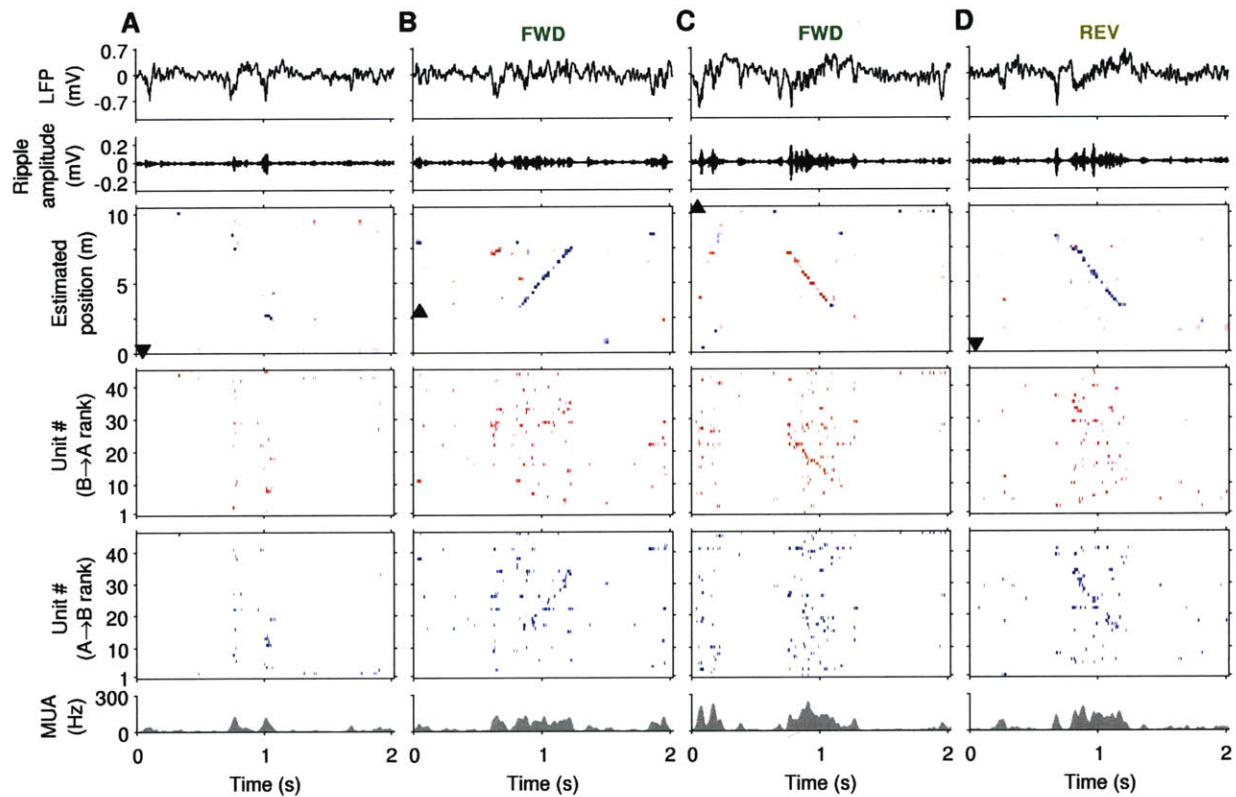
Figure 2.5: Examples of sharp wave-ripple-associated reactivation

A-D, Four examples of short timescale decoding (20 ms bins) of neural activity during population bursts while the rat is stopped. From top to bottom: Local field potential (LFP) trace showing sharp wave activity; filtered LFP (150-250 Hz bandpass) showing ripple activity; joint estimates of animal location and running direction, with black arrowhead indicating the animal's current position and heading direction; raster plots of individual hippocampal units sorted by location of peak in the B→A tuning curve (red) and the A→B tuning curve; multi-unit activity (MUA; spiking rate/tetrode) for all detected spikes, including those not used for decoding. FWD: Forward replay; REV: Reverse replay (see text).
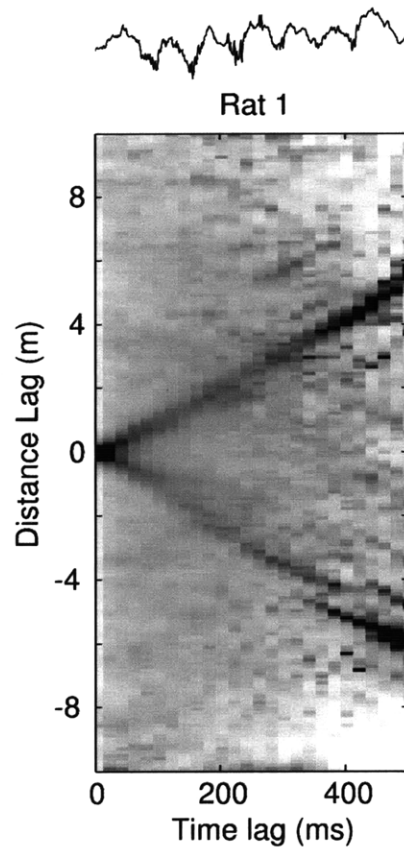
Figure 2.6: Spatiotemporal autocorrelation of estimated position reveals constant-speed replay.

Position estimates were calculated using 20 ms bins during all population bursts. Each column is the average unbiased spatial cross-correlation for pairs of position estimates separated by the indicated time lag. The location of peak correlation shifts to more distant locations with increased time lag, indicating extended, constant-speed replay. A train of sharp wave-ripples is plotted above for reference, using the same time scale.
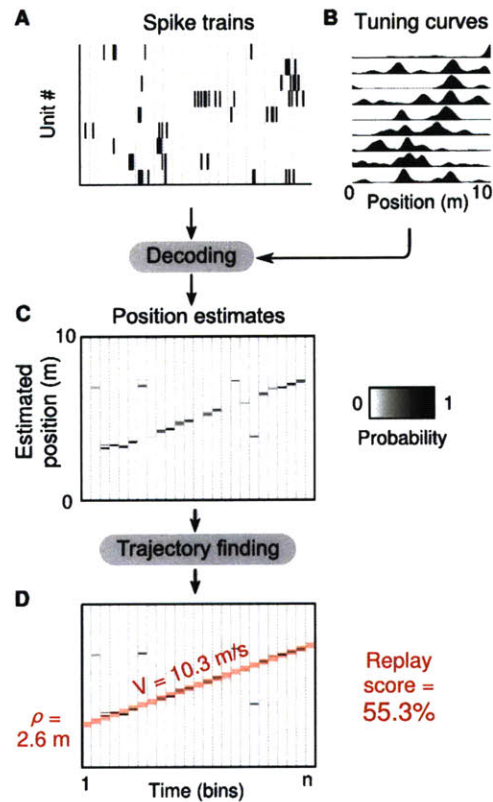
Figure 2.7: Probabilistic detection of linear replay trajectories

Schematic of replay detection procedure. Ensemble activity during a population burst (**A**) and corresponding pre-computed tuning curves for each cell (**B**) are used to generate a series of independent estimates of position (**C**). Since each column is a probability density function over location, we can evaluate the likelihood that the replay followed any given trajectory by integrating the corresponding regions of the PDFs. We hypothesize that replay proceeds at a constant speed, and so we search for the most likely one and report it as the candidate replay trajectory (30 cm-wide pink line in **D**). This trajectory is defined by its start location (ρ) and velocity (V). The mean likelihood that the replay was on this trajectory (i.e. the mean of the density within the pink regions in each column) is reported as the "Replay score".

Chapter 3:

Hippocampal replay of extended experience[1]

Summary

During pauses in exploration, ensembles of place cells in the rat hippocampus re-express firing sequences corresponding to recent spatial experience. Such 'replay' co-occurs with ripple events: short-lasting (~50-120 ms), high frequency (~200 Hz) oscillations that are associated with increased hippocampal-cortical communication. In previous studies, rats explored small environments, and replay was found to be anchored to the rat's current location, and compressed in time such that replay of the complete environment occurred during a single ripple event. It is not known whether or how longer behavioral sequences are replayed in the hippocampus. Here we show, using a neural decoding approach, that firing sequences corresponding to long runs through a large environment are replayed with high fidelity (in both forward and reverse order), and that such replay can begin at remote locations on the track. Extended replay proceeds at a characteristic virtual speed of ~8 m/s, and remains coherent across trains of ripple events. These results suggest that extended replay is composed of chains of shorter subsequences, which may reflect a strategy for the storage and flexible expression of memories of prolonged experience.

Introduction

Place cells in the hippocampal formation fire selectively when an animal moves through particular locations ("place fields") in the environment (O'Keefe and Dostrovsky, 1971; Wilson and McNaughton, 1993). As a consequence, when the animal travels along a given trajectory hippocampal cells with place fields on that trajectory are activated in sequence. During pauses in locomotion, and during slow-wave sleep, many place cells are recruited in intermittent population bursts, which are accompanied by ripples in the hippocampal local field potential (Buzsáki et al., 1992; Chrobak and Buzsáki, 1996). The firing order of place cells during those bursts reflects a memory for the order in which they were activated during previous exploration. Such 'replay' has been observed during slow-wave sleep (Ji and Wilson, 2007; Lee and Wilson, 2002; Nadasdy et al., 1999; Wilson and McNaughton, 1994) as well as during immobility on linear tracks (Diba and Buzsáki, 2007; Foster and Wilson, 2006) and in an open field (Csicsvari et al., 2007).

During replay events in rats, place cell firing sequences are re-expressed at a faster rate than during actual experience (Diba and Buzsáki, 2007; Foster and Wilson, 2006; Ji and Wilson, 2007; Lee and Wilson, 2002). For the small 1-2 meter long linear tracks used in previous studies, the firing sequence of a set of place cells that spans the complete environment can be re-expressed at the same time scale of a single ripple (50-120 milliseconds (Ylinen et al., 1995)). These results can be accounted for by a model in which sensory input drive to place cells is 'read out', possibly by a sweeping release of inhibition during a single sharp wave-ripple (Csicsvari et al., 2007; Diba and Buzsáki, 2007; Foster and Wilson, 2006).

The limited duration of single ripple events suggest that awake replay in a large environment should be limited to a small region of space. In the wild, however, rats typically navigate over tens, or even hundreds of meters (Jackson, 1982). Can the hippocampus support replay across larger spatial scales? If so, is such extended replay further compressed in time or is there a fixed rate at which replay progresses? If the latter, how are longer sequences mapped onto short-lasting ripple events?

Results

Extended replay detected by neural decoding

Simultaneous recordings of multiple single units in hippocampal area CA1 were made (n=47, 34, 23 and 32 units with consistent place-related firing in rats 1-4; see Experimental Procedures) while rats explored a 10.3-meter long track (Fig. 3.1A-B). Food reward was provided at both ends of the track, but since rats were not pre-trained, behavior was variable and the animals frequently paused at many locations on the track (Fig. 3.1C). We linearized the animal's position, such that it represented the distance from one end of the track (Fig. 3.1C), and the behavior of the rat was classified as either 'RUN' (linear speed > 15 cm/s) or 'STOP' (linear speed < 5 cm/s) (Fig. 3.1E). Candidate replay events ('CAND') were identified as periods during STOP with elevated multi-unit activity across all electrodes (Fig. 3.1D-E; see Experimental Procedures; mean rate during STOP of 0.36, 0.40, 0.32, 0.40 events/s in rats 1-4). Candidate events were characterized by sharp on- and offsets (Fig. 3.1G-H), and event durations ranged

from 40-1018 ms (Fig. 3.1F), with 19% of events (17%, 22%, 16%, 14% in rats 1-4) characterized as 'long' (>250 ms; chosen to be more than twice the typical duration of a single ripple (Ylinen et al., 1995)).

To evaluate whether candidate events contained replayed spatial memory sequences, we employed a probabilistic neural decoding strategy to estimate the animal's position on the track from the ensemble of spike trains (Brown et al., 1998; Wilson and McNaughton, 1993; Zhang et al., 1998). We reasoned that during reactivation of previous experience the position estimate would deviate systematically from the actual position (Johnson and Redish, 2007).

Our algorithm does not require that each cluster used for decoding contains only spikes emitted by a single neuron; successful estimation requires only that the spatial tuning of each unit is stable across the training and decoding epochs. This property of the decoder allows us to make optimal use of the spatial information present in the neural data by including units that are less well isolated, but which nevertheless have a stable spike amplitude signature and carry consistent spatial information. We therefore interpret our results in terms of the behavior of the hippocampal ensemble rather than that of individual place cells; we use the term 'unit' rather than 'cell' throughout the paper to emphasize this distinction. All reported results were qualitatively similar when calculated using only well-isolated units (see below).

We first confirmed that we could use our decoder to accurately estimate the animal's position during RUN using 500 ms time windows. Median error for rats 1-4 was 7, 9, 8, and 8 cm, with good performance across the entire environment (Fig. 3.2A-C; Fig. 3.6 and Movie 1).

We next applied the decoding algorithm to non-overlapping 20 ms time windows in all candidate events lasting at least 100 ms. During many candidate events the sequence of position estimates described a rapid traversal of a section of the track at a relatively constant speed, even though the animal was stationary (Figs. 2, 3A-C). The decoding algorithm we use is memoryless, and therefore the observed trajectories are not the result of temporal smoothing across neighboring estimates.

In order to characterize individual events, we determined the best linear fit to the observed pattern of position estimates for each candidate event by maximizing a 'replay score.' The resulting fit specifies the most likely constant-speed trajectory being replayed, and the replay score corresponds to the mean estimated likelihood that the rat was on the specified trajectory (see Experimental Procedures and Fig. 3.7. To test for statistical significance, we compared the observed replay score for each event to sample distributions of scores obtained after shuffling the original data. Three distinct shuffling regimes were employed to control for non-specific factors possibly contributing to the replay score (Fig. 3.7). First, to control for the chance linear alignment of position estimates, we circularly shifted the estimate at each time point by a random distance ("column-cycle shuffle"). Second, to control for the contribution to the replay score of firing characteristics of single units (e.g. bursting), we randomly permuted the mapping between spiking records and spatial tuning curves ("unit identity shuffle"). Third, to control for a bias of the decoding procedure towards particular locations, we constructed artificial candidate events by combining position estimates taken randomly from the complete set of candidate events in each session ("pseudo-event shuffle"). We performed each shuffle 1500 times and conservatively consider only events with a Monte Carlo p-value < 0.01 under all three shuffles to be significant.

Using these criteria, 16% of all analyzed candidate events contained significant replay trajectories (118/657, 109/699 , 12/137, 24/163 events significant for rats 1-4; $p < 10^{-7}$ for each rat under a binomial distribution assuming a false positive rate of 1%). Of long (>250 ms) candidate events, 33% were significant (59/145, 60/203, 9/36, 10/38 candidate events in rats 1-4; $p < 10^{-10}$ for each rat). (Using a significance threshold of $p < 0.05$, 29% of all events, and 50% of long events were found to contain replayed memory sequences, however we chose to use a threshold of $p < 0.01$ to reduce the likelihood of false positives.)

Speed and location of replayed trajectories

Individual replay events could cover long sections of the track (Fig. 3.3A-C), and the length of replayed trajectories was linearly correlated with the duration of the events (Fig. 3.3D), indicating a characteristic speed for replay (Fig 3E; median speed of 8.1 m/s for all significant events; median of 8.7/7.6/10.8/10.5 m/s in rats 1-4). These replay speeds are 15-20 times faster than a typical rat running speed (~0.5 m/s), consistent with previous reports of compression factors for shorter-duration replay events (Lee and Wilson, 2002; Nadasdy et al., 1999).

We next analyzed the relation between the actual position of the animal during replay and the location of the replayed trajectories on the track. Replay occurred while the animal stopped at the ends of the track to consume reward as well as at other locations (Fig. 3.8A-D). Replay in both the A→B and B→A directions was common, with no clear trend across rats (58%, 51%, 25%, 21% of replays from A→B in rats 1-4). Since rats spent a significant amount of time at the reward sites facing away from the track, a higher proportion of replayed trajectories occurred behind the animal (Fig. 3.8; 35% ahead, 65% behind). This bias was not significantly different from chance (33% ahead; $p = 0.58$, 2-sided Monte Carlo p-value), computed under the null hypothesis that there is no relation between the stopping location of the rat and the position of the replayed trajectory.

Locally and Remotely initiated replay

Previous reports suggested that replay might be influenced by strong local place-related inputs (Csicsvari et al., 2007; Diba and Buzsáki, 2007; Foster and Wilson, 2006). Consistent with this model for replay generation, we found that the start locations of the replay trajectories were strongly biased towards the rat's current location (Fig. 3.3F), with 40% of significant replay trajectories starting within 50 cm of the rat's current location, which we refer to as 'local replay' (chance = 17%, calculated by bootstrapping under the null hypothesis that replay trajectories and the rat's position are uncorrelated, $p < 0.0005$ pooled across rats). The location of the ends of replay trajectories were not similarly biased towards current location (Fig. 3.3G), with only 5% ending nearer than 50 cm (chance = 8%, $p = 0.99$ pooled across rats).

We also observed many significant replay trajectories that began at remote locations (Fig. 3.3F; see examples in Fig. 3.3B, Fig. 3.4C-D). Indeed, 51% of events started at least 1m away from the rat's current location. Such trajectories could be artifacts of our replay detection method, if we made an error in determining the start time of the candidate event. A remote event could be either a truncated fragment of a long trajectory that actually begins at the current location (i.e. event start time too late; see Fig. 3.12C for a possible example), or an incorrect extrapolation of a shorter trajectory that actually begins at the current location (i.e. start time too early). We conservatively exclude these two classes of possible errors by selecting only trajectories that proceed from a remote location towards the animal; and that never proceed past the current location. Twenty percent of significant replay events (52 / 263) meet these more stringent criteria and are termed 'remote replay' (dashed lines in Fig. 3.8). There are significantly more remote replay events than the number of false positives expected to be generated by our replay detection procedure (52 of 1656 events; $p < 10^{-11}$ under a binomial distribution, using false positive rate of 1%).

Forward- and reverse-ordered replay

Previous studies have taken advantage of the joint tuning of many CA1 cells to running direction and location on the linear track (McNaughton et al., 1983; Muller et al., 1994) to demonstrate that spiking sequences can be replayed in either the forward or reverse temporal order (Diba and Buzsáki, 2007; Foster and Wilson, 2006). In order to determine the ordering of the observed replay trajectories, we extended our decoding procedure to estimate both the rat's position and its *instantaneous* running direction (i.e. whether the rat is running from 'A→B' or 'B→A') from the entire ensemble (Fig. 3.4A; see Experimental Procedures). Running direction was estimated correctly during RUN 89%, 83%, 83%, and 89% of time in rats 1-4 (chance 50%; $p < 10^{-12}$ in each rat).

For each replay event, we computed a 'replay order score' which reflected the degree to which our estimate of instantaneous running direction agreed with (+1, forward replay) or disagreed with (-1, reverse replay) the overall direction of the trajectory being replayed (see Experimental Procedures). For example, the reverse replay event in Figure 3.4D proceeds in the

B→A direction (from 7.5 to 2.5 m on the track), but uses the hippocampal ensemble code associated with running in the opposite A→B direction, as indicated by the blue color. Overall, replay order scores were significantly biased away from 0 and towards -1 and 1 (Fig. 3.4G; $p <$ 0.02 for each rat, one-sided Kolmogorov-Smirnov two-sample test, compared to pseudo-event shuffle distribution), indicating that the hippocampal ensemble tends to represent one running direction consistently within a replay event. Statistical significance of the replay order score was next tested for each event by comparison to a distribution of order scores obtained from shuffled data. Significantly ($p < 0.05$) forward- and reverse-ordered replays were observed in all sessions (Fig. 3.4B-D, Figs. 10-12, and Movie 2) across the full range of event durations (Fig. 3.4G).

Forward replay is significantly more frequent than reverse replay ($p < 0.005$ by two-sided binomial test), consistent with a previous report (Diba and Buzsáki, 2007). Of all replay events, 40% (106/263) were significantly forward-ordered, 26% (68/263) were significantly reverse-ordered, with the remaining 33% of events (89/263) classified as 'mixed' replay. This difference was also significant ($p < 0.002$) among replay events longer than 250ms: 48% (66/138) were significantly forward-ordered, and 25% (34/138) were significantly reverse-ordered. Mixed replay events contain significant replay in decoded position, but do not exhibit a consistent directional estimate. Most mixed events exhibit weak or variable direction tuning (e.g. Fig. 3.4E), but we also occasionally observe events which apparently switch represented directions in mid-replay (e.g. Fig. 3.9F, Figs. 10D & 12C).

We observed no significant difference in the speeds of forward and reverse replay trajectories (Fig. 3.4H; median speed 8.6 vs. 7.9 m/s ; Kolmogorov-Smirnov two-sample test, $p =$ 0.24). Forward and reverse replay trajectories did not preferentially correspond to runs in the A→B or B→A direction (47% of forward, and 52% of reverse replay trajectories proceeded from A→B). Similar proportions of both forward and reverse replay trajectories were initiated locally or remotely ($p > 0.5$ by two-proportion z-test): 42% of forward and 38% of reverse replay events were local (greater than the respective chance levels of 17% and 19%, $p < 0.0005$); and 19% of forward and 20% of reverse replay events were remote.

Locally-initiated forward replay will reflect possible future paths, while locally-initiated reverse replay will reflect possible approaches to the animal's current location. Do such replayed trajectories preferentially express the animal's *actual* past and future paths, rather than the paths not taken? To address this question, we analyzed periods when the animal was stopped in the middle of the track (at least 2 m from either end), where there are two possible paths associated with the animal's current location. We did not find a strong bias for locally-initiated forward replay trajectories to represent the actual future path (15 actual future path vs. 12 opposite direction). Similarly, there was no strong bias for locally-initiated reverse replay to represent the actual path taken by the animal to reach the current location (9 actual past path vs. 7 opposite direction).

Relationship between extended replay and ripples

Replay events have consistently been found to co-occur with ripple oscillations in the hippocampal local field potential (Diba and Buzsáki, 2007; Foster and Wilson, 2006; Ji and Wilson, 2007; Lee and Wilson, 2002; Nadasdy et al., 1999). Consistent with these reports, we found that ripple emission rate was much higher during replay events than during non-candidate event STOP periods (8.8-11.8 s^{-1} vs. 0.17-0.27 s^{-1}; $p < 10^{-4}$ in each rat). Detected ripples were associated with transient deflections in the LFP ('sharp waves'; Fig. 3.5B-D) and with transient increases in multi-unit activity (78-88% increase; $p < 10^{-7}$ for each rat; Fig. 3.5C,E), and single unit firing rate (81-94% increase; $p < 0.0002$ for each rat). These effects each lasted approximately 50 ms, which is comparable to the duration of single sharp wave-ripple complexes as described previously (Ylinen et al., 1995). In order to characterize the relationship between ripple events and extended replay, we performed a linear regression and found a strong positive correlation between the number of emitted ripples and the duration of the replay event (Fig. 3.5A; $R^2 = 0.65, 0.45, 0.70, 0.67$ for rats 1-4, $p < 0.001$ for each rat). These results demonstrate that extended replay spans trains of discrete sharp wave-ripple events.

Next we explored whether the confidence of the position reconstruction during replay events was uniform across the ripple trains. We find that reconstruction quality during replay, as measured by the mode of the position estimate, is significantly elevated at ripple peak times (Fig.

3.5F; 0.20-0.26 vs. 0.32-0.41; $p < 10^{-4}$ for each rat). We also find that the error between the replayed trajectory and the estimated position is significantly lower at ripple peak times (Fig. 3.5G.; 87-131 cm vs. 182-227 cm, $p < 0.003$ for each rat). These data show that replay integrity is not uniform across the duration of an event, but that it is modulated in association with ripple trains, suggesting that extended replay consists of chains of shorter ripple-associated subsequences.

Effect of unit isolation criteria

We repeated all analyses using conservative unit isolation criteria (clear cluster boundary separation, clear refractory period in ISI, high complex-spike index [CSI], and low cross-CSI between clusters), obtaining between 17-32 simultaneously-recorded units with clear spatial modulation during behavior per session, rather than the 23-47 units used in the main paper.

The exclusion of marginally-isolated units resulted in a marked decrease in position decoding accuracy during RUN: median position error increased from 8.2 to 14.1 cm; the fraction of correct estimates of running direction decreased from 87% to 77%; and evenness of coverage over the long track was also poorer.

The same candidate event times were used since these depend only on unclustered multi-unit activity. Significant replay was detected in 8.6% of events, rather than the 16% we report in the main results (for events lasting at least 250 ms, the detection frequency was 20% rather than 33%). All results are significantly higher than the 1% expected by chance ($p < 0.01$ per session, $p < 10^{-13}$ pooled across sessions for both short and long events).

The characteristics of the detected replay were little changed. Median replay speed was 7.8 m/s rather than 8.1 m/s. Replay start locations remained significantly biased towards current location (39% local replay, $p < 0.0005$, rather than 40%). Remote replay accounted for 25% (rather than 20%) of detected events, and is observed significantly more frequently than expected by chance ($p < 0.0001$ pooled across sessions). Replay order scores remained significantly biased towards forward and reverse ($p < 10^{-47}$ pooled), and forward-ordered was more common than reverse-ordered replay (41% and 20% of all significant replay events, respectively, rather than 40% and 26%).

Discussion

We have shown that time-compressed forward and reverse hippocampal replay of long behavioral sequences is common during pauses in exploration of a large environment, and is associated with trains of ripple events. In contrast to studies conducted in smaller environments, we find that replay is neither limited to locations where reward is consumed, nor exclusively tied to the animal's current location.

We developed and used a neural decoding approach for replay detection. Performing replay detection in the decoded spatial domain affords advantages over methods that examine firing order across individual units, such as pair-wise correlation (Wilson and McNaughton, 1994) and spike sequence detection (Diba and Buzsáki, 2007; Foster and Wilson, 2006; Lee and Wilson, 2004). Our method allowed us to examine the fine spatial structure of replayed trajectories in a statistically rigorous manner, and makes optimal use of the spatial information contained in hippocampal spikes, including the activity of cells with irregularly-shaped place fields (such as those with multiple firing fields (Fenton et al., 2008); see chapter 2).

Using this decoding approach, we demonstrated that behavioral sequences spanning long sections of a 10 m track are re-expressed during population bursts lasting up to 700 ms. Replay trajectories proceed at a constant speed of ~8 m/s, approximately 15-20 times faster than typical rat running speeds. Such values are consistent with the compression factors determined previously by analysis of spike time lags in smaller environments (Diba and Buzsáki, 2007; Lee and Wilson, 2002) . The constant speed of replay contrasts strongly with the rats' highly irregular behavior on the track, suggesting that the sequential structure of the behavioral experience, rather than the detailed time course of particular episodes, is re-expressed during replay. Constant-speed replay is also reminiscent of studies in humans showing that response times are linearly dependent on distance traveled across an imagined map (Kosslyn et al., 1978), or on the magnitude of mental rotation of 3-dimensional objects (Shepard and Metzler, 1971).

We confirmed previous reports that awake replay events are associated with sharp wave-ripples in the local field potential (Csicsvari et al., 2007; Diba and Buzsáki, 2007; Foster and Wilson, 2006). However, the extended replay sequences we report last much longer than the

duration of a single sharp wave-ripple event and we demonstrate that they span trains of sharp wave-ripples. Such trains have been noted since the first reports of ripples (Buzsáki et al., 1983; O'Keefe and Nadel, 1978; Suzuki and Smith, 1987), but no function has previously been ascribed to this phenomenon. CA1 unit activity is highest at the peak of individual ripples, corresponding to an increased confidence of the position estimate, which suggests that extended hippocampal replay may consist of chains of subsequences, each with a spatial extent of ~50 cm (based on a ripple duration of 60 ms, and a replay speed of 8 m/s)..

Both theta sequences (see chapter 4 and (Mehta et al., 2002)) and ripple-associated replay (Foster and Wilson, 2006) have been proposed to arise from a translation of place cell excitability into a phase or latency offset by a sweeping decrease in inhibition. This model predicts that ripple-associated sequences in the hippocampus should be limited to roughly the spatial scale of a single place field, and would therefore require that longer sequences consist of several sharp wave-ripple associated subsequences. One possible mechanism for the generation of trains of subsequences is suggested by the re-entrant loops in the hippocampal-entorhinal circuitry (Canto et al., 2008; Kloosterman et al., 2004). Following each ripple, current hippocampal output to the entorhinal cortex could be fed back to the hippocampus, providing the inputs required for expression of the next subsequence. Alternatively, extended replay may reflect the continuous operation of an auto-associative network, possibly in area CA3 (August and Levy, 1999). Recordings across multiple brain regions will be necessary to test these hypotheses.

We found a bias for both forward- and reverse-ordered replay trajectories to begin near the animal, which suggests that such events could be used for evaluation of immediate future and past paths. We also found that when the rat was stopped in the middle of the track, where there are multiple possible paths away from (and possible approaches to) the current location, replayed trajectories were not strongly correlated with the animal's actual behavior. In particular, forward replay trajectories were not predictive of the upcoming path, and reverse replay did not preferentially reflect the path just taken by the animal. These results suggest that replayed trajectories represent the set of possible future or past paths linked to the animal's current position, rather than the actual paths. Further study of the correspondence between replay order

and behavior may benefit from the use of tasks that place specific demands on the animal's evaluation of past and future experience.

Diba and Buzsáki (2007) found that forward replay beginning at the present location, moving along the upcoming path, was more common than forward replay beginning at a remote location and proceeding towards the animal's current location along the preceding path. Similarly, they find a preference for reverse replay events to represent the previous path (which, since it is replayed in reverse, is also initiated locally). These results are consistent with our observation of a bias towards local initiation for both forward and reverse replay.

We also report that a significant number of replay events express trajectories beginning at locations remote from the physical location of the rat. This indicates that during awake replay the hippocampus has access to a broad range of stored memory sequences that are not solely dependent on the current location or the behavior just prior to the replay event. In this respect, awake replay is similar to sequence reactivation during slow-wave sleep (Ji and Wilson, 2007; Lee and Wilson, 2002). In our experiments the rat has visual access to the complete track and it is possible that remotely initiated replay is cued by sensory inputs reaching the hippocampus through cortical pathways. Similarly, during slow-wave sleep cortical inputs may bias or otherwise influence memory reactivation (Ji and Wilson, 2007), given the complex bi-directional interactions between the hippocampus and neocortex (Isomura et al., 2006; Mölle et al., 2006; Siapas and Wilson, 1998; Sirota et al., 2003; Wolansky et al., 2006).

Interestingly, groups of ripples are also present during slow-wave sleep, where they predominantly occur during periodic increases in neocortical population activity ("up-states") (Battaglia et al., 2004; Clemens et al., 2007; Mölle et al., 2006; Sirota et al., 2003) associated with slow oscillations in the cortical EEG (Isomura et al., 2006; Steriade, 2006; Wolansky et al., 2006). During these up-states, coordinated memory reactivation has been observed in the hippocampus and visual cortex (Ji and Wilson, 2007). These data suggest that individual trains of ripples during both slow-wave sleep and in the awake state may constitute a higher level organization, possibly sharing a common mechanism for their generation.

Replay associated with single ripples may represent a building block for the expression of longer, more complex memories. Hippocampal replay has been proposed to contribute to memory consolidation during sleep (Buzsáki, 1989; Marshall and Born, 2007; Stickgold, 2005). During wakefulness, high-speed replay of long sequences of behavior could also support learning processes that would benefit from prospective or retrospective evaluation, such as reinforcement learning (Foster and Wilson, 2006). Extended replay may also support tasks involving memory recall. This last possibility, while speculative, is lent some support by the recent finding of specific reactivation of hippocampal neurons during free recall in humans (Gelbard-Sagiv et al., 2008), and by specific activation of the hippocampus during sequence recall tasks in humans (Lehn et al., 2009). This interpretation of the functional role of awake replay is also consistent with work suggesting a high degree of overlap in the cognitive processes supporting both episodic recall and the evaluation of future events in humans (Buckner and Carroll, 2007; Schacter et al., 2007). The link between awake replay and cognition can be further explored by studying replay in animals engaged in more cognitively demanding tasks, and by the experimental disruption or bias of replay.

Methods

Electrophysiology and behavior

All procedures were approved by the Committee on Animal Care at Massachusetts Institute of Technology and followed US National Institutes of Health guidelines. Microdrive arrays carrying 9-18 independently adjustable gold-plated tetrodes or octrodes (2 octrodes in one animal) aimed at area CA1 of the right dorsal hippocampus (2.4 mm lateral and 4.0 mm posterior, relative to bregma) were implanted under isoflurane anesthesia in 5 male Long-Evans rats (400-500 g). Tetrode and octrode construction is as previously described for tetrodes (Wilson and McNaughton, 1993): each electrode consists of a twisted bundle of 4 or 8 polyimide-insulated microwires, fused and cut to create a blunt tip. Wire used for tetrodes was either 13 µm-diameter nichrome resistance wire (RediOhm-800, Kanthal, Palm Coast, FL) electroplated with gold; or 15 µm-diameter nickel-iron wire (Nickel Alloy-120; California Fine Wire, Grover Beach, CA) with all recording sites plated with gold simultaneously using an electroless dip-plating process (Immersion Gold CF, Transene, Danvers, MA). Octrode wires were polyimide-coated tungsten (8 µm diameter, California Fine Wire, Grover Beach, CA). Electrodes were slowly lowered into the CA1 pyramidal cell layer over the course of 1-2 weeks. Individual units were isolated by manual clustering on peak spike waveform amplitudes across all channels using custom software (xclust; M.A.W.). For each electrode the local field potential (LFP) was recorded from a single channel, filtered from 1-475 Hz and sampled at 2 kHz. All recordings were differential against a reference electrode placed in white matter overlying CA1. A screw in the skull overlying cerebellum served as ground. For each rat, a single electrode showing clear sharp waves was selected for plotting of LFP.

Animals were not introduced to the ~10.3 m long linear track (Fig. 3.1A-B) until stable unit recordings were obtained, and only had one track session per day. Animals received food reward at the track ends ('A' and 'B'; see Fig. 3.1A-B) only after the rat completed a full length of the track. For each animal we selected a session in which the animal ran several complete laps, but behavior was still variable and the number of recorded units was large. These criteria limited the number of sessions we could use for each animal and we chose one session per animal in

order to preserve independence across sessions and to avoid over-representation of one animal in the data (track session 3, 3, 3, 4, and 3 , duration of 60, 55, 26, 28, 74 min. for rats 1-5).

Animal location and head direction were captured at 30 Hz by video tracking of 2 head-mounted LEDs using an overhead camera. The linearized position along the track was found by projecting the x,y coordinates of the animal's position onto a hand-fitted spline model of the track (A = 0 m, B ≃ 10.3 m). Linearized velocity was smoothed with a Gaussian kernel (s.d. = 0.25 s) and epochs during which linearized speed is > 15 cm/s (RUN) or linearized speed is < 5 cm/s (STOP) were detected. Positive velocity indicates movement from A→B.

Place/direction tuning and unit selection

For each unit we constructed a joint tuning curve over linearized position (10 cm bins) and running direction (A→B and B→A) from all spikes emitted during RUN. This curve was smoothed in position (Gaussian kernel; s.d. = 5 cm). We excluded putative interneurons (mean firing rate > 5 Hz) and units with weak place-related firing (peak rate in tuning curve < 3 Hz), leaving 47, 34, 23, 32, and 21 spatially tuned single units in rats 1-5 (cluster quality measures: L-ratio 0.12 ± 0.17; isolation distance 17 ± 9; calculated using peak amplitude on each channel (Schmitzer-Torbert et al., 2005)).

Candidate replay events

A smoothed histogram (1 ms bins; Gaussian kernel, s.d. = 15 ms) was constructed of multi-unit activity (MUA) including all spikes with a peak amplitude greater than 100 μV on any channel, whether or not they are part of an isolated cluster. Mean and standard deviation of MUA during STOP was calculated, and candidate replay events were defined as epochs during which MUA was higher than the mean and peak rate was at least 3 standard deviations above the mean. Only candidate events within 30 s of RUN were analyzed to exclude possible sleep periods.

Position estimation and validation

We used a Bayesian reconstruction algorithm (Zhang et al., 1998) to compute the joint probability distribution of position and running direction from neuronal firing in non-overlapping time bins using the place-by-direction tuning curves described above (see Chapter 2). In cases

where only position estimates were needed we computed the marginal distribution of these estimates over position. In order to validate our estimation procedure, RUN epochs in each session were divided into 'training' and 'testing' periods (alternating 1-second epochs). We calculated tuning curves using data from the training period, and used these to estimate position and direction during the testing period. Confusion matrices were calculated to assess reconstruction accuracy across the track (Fig. 8.1B-F), and maximum likelihood estimates of position and running direction were compared with the rat's true behavior (Fig. 3.6A). Data from rat 5 was excluded from further analysis because of poor position estimation during RUN (median error 23 cm; uneven coverage of track).

Replay detection

We define 'replay' as a sequence of hippocampal firing patterns that encodes a trajectory along the track at a constant velocity (Fig. 3.7). The most likely such trajectory is detected using a line-finding algorithm across the set of position estimates obtained during each candidate event (see below). Each replay trajectory is characterized by its velocity, location on the track, and its likelihood ('replay score'). For each candidate event, replay scores were compared to score distributions of three types of shuffled versions of the data to test for significance (Monte Carlo p-value < 0.01 for each shuffle type).

To detect the constant-velocity trajectory that best describes the series of position estimates during a candidate replay event, we used a method similar to line finding in a 2-dimensional image, employing a modified discrete approximation to the Radon transform (Toft, 1996). Each trajectory is defined by its velocity (V) and starting location (ρ). For a candidate event consisting of n position estimates calculated at an interval of Δt (20 ms), the average likelihood R that the rat is within distance d of a particular trajectory is given by:

$$R(V, \rho) = \frac{1}{n} \sum_{k=0}^{n-1} Pr(|pos - \rho + V \cdot k \cdot \Delta t)| \le d) \qquad (5)$$

(For those time bins k when a trajectory would specify a location beyond the end of the track, the trajectory is defined as specifying 'any location', and the median probability of all

possible locations is taken as the likelihood). The value of d was empirically set to 15 cm to allow for detection of replays with small local variations in velocity.

To determine the most likely replay trajectory, we densely sampled the parameter space to find the values of V_{max} and ρ_{max} that maximize R. The value of R_{max} is a measure of the goodness-of-fit of the detected trajectory and is reported as 'replay score' for the candidate event (ranging between 0-100%).

To test whether the replay score for a particular event is higher than expected by chance, we compared it to the distribution of replay scores obtained after applying the same replay detection procedure to shuffled versions of the position estimates (Supplementary Fig. 2e-g). The shuffles tested were: 1) 'Column-cycle' shuffle: circularly shifting the estimate at each time bin by a random distance (to control for chance alignment of locations along the trajectory); 2) 'Cell identity' shuffle: randomly permuting the mapping from spike trains to tuning curves (to control for patterns generated solely by spiking dynamics, such as static trajectories resulting from prolonged bursting of a single unit, or diagonal trajectories resulting from sequential bursting of any two units, etc.); and 3) 'Pseudo-event' shuffle: assembling 'pseudo-events' of identical duration to the original candidate events by making random draws of position estimates from other candidate events (to control for bias towards particular locations by the estimator). 1500 of each type of shuffle were performed for each event. Only candidate events with a Monte Carlo p-value (Davison and Hinkley, 1997) < 0.01 for all shuffle types were classified as significant replay.

Replay order analysis

To determine if significant replay events represented forward or reverse replay we computed a 'replay order' score. For each candidate event containing significant replay a 'replay order' score is computed as the difference between the probabilities of running in the A→B and B→A directions, conditional on the position being on (within d cm of) the replayed trajectory. Thus for a candidate event consisting of n time bins, we can write:

$$\text{replay order} = \frac{\sum\limits_{k=1}^{n} AB_k - BA_k}{\sum\limits_{k=1}^{n} AB_k + BA_k} \cdot sgn(V_{max})$$

where AB_k (BA_k) is the probability that the position is on the trajectory and running direction is A→B (B→A) for time bin k. The replay order score has a value ranging from -1 to 1. The magnitude of the score reflects the degree to which the replayed trajectory is dominated by either of the directional representations, and the sign of the score reflects whether or not that dominant representation agrees with the direction of the replayed trajectory (positive; 'forward replay') or disagrees with it (negative; 'reverse replay'). Values close to zero indicate that on average both directional representations contribute equally ('mixed replay').

To test if the magnitude of the replay order score for a replay event is statistically significant, it was compared to the distribution of replay order magnitudes of 2000 randomly-generated pseudo-events of the same duration. Pseudo-events used for significance testing were generated by drawing actual AB_k/BA_k pairs at random from all significant replay events within the same session. Replay events with a Monte Carlo p-value < 0.05 are classified as 'forward' or 'reverse' replay; the remaining events are classified as 'mixed' replay. Throughout the paper, 'significant replay' includes forward-, reverse-, and mixed-order replay.

To measure overall bias of replay towards forward and reverse replay (scores of +1 and -1) and away from mixed replay (score of 0), we performed a one-sided two-sample Kolmogorov-Smirnov test on the absolute values of the observed replay order scores and of the scores obtained under pseudo-event shuffling, as described above.

Replay trajectory analysis

For each detected replay trajectory, we calculate whether that trajectory lies mostly ahead of or behind (along the track) the rat's true position, and report the fraction of events lying mostly ahead. We test the significance of this measure by non-parametric bootstrapping. The chance level pooled across animals is estimated by randomly pairing within each session the observed replay trajectories to the location of the rat at the time of the replay events (2000 simulations,

Monte Carlo p-value reported), under the null hypothesis that these two variables are independent. Because the rats spent a significant amount of time at the ends of the track facing away from the track, the chance level for replay trajectories lying behind the animal is higher than for those lying ahead of the animal.

The same approach is used to analyze the relationship between the rat's location and the start and end locations of detected replay trajectories. The test statistic in this case is the fraction of start or end locations within 50 cm of the rat's true location. Similar results were obtained for thresholds ranging from 25 cm to 2 m.

Ripple detection and ripple-triggered analyses

We used a variation of Skaggs' (Skaggs et al., 2007) ripple detection procedure, which allows for the detection of closely-spaced ripples. The ripple amplitude at each recording site was estimated by band-pass filtering the local field potential (LFP) signal between 150-250 Hz, then taking the absolute value of the Hilbert-transformed signal (Siapas et al., 2005). The mean ripple amplitude across all recording sites was then smoothed (Gaussian kernel, s.d. = 12.5 ms) to give a single continuous measure of ripple activity. Individual ripples were detected as local peaks in this signal with an amplitude larger than 2.5 s.d. above the mean (mean and s.d. measured during STOP epochs). Ripple emission rates were calculated separately for each replay event, and compared with the mean ripple emission rate across all non-CAND STOP periods using a one-sample *t*-test. For plots in Figs. 5B and 5D (but not for any statistical analysis), ripple times were aligned (+/- ~2 ms) to the closest ripple cycle peak of the channel being plotted, in order to show local ripple structure.

Comparisons between ripple and non-ripple times during replay events were performed using either: a 1-sided 2-sample *t*-test assuming unequal variances (used for MUA; single unit firing rate); or, if the data did not appear to be normally distributed, the Mann-Whitney U test (used for mode of estimate; replay line error). For all tests, we used the same 20 ms non-overlapping time bins used for position reconstruction, and the comparison was between bins that contained a detected ripple and those that did not.

References:

August, D.A., and Levy, W.B. (1999). Temporal sequence compression by an integrate-and-fire model of hippocampal area CA3. J. Comput. Neurosci. 6, 71-90.

Battaglia, F.P., Sutherland, G.R., and McNaughton, B.L. (2004). Hippocampal sharp wave bursts coincide with neocortical "up-state" transitions. Learn. Mem. 11, 697-704.

Brown, E.N., Frank, L.M., Tang, D., Quirk, M.C., and Wilson, M.A. (1998). A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. J. Neurosci. 18, 7411-7425.

Buckner, R.L., and Carroll, D.C. (2007). Self-projection and the brain. Trends Cogn. Sci. 11, 49-57.

Buzsáki, G. (1989). Two-stage model of memory trace formation: a role for "noisy" brain states. Neuroscience 31, 551-570.

Buzsáki, G., Horvath, Z., Urioste, R., Hetke, J., and Wise, K. (1992). High-frequency network oscillation in the hippocampus. Science 256, 1025-1027.

Buzsáki, G., Leung, L.W., and Vanderwolf, C.H. (1983). Cellular bases of hippocampal EEG in the behaving rat. Brain Res. Rev. 6, 139-171.

Canto, C.B., Wouterlood, F.G., and Witter, M.P. (2008). What does the anatomical organization of the entorhinal cortex tell us? Neural Plast. 2008, 381243.

Chrobak, J.J., and Buzsáki, G. (1996). High-frequency oscillations in the output networks of the hippocampal-entorhinal axis of the freely behaving rat. J. Neurosci. 16, 3056-3066.

Clemens, Z., Mölle, M., Eross, L., Barsi, P., Halász, P., and Born, J. (2007). Temporal coupling of parahippocampal ripples, sleep spindles and slow oscillations in humans. Brain 130, 2868-2878.

Csicsvari, J., O'Neill, J., Allen, K., and Senior, T. (2007). Place-selective firing contributes to the reverse-order reactivation of CA1 pyramidal cells during sharp waves in open-field exploration. Eur. J. Neurosci. 26, 704-716.

Davison, A.C., and Hinkley, D.V. (1997). Bootstrap methods and their application (Cambridge, UK ; New York, NY, USA: Cambridge University Press).

Diba, K., and Buzsáki, G. (2007). Forward and reverse hippocampal place-cell sequences during ripples. Nat. Neurosci. 10, 1241-1242.

Fenton, A.A., Kao, H.Y., Neymotin, S.A., Olypher, A., Vayntrub, Y., Lytton, W.W., and Ludvig, N. (2008). Unmasking the CA1 ensemble place code by exposures to small and large environments: more place cells and multiple, irregularly arranged, and expanded place fields in the larger space. J. Neurosci. 28, 11250-11262.

Foster, D.J., and Wilson, M.A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. Nature 440, 680-683.

Gelbard-Sagiv, H., Mukamel, R., Harel, M., Malach, R., and Fried, I. (2008). Internally Generated Reactivation of Single Neurons in Human Hippocampus During Free Recall. Science 322, 96-101.

Isomura, Y., Sirota, A., Ozen, S., Montgomery, S., Mizuseki, K., Henze, D.A., and Buzsáki, G. (2006). Integration and segregation of activity in entorhinal-hippocampal subregions by neocortical slow oscillations. Neuron 52, 871-882.

Jackson, W.B. (1982). Norway Rat and Allies. In Wild Mammals of North America, Chapman, J. A. and Feldhamer, G. A. eds., (Baltimore, Maryland: Johns Hopkins University Press) pp. 1077-1088.

Ji, D., and Wilson, M.A. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. Nat. Neurosci. 10, 100-107.

Johnson, A., and Redish, A.D. (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. J. Neurosci. 27, 12176-12189.

Kloosterman, F., van Haeften, T., and Lopes da Silva, F.H. (2004). Two reentrant pathways in the hippocampal-entorhinal system. Hippocampus 14, 1026-1039.

Kosslyn, S.M., Ball, T.M., and Reiser, B.J. (1978). Visual images preserve metric spatial information: evidence from studies of image scanning. J. Exp. Psychol. Hum. Percept. Perform. 4, 47-60.

Lee, A.K., and Wilson, M.A. (2004). A combinatorial method for analyzing sequential firing patterns involving an arbitrary number of neurons based on relative time order. J. Neurophysiol. 92, 2555-2573.

Lee, A.K., and Wilson, M.A. (2002). Memory of sequential experience in the hippocampus during slow wave sleep. Neuron 36, 1183-1194.

Lehn, H., Steffenach, H.A., van Strien, N.M., Veltman, D.J., Witter, M.P., and Haberg, A.K. (2009). A specific role of the human hippocampus in recall of temporal sequences. J. Neurosci. 29, 3475-3484.

Marshall, L., and Born, J. (2007). The contribution of sleep to hippocampus-dependent memory consolidation. Trends Cogn. Sci. 11, 442-450.

McNaughton, B.L., Barnes, C.A., and O'Keefe, J. (1983). The contributions of position, direction, and velocity to single unit activity in the hippocampus of freely-moving rats. Exp. Brain Res. 52, 41-49.

Mehta, M.R., Lee, A.K., and Wilson, M.A. (2002). Role of experience and oscillations in transforming a rate code into a temporal code. Nature 417, 741-746.

Mölle, M., Yeshenko, O., Marshall, L., Sara, S.J., and Born, J. (2006). Hippocampal sharp wave-ripples linked to slow oscillations in rat slow-wave sleep. J. Neurophysiol. 96, 62-70.

Muller, R.U., Bostock, E., Taube, J.S., and Kubie, J.L. (1994). On the directional firing properties of hippocampal place cells. J. Neurosci. 14, 7235-7251.

Nadasdy, Z., Hirase, H., Czurko, A., Csicsvari, J., and Buzsáki, G. (1999). Replay and time compression of recurring spike sequences in the hippocampus. J. Neurosci. 19, 9497-9507.

O'Keefe, J., and Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain Res. 34, 171-175.

O'Keefe, J., and Nadel, L. (1978). The hippocampus as a cognitive map (Oxford; New York: Clarendon Press; Oxford University Press).

Schacter, D.L., Addis, D.R., and Buckner, R.L. (2007). Remembering the past to imagine the future: the prospective brain. Nat. Rev. Neurosci. 8, 657-661.

Schmitzer-Torbert, N., Jackson, J., Henze, D., Harris, K., and Redish, A.D. (2005). Quantitative measures of cluster quality for use in extracellular recordings. Neuroscience 131, 1-11.

Shepard, R.N., and Metzler, J. (1971). Mental rotation of three-dimensional objects. Science 171, 701-703.

Siapas, A.G., Lubenov, E.V., and Wilson, M.A. (2005). Prefrontal phase locking to hippocampal theta oscillations. Neuron 46, 141-151.

Siapas, A.G., and Wilson, M.A. (1998). Coordinated interactions between hippocampal ripples and cortical spindles during slow-wave sleep. Neuron 21, 1123-1128.

Sirota, A., Csicsvari, J., Buhl, D., and Buzsáki, G. (2003). Communication between neocortex and hippocampus during sleep in rodents. Proc. Natl. Acad. Sci. U. S. A. 100, 2065-2069.

Skaggs, W.E., McNaughton, B.L., Permenter, M., Archibeque, M., Vogt, J., Amaral, D.G., and Barnes, C.A. (2007). EEG sharp waves and sparse ensemble unit activity in the macaque hippocampus. J. Neurophysiol. 98, 898-910.

Steriade, M. (2006). Grouping of brain rhythms in corticothalamic systems. Neuroscience 137, 1087-1106.

Stickgold, R. (2005). Sleep-dependent memory consolidation. Nature 437, 1272-1278.

Suzuki, S.S., and Smith, G.K. (1987). Spontaneous EEG spikes in the normal hippocampus. I. Behavioral correlates, laminar profiles and bilateral synchrony. Electroencephalogr. Clin. Neurophysiol. 67, 348-359.

Toft, P.A. (1996). The Radon Transform - Theory and Implementation. PhD thesis, Technical University of Denmark. URL: http://petertoft.dk/PhD/

Wilson, M.A., and McNaughton, B.L. (1994). Reactivation of hippocampal ensemble memories during sleep. Science 265, 676-679.

Wilson, M.A., and McNaughton, B.L. (1993). Dynamics of the hippocampal ensemble code for space. Science 261, 1055-1058.

Wolansky, T., Clement, E.A., Peters, S.R., Palczak, M.A., and Dickson, C.T. (2006). Hippocampal slow oscillation: a novel EEG state and its coordination with ongoing neocortical activity. J. Neurosci. 26, 6213-6229.

Ylinen, A., Bragin, A., Nadasdy, Z., Jando, G., Szabo, I., Sik, A., and Buzsáki, G. (1995). Sharp wave-associated high-frequency oscillation (200 Hz) in the intact hippocampus: network and intracellular mechanisms. J. Neurosci. 15, 30-46.

Zhang, K., Ginzburg, I., McNaughton, B.L., and Sejnowski, T.J. (1998). Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. J. Neurophysiol. 79, 1017-1044.
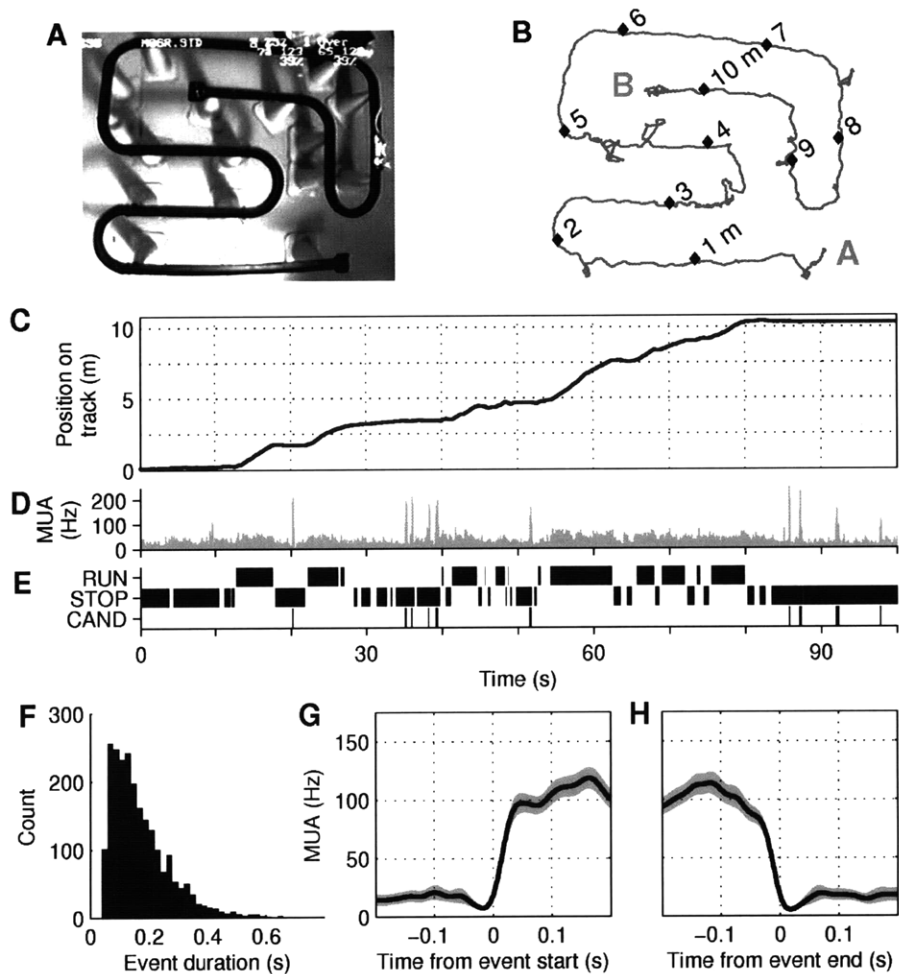
Figure 3.1: Behavior and candidate replay events

A, Top view of the 10.3 m-long track. Rat visible at right. **B,** Head position during 100 seconds of exploration. Labels 'A' and 'B' denote the 2 ends of the track as used throughout the paper. **C,** Linearized position (meters from 'A'); same data as (B). **D,** Multi-unit activity (MUA) across all electrodes. Note distinct peaks corresponding to elevated population activity. **E,** Identified periods of RUN (>15 cm/s), STOP (<5 cm/s), and candidate replay events (CAND; extracted from the MUA in D). **F,** Histogram of candidate event durations in all rats. **G,H,** Average MUA aligned to start (G) and end (H) of long (>250 ms) candidate events for rat 1. Note steep onset and offset of events.
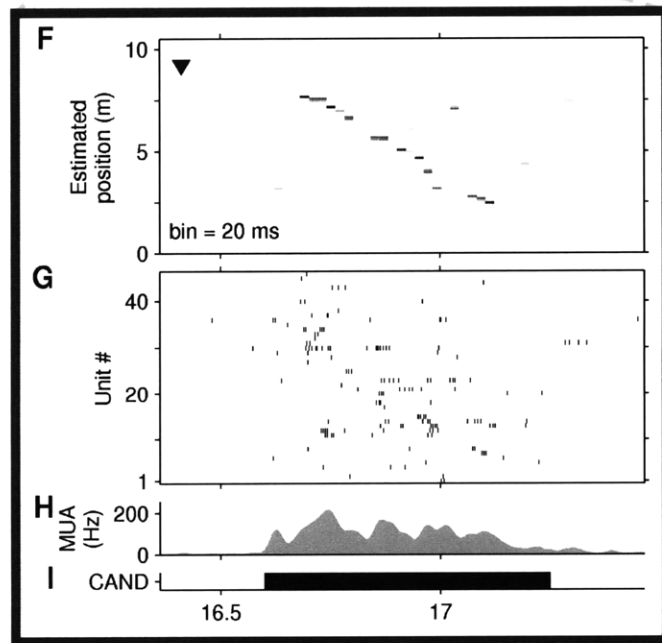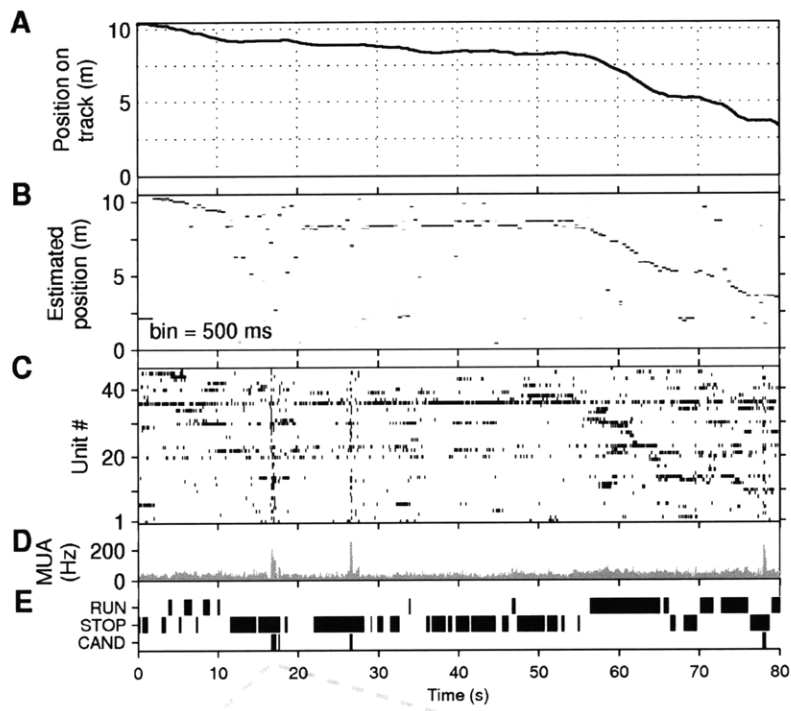
Figure 3.2: Replay detection using position reconstruction.

Figure 3.2: Replay detection using position reconstruction.

A-E, Behavior and position reconstruction for an 80-second epoch during which rat 1 runs approximately 7 m (from 10.3m to 3.5m), while pausing frequently. **A,** True position of animal. **B,** Estimated position. Each column is a probability density function estimated from unit activity in a 500 ms window. (White: p=0; black: p=1). **C,** Raster plots of spike times. Units are ranked by their preferred firing location; unit 1 has a place field closest to 0 m. Note bursts at 17 s, 27 s, and 78 s, which recruit a large fraction of all units. **D,** Multi-unit activity (MUA; average spike rate per tetrode, including unclustered spikes). **E,** Identified periods of RUN and STOP and candidate replay events (CAND). **F-I,** Position reconstruction applied to a candidate event revealing extended replay. **F,** Estimated position (20 ms bins) describes a trajectory from 8 m to 2 m while the animal remains stationary at 9 m (black arrowhead). The direction of the arrowhead indicates that the animal is facing in the B→A direction. **G,** Raster plot of unit firing. **H,** MUA. **I,** Candidate event. This event is the third example shown in Movie 2.
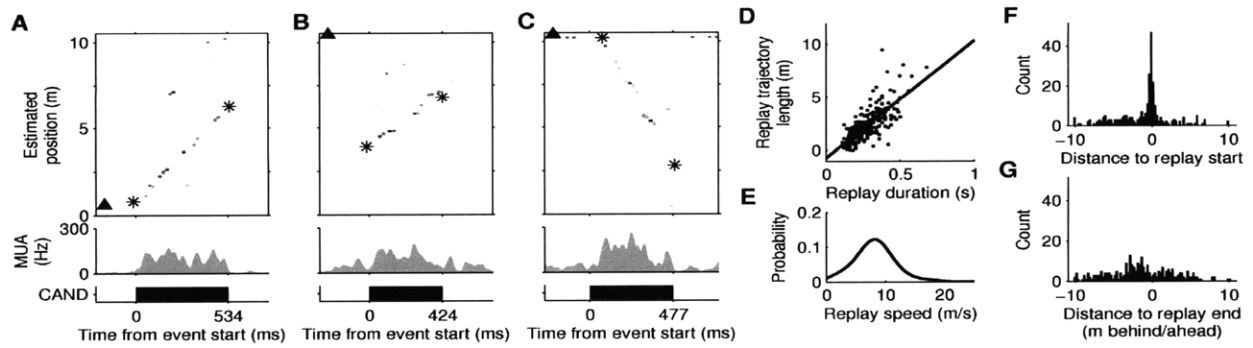
Figure 3.3: Extended replay

A-C, Examples of extended replay from rats 1-3. *Top:* Estimated position (asterisks indicate start/end of detected linear trajectory). *Middle:* MUA. *Bottom:* extent of replay event. **D,** Length of replayed trajectory vs. event duration for all replays. Solid line: linear regression (slope = 11.1 m/s; R^2 = 0.59; $p<10^{-10}$). **E,** Kernel density estimate (Gaussian kernel, width = 1.5 m/s) of the distribution of replay speeds. **F-G,** Distribution of start (F) and end (G) locations of replay trajectories relative to the animal's position and heading on the track. A negative distance indicates the replayed trajectory starts or ends behind the animal (along the track).
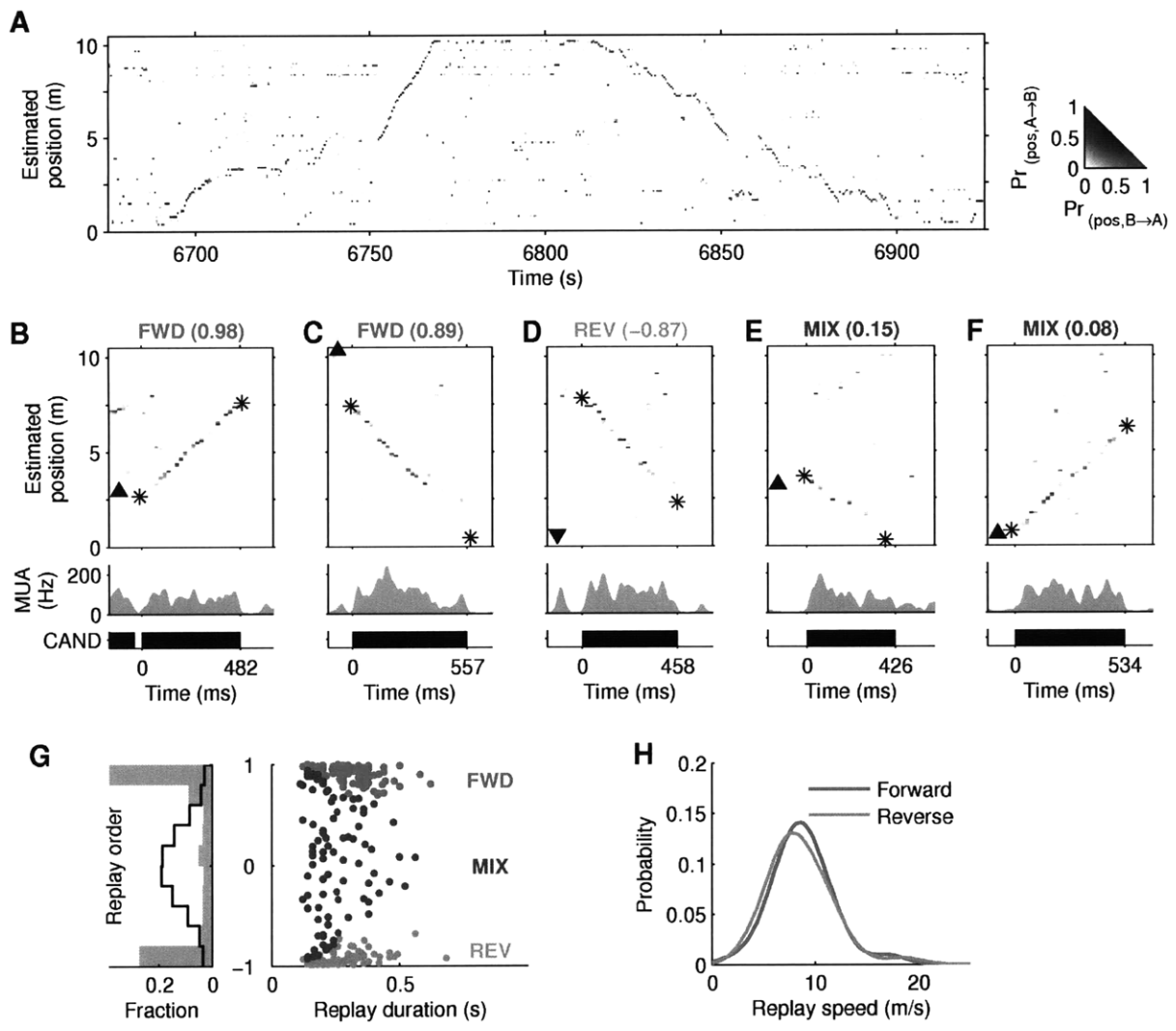
Figure 3.4: Forward and reverse extended replay.

Figure 3.4: Forward and reverse extended replay.

A, Joint reconstruction of position and running direction (500 ms bins). Color indicates estimated running direction (see inset). Direction is correctly estimated for both the A→B (6750-6770 s) and B→A directions (6820-6850 s). **B-F,** Examples of forward (FWD), reverse (REV), and mixed (MIX) replay from rat 1, each labeled with its replay order score. The events in (B) and (C) are the first 2 examples shown in Movie 2. *Top:* Joint position and direction estimates (20 ms bins). Arrowhead indicates animal's position and facing direction. Asterisks indicate start and end of detected replay trajectory. *Middle:* Multi-unit activity. *Bottom:* extent of replay event. **B,** Forward replay in the A→B direction proceeding ahead of the animal. **C,** Forward replay in the B→A direction, starting 2 m behind the animal and proceeding behind the animal. **D,** Reverse replay, starting remotely and proceeding towards the animal. Trajectory is similar to (C), but this is a reverse-ordered replay because the estimated running direction (i.e. A→B (blue)) does not agree with the direction in which the replay proceeds (i.e. from B→A). **E,** Mixed replay proceeding behind the animal. **F,** Mixed replay proceeding ahead of the animal. This event begins as an apparently forward-ordered replay then switches to reverse-ordered after ~240 ms. **G,** *Left:* distribution of observed (gray bars) and expected (pseudo-event shuffles; black line) replay order scores. *Right:* scatter plot of replay order score and replay duration for all significant replay events in all animals. green: forward replay; yellow: reverse replay; gray: mixed replay. **H,** Kernel density estimates (Gaussian kernel, width = 1.5 m/s) of the distribution of replay speeds for forward and reverse replay.
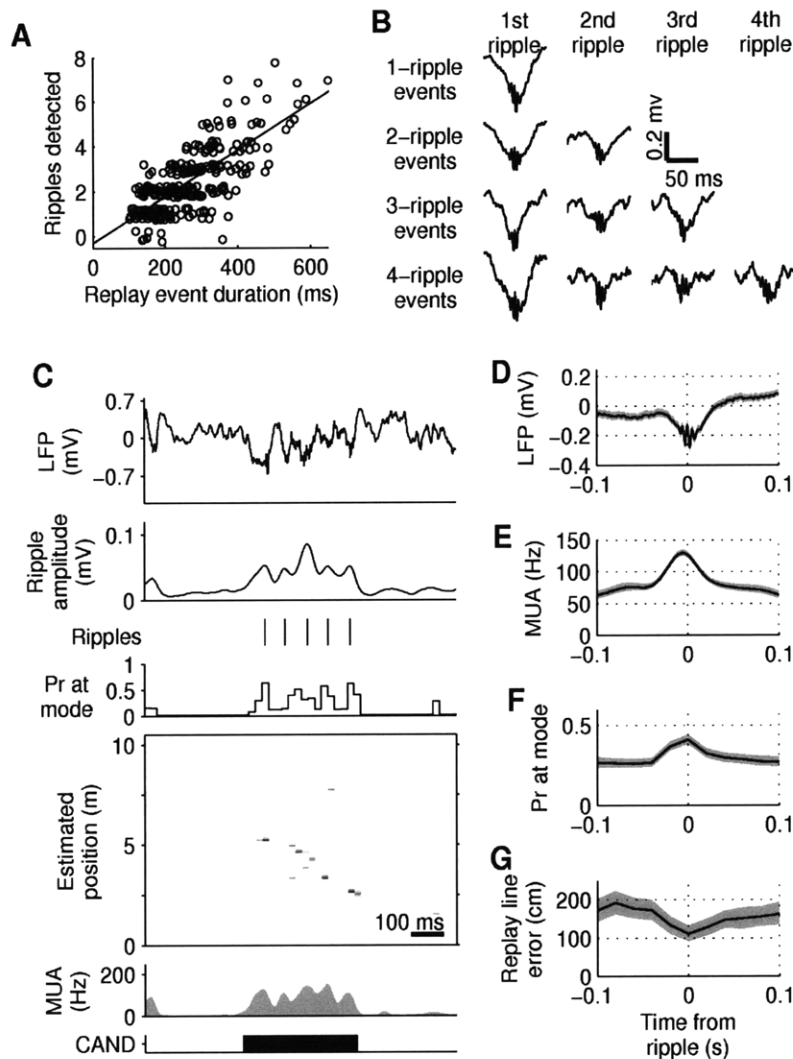
Figure 3.5: Replay spans multiple ripples.

A, Scatter plot of number of detected ripples during significant replay events in all animals as a function of replay event duration. Random jitter added in y-axis for visualization. Linear regression: 9.9 ripples/s, R2 = 0.56. B, Ripple-triggered averages of wide-band hippocampal local field potential (LFP) during replay events in rat 1. Even in multiple-ripple events, each ripple is associated with a sharp wave in the LFP. C, Example of multiple ripples during a single extended replay event. From top to bottom: LFP; average amplitude in the ripple band (150-250 Hz) across all electrodes; detected ripples; probability at the mode of position estimate; position estimate; MUA; candidate event time. D-G, Ripple-triggered averages of: wide band LFP (D), MUA (E), mode of position estimate (F), and error between estimated location and replay trajectory (G) for replay events in rat 1. Shaded regions: 95% confidence intervals for the mean.
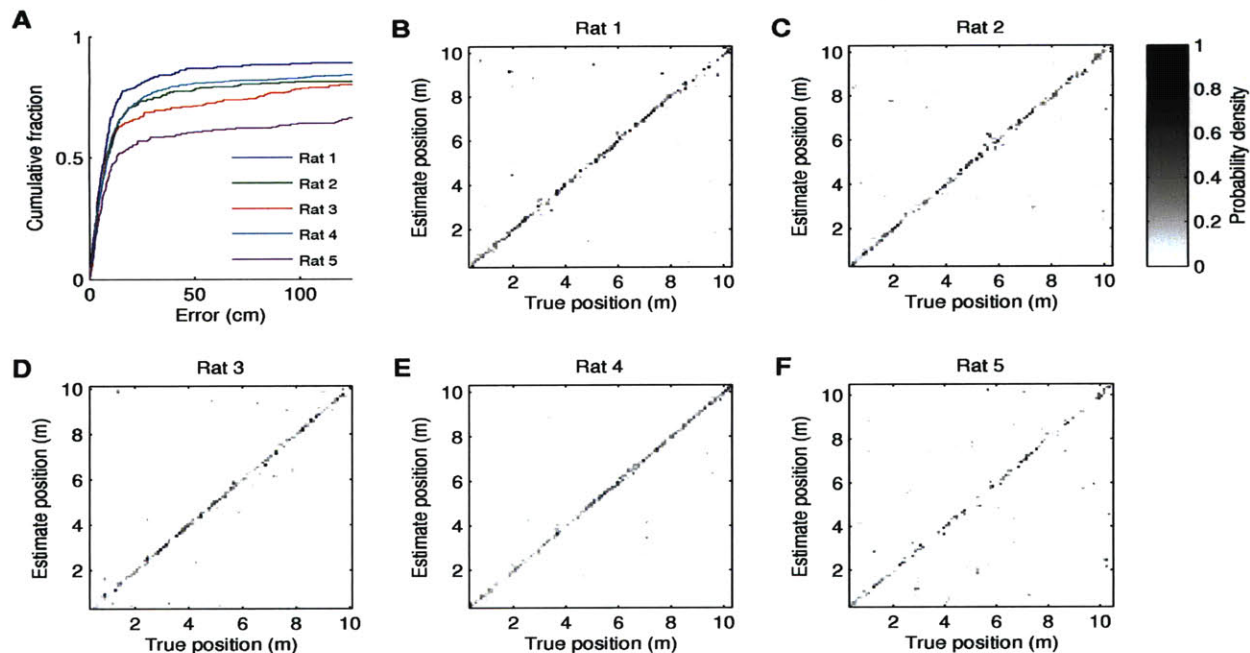
Figure 3.6: Position estimation error

Error measures for position estimation during RUN for rats 1-5 (500 ms time bins; test set validation). **A,** Cumulative distribution functions of the error of the maximum likelihood estimate of position **B-F,** Confusion matrices for rats 1-5. Each column corresponds to the average estimated probability distribution over position for a given true position (10 cm bins). Systematic estimation errors (i.e. 'confusing' one location for another) would appear as power off the identity diagonal. Notice that for animals 1-4 we can accurately estimate the rat's position for most locations on the track. Rat 5 was excluded from replay analysis due to poor reconstruction quality during RUN.
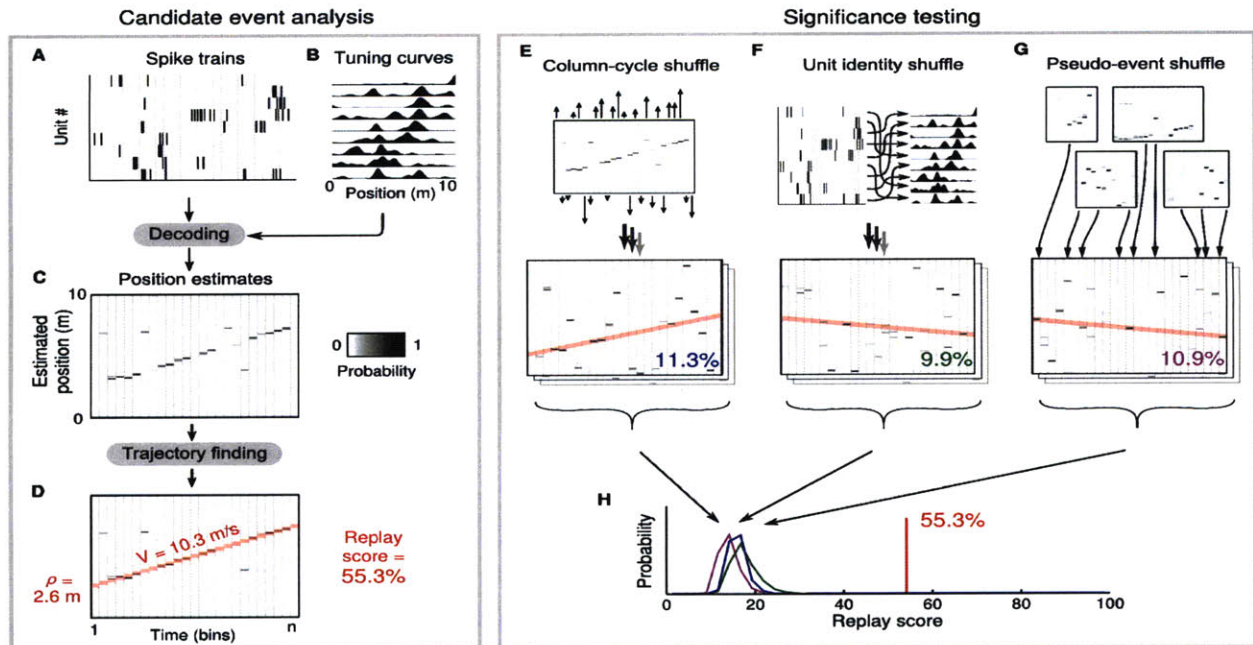
Figure 3.7: Neural decoding, replay detection and significance testing

A-D, Neural decoding and replay detection for a single candidate event. Spiking activity is divided into non-overlapping 20 ms bins (A). We estimate position within each time bin by decoding the ensemble activity using the spatial tuning curves (B) obtained during RUN. This gives an estimated probability distribution over position (PDF) for each time bin (plotted as columns in panel C). Replay appears as a systematic shift of the peak in the PDFs over time. To detect linear replay trajectories, we densely sample all possible lines (30 cm wide) and for each line compute a score equal to the mean likelihood that the rat was on the specified trajectory throughout the event. The line with the highest score (red shaded line in D) is selected as the candidate replay trajectory, and its score is reported as the 'replay score' for the event. **E-G,** The statistical significance of a candidate replay trajectory is assessed by repeating the scoring procedure on shuffled versions of the same data. **E,** In the 'column-cycle' shuffle, the PDF at each time bin is circularly shifted by a random distance. **F,** In the 'unit identity' shuffle, the mapping from each unit's spiking record to its spatial tuning curve is randomly permuted. **G,** The 'pseudo-event' shuffles are constructed by substituting each PDF in a candidate event for a PDF drawn at random from any other candidate event. **H,** Each type of shuffle is repeated 1500 times for each event to obtain three sample distributions. The true replay score (red line) is compared with each of these distributions, and the largest of the 3 resulting Monte Carlo p-values is conservatively reported as the significance level for the event.
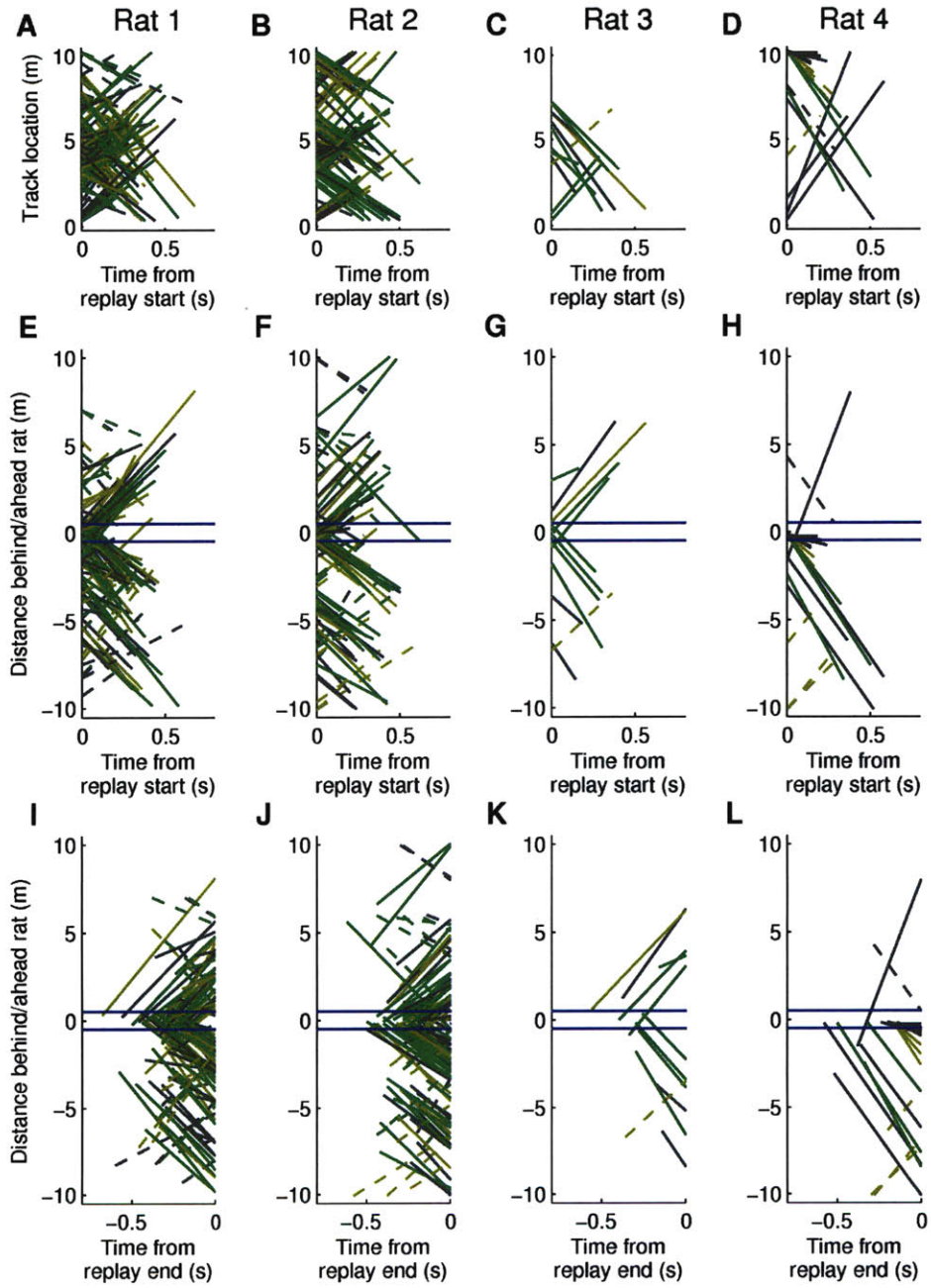
Figure 3.8: Replay trajectory locations

Figure 3.8: Replay trajectory locations

A-D, Significant (p<0.01) replay trajectories, aligned to the start time of the replay event. Note that replay can begin and end at any location on the track, not only at reward locations. Green: forward replay; yellow: reverse replay; gray: mixed replay; dashed lines: strictly remote replay (see text). **E-L,** The same replay trajectories as in A-D, but plotted relative to the animal's location at the time of the replay (positive distance values = locations ahead of the animal along the track; negative distance values = behind the animal). Replays plotted aligned to replay start (E-H) or end (I-L) time. Replay start locations, but not end locations, are significantly biased to start within 0.5 m of the current location (blue lines). Line colors as in A-D.
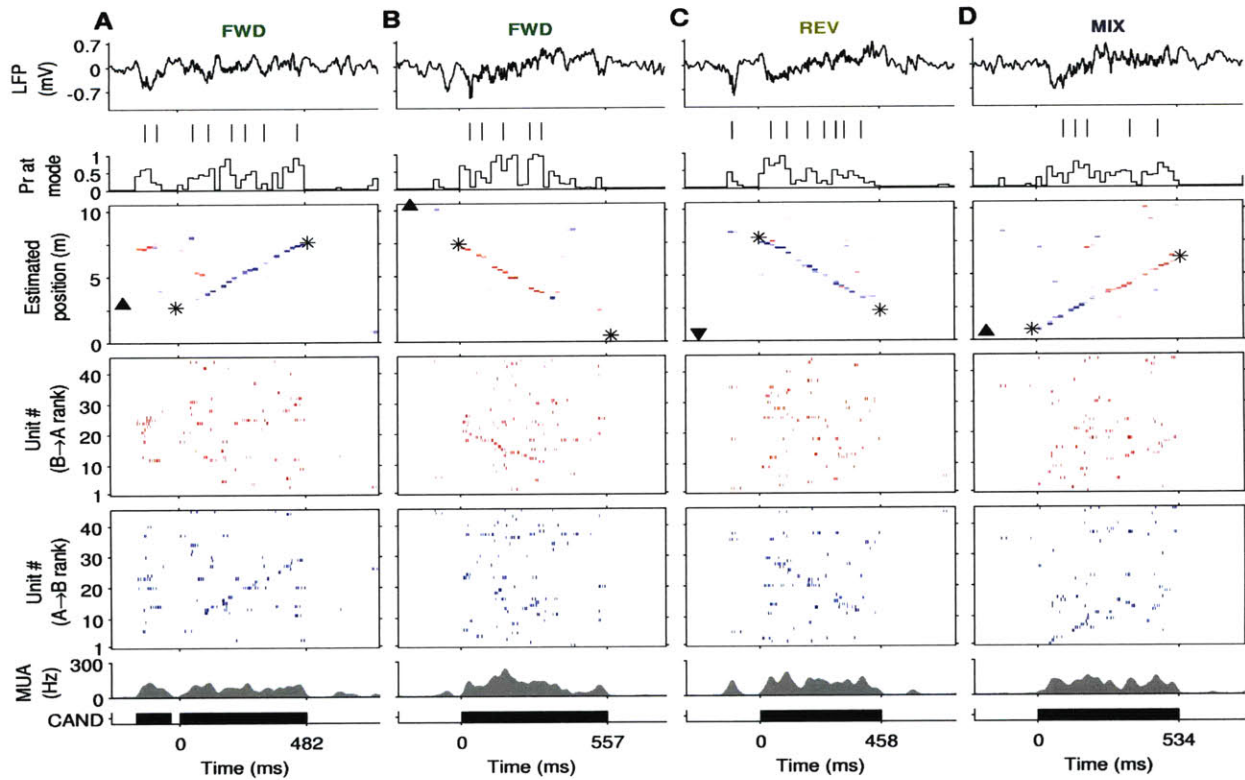
Figure 3.9: Replay examples for rat 1

A-D, Examples of candidate events containing forward (FWD), reverse (REV), and mixed (MIX) replay. For each example the following is shown (top to bottom): wide band LFP; detected ripples; mode of position reconstruction; position reconstruction; spike raster plot of all units sorted according to their preferred firing location when rat is running B→A direction; spike raster plot of all units sorted according to their preferred firing location when rat is running in A→B direction; Multi-unit activity across all tetrodes (MUA; average rate per tetrode including unclustered spikes); candidate event (CAND) start and end time. Black triangle in top panel indicates true location and facing direction of the rat throughout the event. Asterisks in top panel indicate start and end of detected replay trajectory. (Events are the same as those in Fig. 3.4B, C, D & F).
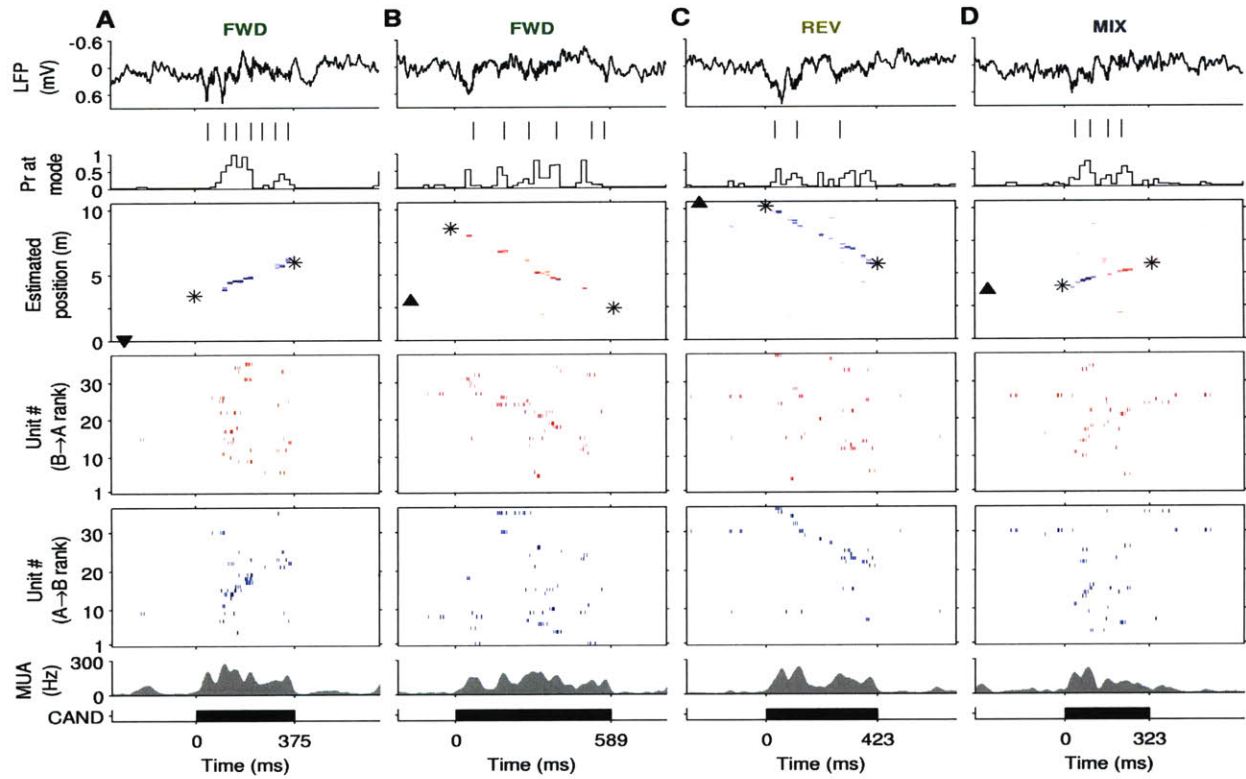
Figure 3.10: Replay examples for rat 2

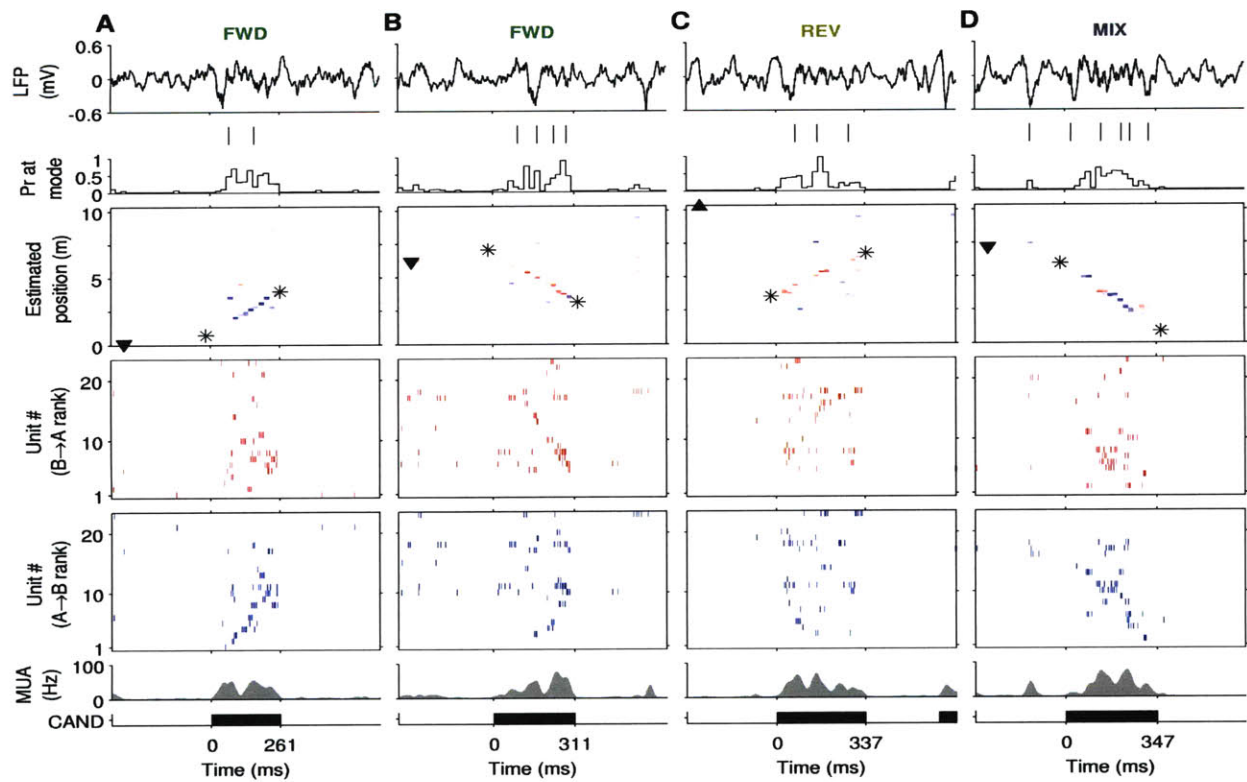All panels as for Fig. 3.9. LFP scale inverted to show sharp waves as downward-going.

Figure 3.11: Replay examples for rat 3

All panels as for Fig. 3.9. LFP high-pass filtered (5 Hz cutoff) to remove movement artifacts.
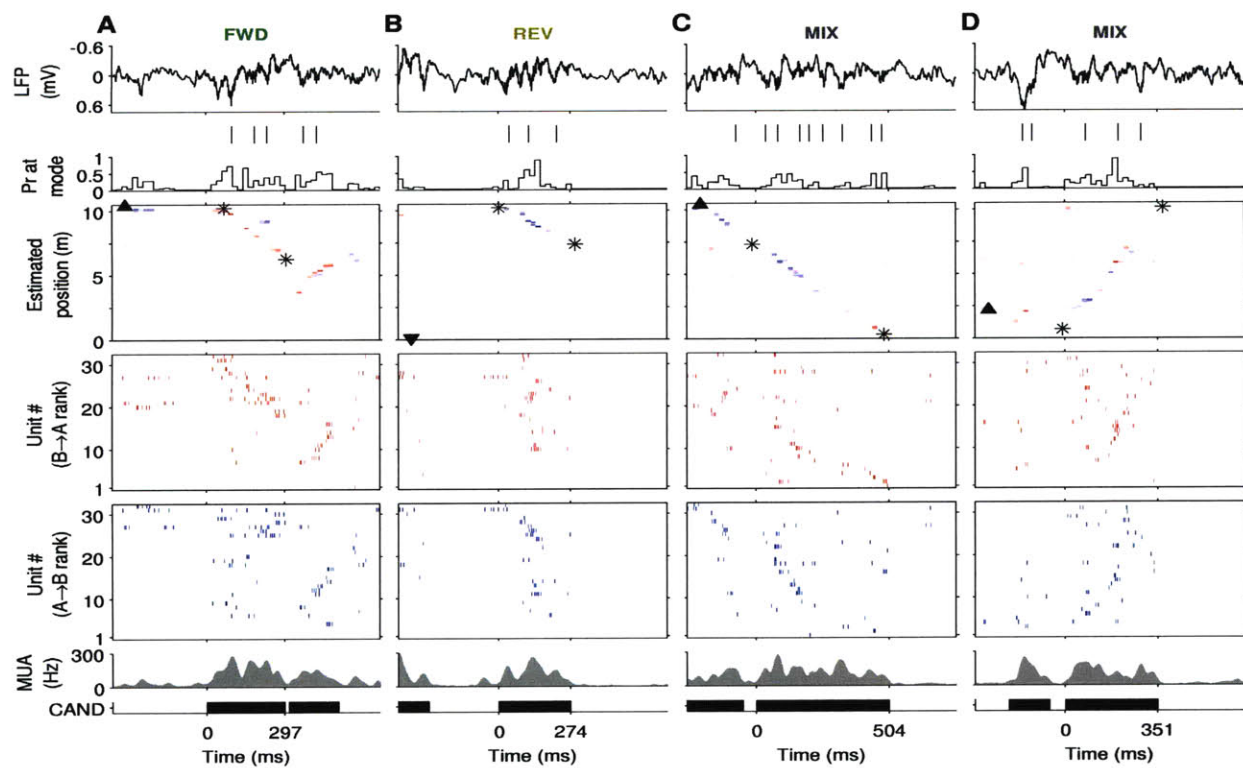
Figure 3.12: Replay examples for rat 4

All panels as for Fig. 3.9. LFP scale inverted to show sharp waves as downward-going.

Movie 1: Position estimation during RUN

Rat 1 exploring the track. Audio represents spiking of all units used for decoding. Moving green circle indicates the rat's true head position. Arrowheads at each location on the track (10 cm spacing) are scaled according to the estimated probability that the rat is occupying that location, and colored according to the estimated direction of travel (blue: A→B, red: B→A). Estimates are based only on spiking in non-overlapping 500 ms time windows. Note that when the rat stops running the firing pattern becomes irregular and the accuracy of the position estimate degrades.

Movie 2: Examples of extended replay during STOP

Slow-motion (0.2x) video of extended replay events observed when rat 1 was stopped on the track. Audio and video are the same as in Supplementary Movie 1, except that estimates are calculated using 20 ms time bins. Time is measured from the start of the candidate event. The three events shown are examples of: forward replay of an A→B trajectory (same event as Figure 3.4B); remote forward replay of a B→A trajectory (same event as Figure 3.4C); and reverse replay of an A→B trajectory (proceeds from B→A; same event as Figure 3.2F-I).

Chapter 4:

Discussion

Contributions of the thesis

We developed a novel probabilistic decoding method for the detection of neural reactivation in large ensembles of neurons during off-line states (Chapter 2). An important advantage of this method is that it enables the characterization of the mnemonic content of individual reactivation events (in our case replayed spatial trajectories). This content could then be correlated to ongoing behavior and electrophysiological states.

We then used this method to study hippocampal reactivation during an ethologically relevant, but previously unstudied behavior: the exploration of a large environment (Chapter 3). We found that hippocampal replay was common during exploration of a large environment. Analysis of the replayed trajectories showed that these could proceed over long distances, and that the hippocampal ensemble could recapitulate previous experience in either the forward or reverse order. Replay was highly temporally compressed, and replayed trajectories could begin at remote locations. We also found that extended replay spanned trains of closely-spaced sharp wave-ripple (SWR) events, suggesting that these longer sequences may be composed of several shorter subsequences.

In this chapter we will first consider the implications of these results for existing models of replay generation, and propose a novel hippocampal-cortical mechanism that accounts for the observed extended replay. Lastly, we will discuss how our results relate to the evolving views of the potential cognitive role for hippocampal replay.

Mechanisms of extended SWR-associated replay

SWRs have long been hypothesized to represent the 'quantum' of the hippocampo-neocortical dialogue (Buzsáki, 1989; Buzsáki, 1996). Structured hippocampal replay has consistently been shown to occur in association with single SWRs, and SWRs have been shown to be temporally correlated with cortical spindling (Siapas and Wilson, 1998) and the neocortical slow oscillation (Battaglia et al., 2004; Mölle et al., 2006). Further, SWRs are associated with increased activity in neocortical cells.

Since the first reports of SWRs, it has been noted that in addition to appearing in isolation, they can appear in closely-spaced trains (up to 7 ripples at ~10Hz) during quiet wakefulness (Buzsáki et al., 1983; O'Keefe and Nadel, 1978; Suzuki and Smith, 1987). However, no detailed

study has been made of this phenomenon, and no function has yet been ascribed to these trains.

We have shown that closely-spaced SWRs can contain extended replay of hippocampal spike sequences. The seemingly discrete nature of the individual events suggests a model in which repeated iterations of a simple, single-ripple replay mechanism might generate the replay of longer sequences.

Neuroanatomical substrate for SWR trains

An attractive hypothesis for the generation of SWR trains is suggested by the neuroanatomy of the hippocampus and the adjoining entorhinal cortex (EC), which together form two powerful re-entrant loops. A wide range of inputs from neocortical multimodal association areas converge onto the EC, which in turn projects to the various subfields of the hippocampus. These inputs pass chiefly along two unidirectional pathways that loop through the hippocampus, and give rise to strong reciprocal projections back to the EC. The two pathways are known as the 'indirect' (or 'trisynaptic') pathway, and the 'direct' pathway (also known as the 'monosynaptic' or 'temporoammonic' pathway). [1]

In support of this hypothesis, *in vitro* preparations in which the entorhinal cortex remains attached to the hippocampus can display oscillatory activity with strong similarities to the observed trains of SWRs. In one slice preparation, SWR-like population spikes were observed originating in area CA3, propagating through CA1 to the EC, then re-invading the DG and continuing through the trisynaptic pathway a second time (Borck and Jefferys, 1999). Even more strikingly, in an isolated whole guinea pig brain preparation, trains of population spikes were observed traveling around the entorhinal-hippocampal loop with a frequency of 16-25Hz (loop time 40-80 ms) in trains lasting a few hundred ms, accompanied by rhythmic firing of the principal neurons of CA3 and CA1(Pare et al., 1992). These preparations were put forward as models for epileptogenic activity, but our results suggest they may support normal brain function.

The looping model depends upon the propagation of signals from the deep layers of EC

---

1   The indirect pathway consists of projections from cells in the superficial layers of the EC to the granule cells of the dentate gyrus; these project in turn to pyramidal cells in area CA3 via the mossy fibres; which in turn project to the pyramidal cells of area CA1 via the Schaffer collaterals. The direct pathway consists of projections that 'shortcut' directly from superficial EC to area CA1. The output of both pathways is returned from CA1 to the deep layers of the EC, either directly or via the subiculum. Connections between the deep and superficial layers of EC 'close the loop.' SWR-related bursts of activity (whether due to EC input or generated within the HPC) could in principle propagate around either of these loops.

(where hippocampal outputs are received) to the superficial layers (where hippocampal inputs are generated). While firing in superficial EC in response to deep EC activation has been found to be weak under some experimental circumstances, (Chrobak and Buzsáki, 1996), recent electrophysiological studies have shown that the required neuronal connectivity is present in vivo (Kloosterman et al., 2004).

Several other functions have already been proposed for the hippocampal-entorhinal loop. It has been hypothesized that reverberant activity in the loop could support a working memory, either for single items or for short sequences. The EC is also uniquely positioned to compare hippocampal inputs and outputs, and this 'comparator' function has been hypothesized to support novelty detection on sensory inputs. Such models are not necessarily incompatible with the model we propose. For instance, the EC might generate a 'novelty signal' during exploration, and support chained replay while the animal is at rest.

Replay within single SWRs

The 'chained replay' hypothesis outlined above does not specify how replay is generated during single SWRs. Existing models of hippocampal replay during SWRs fall into two classes: recurrent neural network (RNN) models, and competitive queuing (CQ) models. RNN models involve learning each 'link' in a chain of stimuli during training (i.e. exploration), and storing these temporal associations in a highly recurrent (or self-connected) network. After learning, replay can be triggered by presentation of any learned input pattern, which triggers the next element in the sequence, and so on.

In CQ models, by contrast, the replayed sequence is not learned during behavior. Instead, it is encoded in the initial input presented to the network. The most active element (i.e. the cell receiving the strongest synaptic drive) is read out first, followed by the next most active element, and so on. If input strength is correlated with position in the sequence, then the result of this process is sequence readout. (CQ models of sequence generation have been proposed as models of working memory (Bullock and Rhodes, 2003)). In this model, cells with nearby place fields will be more excitable than those with more distant place fields, and so will fire at a lower latency, resulting in a near-to-far sequence as global inhibition is decreased during a sharp wave (Diba and Buzsáki, 2007; Foster and Wilson, 2006). The downward sweep of inhibition 'queries' the cells in the network, with the cells answering the query in decreasing order of input drive

received.

The CQ model of replay has important limitations. While an RNN can in principle replay a sequence of arbitrary length, sequences generated by CQ can only include cells that are already receiving some place-related input. We observe replay that proceeds for close to 10 meters, which is far larger than the scale of single place fields, and likely larger even than any sub-threshold place-related input. We also observe replay that continues across trains of ripples, which is not consistent with a prolonged query of a single input strength gradient. After discussing how these sequences may be similar to those seen during individual theta cycles, we will put forward a model to explain how multiple short sequences could be linked together.

Similarity to theta sequences

Short, temporally compressed sequences of place cell firing are also seen during running in association with the theta oscillation (Foster and Wilson, 2007), and a similar excitability-to-latency model has been proposed to explain these sequences (Mehta et al., 2002). In order to compare these two phenomena, we performed short-timescale decoding of the ensemble activity during running (Figure 4.1). As expected, we frequently observed instances where the estimated position sweeps ahead of the animal at a speed only slightly less than seen during SWR-associated replay. This is also apparent in a theta peak-triggered average of the position estimate about the current location (Fig. 4.1C-E). The distance covered during a theta sequence is ~30-50 cm over a time span of approximately 60 ms. Single SWRs have a similar duration and at ~8 m/s a replayed trajectory will also cover about 50 cm. These similarities suggest that a shared mechanism could underlie sequence expression in both states.

Several studies have calculated compression factors during theta and/or replay sequences. Both Skaggs et al. (1996) and Dragoi and Buzsáki (2006) report a compression factor of ~10 (relative to the behavioral time scale) during theta sequences. Lee and Wilson (2002) found that for SWR-associated replay during slow wave sleep, the compression factor was in the range of 16-20 (which is roughly in agreement with the ~8 m/s replay speed we observe, assuming a 40-50 cm/s run speed). Diba and Buzsáki (2007) directly compared the peak latency in the cross-correlation between cell pairs during SWR-associated replay and theta. From their figure 2 we can estimate that theta sequences were less compressed than replay sequences (~0.7x), roughly consistent with the above figures. Since the speed of sequence expression in the CQ model is a

function of how quickly the gradient is queried, such differences in 'virtual speed' could be the result of a faster lowering of inhibitory tone during an SWR.

Linking SWRs: replay, invert, repeat

These observations lead us to propose that prolonged replay is made up of multiple, chained, single-SWR replay events. As we have noted, the CQ replay model depends on the readout of structured input to place cells. After the first set of inputs has been queried by the first ripple, how are the second and subsequent activity gradients set up for readout?

As noted above, the EC both generates the bulk of the cortical input to the hippocampus, and also has continuous access to hippocampal outputs. We propose that during exploration, in addition to relaying sensory inputs to the hippocampus, projection cells in the superficial EC also learn which patterns of hippocampal output are associated with their activity. This learning could take place under the a simple time-symmetric competitive Hebbian learning rule (though synchronization of EC and hippocampal theta suggests that perhaps finer-grained temporal structure could also be exploited). The learning of this mapping would allow the EC, when later presented with a hippocampal output pattern, to invert this pattern in order to re-express the corresponding input pattern, even in the absence of any sensory input from higher cortical areas.

At the end of a single-SWR replay event that covered, for example, 1 m of track, the hippocampal outputs could be used by the EC to reconstruct the hippocampal inputs that would be present if the rat were actually present in that new location. Gradient replay on this new set of inputs could then proceed as if the rat were in the new location, and the process could be iterated to re-express longer sequences.

In this model, the initial input state need not correspond to the current location. The 'remote replay' trajectories that we observe could instead be seeded by reactivated cortical memory traces which contain enough structured sensory information to be capable of driving CQ replay.

Predictions and future experiments.

This model predicts that during prolonged hippocampal replay, position reconstruction based on recordings of EC grid cells (Fyhn et al., 2004) should encode identical trajectories through space.

Intracellular recordings of place cells (Lee et al., 2006; Lee et al., 2009) during replay would

help to distinguish between the RNN and gradient models of replay during single ripples. In gradient-based replay, but not in RNN replay, the level of activity at the beginning of a ripple would be expected to predict latency to firing.

Implications for hippocampal function

The neural decoding techniques we have employed work because the hippocampus continuously calculates a sparse code for the current spatial context. Indeed, it seems likely that the place cell network may be optimized for performing such a function.

Of what use is such a code? If it is available to higher brain areas during both behavior and replay, then it can function as an 'index' of a particular behavioral context (Teyler and DiScenna, 1986; Teyler and Rudy, 2007). Cells in higher cortical sensory areas that receive input from the hippocampus during behavior could learn the association between their activity and the corresponding hippocampal outputs. Later, during replay, when sequences of those hippocampal output patterns are generated, the activity of the cortical sensory cells could again be triggered, reactivating the original cortical representations.

Such reactivation has frequently been suggested as a means of supporting off-line memory consolidation processes, but it seems clear that such a phenomenon might also support memory recall processes. This latter view is less tenable if replay only occurs during sleep, and if hippocampal damage (and presumably impaired replay) does not affect recall of previously learned material.

It is now clear that in rodents, the behavioral contexts in which replay is observed (during brief pauses in exploration) are consistent with a role for replay in the evaluation or use of learned content. And it is becoming clear that human patients with impaired hippocampal function exhibit deficits in a constellation of cognitive abilities—episodic memory recall, imagination of future events, and spatial navigation—involving "self-projection" (Buckner and Carroll, 2007).

Could these deficits be due to the loss of ripple-associated replay in humans? Some evidence supports this view. Ripple oscillations are seen in the hippocampus of awake humans, and modulate the firing of individual hippocampal neurons (Le Van Quyen et al., 2008). In recordings taken between the learning and recall phases of a memory task, Axmacher et al. (2008) recently reported that ripples were more frequent during wakefulness than during sleep.

And recent single-unit recordings in the human hippocampus have demonstrated memory-specific neuronal reactivation during free recall of recently experienced episodes (Gelbard-Sagiv et al., 2008).

References:

Axmacher, N., Elger, C.E., and Fell, J. (2008). Ripples in the medial temporal lobe are relevant for human memory consolidation. Brain *131,* 1806-1817.

Battaglia, F.P., Sutherland, G.R., and McNaughton, B.L. (2004). Hippocampal sharp wave bursts coincide with neocortical "up-state" transitions. Learn. Mem. *11,* 697-704.

Borck, C., and Jefferys, J.G. (1999). Seizure-like events in disinhibited ventral slices of adult rat hippocampus. J. Neurophysiol. *82,* 2130-2142.

Buckner, R.L., and Carroll, D.C. (2007). Self-projection and the brain. Trends Cogn. Sci. *11,* 49-57.

Bullock, D., and Rhodes, B.J. (2003). Competitive Queuing for Planning and Serial Performance. In The Handbook of Brain Theory and Neural Networks, Arbib, Michael A. ed., (Cambridge, Mass.: MIT Press) pp. 241-244.

Buzsáki, G. (1996). The hippocampo-neocortical dialogue. Cereb. Cortex *6,* 81-92.

Buzsáki, G. (1989). Two-stage model of memory trace formation: a role for "noisy" brain states. Neuroscience *31,* 551-570.

Buzsáki, G., Leung, L.W., and Vanderwolf, C.H. (1983). Cellular bases of hippocampal EEG in the behaving rat. Brain Res. Rev. *6,* 139-171.

Chrobak, J.J., and Buzsáki, G. (1996). High-frequency oscillations in the output networks of the hippocampal-entorhinal axis of the freely behaving rat. J. Neurosci. *16,* 3056-3066.

Diba, K., and Buzsáki, G. (2007). Forward and reverse hippocampal place-cell sequences during ripples. Nat. Neurosci. *10,* 1241-1242.

Dragoi, G., and Buzsáki, G. (2006). Temporal encoding of place sequences by hippocampal cell assemblies. Neuron *50,* 145-157.

Foster, D.J., and Wilson, M.A. (2007). Hippocampal theta sequences. Hippocampus *17,* 1093-1099.

Foster, D.J., and Wilson, M.A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. Nature *440,* 680-683.

Fyhn, M., Molden, S., Witter, M.P., Moser, E.I., and Moser, M.B. (2004). Spatial representation in the entorhinal cortex. Science *305,* 1258-1264.

Gelbard-Sagiv, H., Mukamel, R., Harel, M., Malach, R., and Fried, I. (2008). Internally Generated Reactivation of Single Neurons in Human Hippocampus During Free Recall. Science *322,* 96-101.

Kloosterman, F., van Haeften, T., and Lopes da Silva, F.H. (2004). Two reentrant pathways in the hippocampal-entorhinal system. Hippocampus *14*, 1026-1039.

Le Van Quyen, M., Bragin, A., Staba, R., Crepon, B., Wilson, C.L., and Engel, J.,Jr. (2008). Cell type-specific firing during ripple oscillations in the hippocampal formation of humans. J. Neurosci. *28*, 6104-6110.

Lee, A.K., Epsztein, J., and Brecht, M. (2009). Head-anchored whole-cell recordings in freely moving rats. Nat. Protoc. *4*, 385-392.

Lee, A.K., Manns, I.D., Sakmann, B., and Brecht, M. (2006). Whole-cell recordings in freely moving rats. Neuron *51*, 399-407.

Lee, A.K., and Wilson, M.A. (2002). Memory of sequential experience in the hippocampus during slow wave sleep. Neuron *36*, 1183-1194.

Mehta, M.R., Lee, A.K., and Wilson, M.A. (2002). Role of experience and oscillations in transforming a rate code into a temporal code. Nature *417*, 741-746.

Mölle, M., Yeshenko, O., Marshall, L., Sara, S.J., and Born, J. (2006). Hippocampal sharp wave-ripples linked to slow oscillations in rat slow-wave sleep. J. Neurophysiol. *96*, 62-70.

O'Keefe, J., and Nadel, L. (1978). The hippocampus as a cognitive map (Oxford; New York: Clarendon Press; Oxford University Press).

Pare, D., deCurtis, M., and Llinas, R. (1992). Role of the hippocampal-entorhinal loop in temporal lobe epilepsy: extra- and intracellular study in the isolated guinea pig brain in vitro. J. Neurosci. *12*, 1867-1881.

Siapas, A.G., and Wilson, M.A. (1998). Coordinated interactions between hippocampal ripples and cortical spindles during slow-wave sleep. Neuron *21*, 1123-1128.

Skaggs, W.E., McNaughton, B.L., Wilson, M.A., and Barnes, C.A. (1996). Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. Hippocampus *6*, 149-172.

Suzuki, S.S., and Smith, G.K. (1987). Spontaneous EEG spikes in the normal hippocampus. I. Behavioral correlates, laminar profiles and bilateral synchrony. Electroencephalogr. Clin. Neurophysiol. *67*, 348-359.

Teyler, T.J., and DiScenna, P. (1986). The hippocampal memory indexing theory. Behav. Neurosci. *100*, 147-154.

Teyler, T.J., and Rudy, J.W. (2007). The hippocampal indexing theory and episodic memory: updating the index. Hippocampus *17*, 1158-1169.
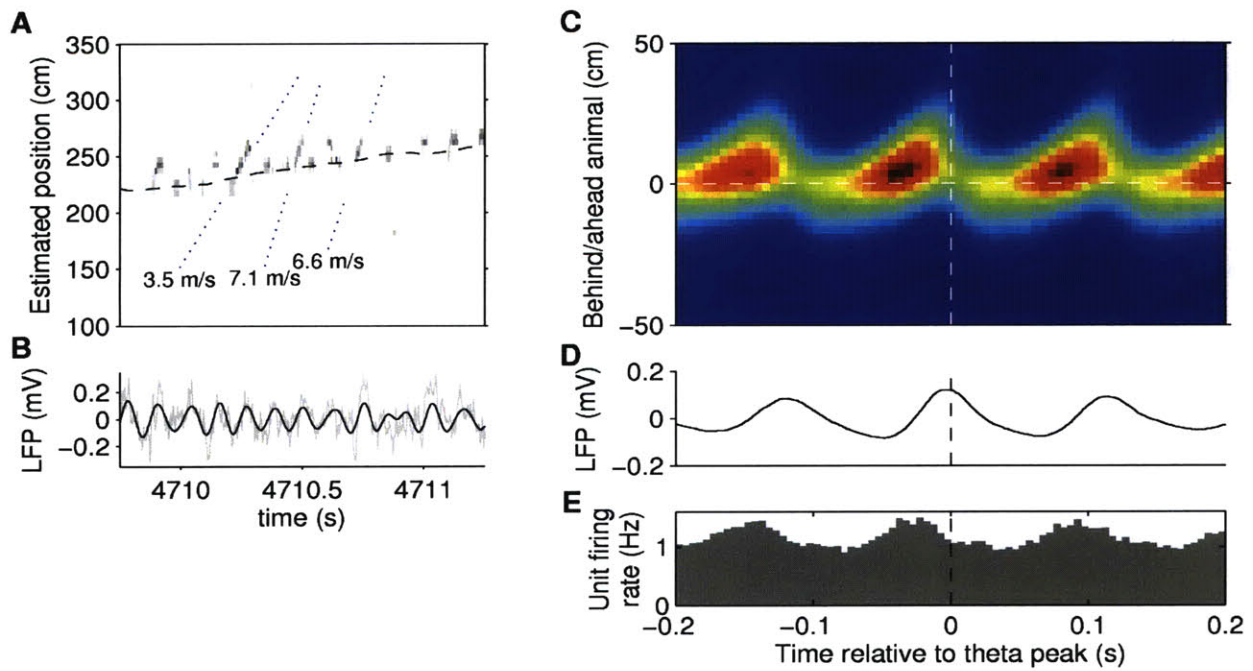
Figure 4.1: Spatial structure of theta sequences

Theta sequences are present in fine timescale position estimates (15 ms time bins, 5 ms steps) calculated during RUN. **A,** Short sequences proceed ahead of the rat's true position (dashed black line). Selected slopes (blue dotted lines) are computed from 62.5 ms around zero-crossings of the theta oscillation. **B,** Raw (gray) and theta-filtered (black; 6-12 Hz bandpass) hippocampal LFP. **C-E,** Average theta sequence across all RUN periods. Theta-triggered averages of: rat-aligned position estimate (**C**); raw LFP (**D**), and place cell firing rate (**E**).