Hexagonal Grid Fields Optimally Encode Transitions in Spatiotemporal Sequences

Nicolai Waniek
nicolai.waniek@tum.de
Neuroscientific System Theory, Technical University of Munich,
80333 Munich, Germany

Grid cells of the rodent entorhinal cortex are essential for spatial navigation. Although their function is commonly believed to be either path integration or localization, the origin or purpose of their hexagonal firing fields remains disputed. Here they are proposed to arise as an optimal encoding of transitions in sequences. First, storage requirements for transitions in general episodic sequences are examined using propositional logic and graph theory. Subsequently, transitions in complete metric spaces are considered under the assumption of an ideal sampling of an input space. It is shown that memory capacity of neurons that have to encode multiple feasible spatial transitions is maximized by a hexagonal pattern. Grid cells are proposed to encode spatial transitions in spatiotemporal sequences, with the entorhinal-hippocampal loop forming a multitransition system.

1 Introduction

Decades of research unearthed neurons that represent spatial information. For instance, place cells (PCs) encode mostly singular locations (O'Keefe & Dostrovsky, 1971; O'Keefe, 1979), head direction cells (HDs) show preferential tuning toward head directions (Chen, Lin, Green, Barnes, & McNaughton, 1994; Ranck, 1984; Preston-Ferrer, Coletta, Frey, & Burgalossi, 2016), and grid cells (GCs) fire at hexagonally arranged locations of an environment (Hafting, Fyhn, Molden, Moser, & Moser, 2005). Together, they are thought to form a cognitive map (Fyhn, Solstad, & Hafting, 2008; Moser, Kropff, & Moser, 2008), anticipated by Edward Tolman as early as 1948 (Tolman, 1948).

GCs, stellate cells of the entorhinal cortex, are believed to convey critical metric information during spatial navigation (Fyhn et al., 2008; Moser & Moser, 2008). It was discovered recently that their fields of activity with respect to the environment, called grid fields, are not only hexagonally distributed but also that the sizes of grid fields vary in discrete steps (Stensola, Stensola, Froland, Moser, & Moser, 2012). GCs are typically characterized by their firing field sizes, orientation, and shift of the hexagonal pattern

relative to an arbitrary coordinate system. Cells that share the same orientation and field sizes are denoted as belonging to the same grid module. Close to the peak of the hierarchical organization of the cortex, they are considered to be an ideal vehicle to understand abstract cortical representations (Moser et al., 2014). However, the origin and purpose of the hexagonal fields and discrete scales are still insufficiently understood and controversial. While several models propose recurrent dynamics as the origin for the arrangement (Fuhs & Touretzky, 2006; Burak & Fiete, 2009; Couey et al., 2013), others use integration of oscillations to form hexagonal fields (Burgess, Barry, & O'Keefe, 2007). Yet others suggest that the fields form due to spatially modulated afferents and as a result of a self-organization process (Gorchetchnikov & Grossberg, 2007; Kropff & Treves, 2008; Stepanyuk, 2015). Extended overviews of such models can be found in Giocomo, Moser, and Moser (2011); Zilli (2012) and Shipston-Sharman, Solanka, and Nolan (2016).

Most researchers assume that GCs perform one of two functions, although both have subtle, yet significant, issues. First, their hexagonal pattern was reported to be suitable for path integration (Burak & Fiete, 2009) and even provide an error-correction mechanism (Sreenivasan & Fiete, 2011). Real-world experiments showed that such models quickly accumulate noise and require external resetting (Mulas, Waniek, & Conradt, 2016), though. Second, theoretical studies showed that GCs grossly outperform PCs during localization when the location is decoded using Bayesian inference (Mathis, Herz, & Stemmler, 2012; Stemmler, Mathis, & Herz, 2015). However, PCs are known to play an essential role during localization and navigation (Morris, Garrud, Rawlins, & O'Keefe, 1982). It remains unclear why there should be two subsystems for localization—GCs and PCs—especially given that neural networks are energetically expensive (Niven & Laughlin, 2008). In either of the two cases, researchers are in disagreement about how downstream neurons should resolve the ambiguities that are due to the hexagonally repeating pattern. Many models use integration of multiple scales to form PCs (Solstad, Moser, & Einevoll, 2006) and simply add more scales to resolve said ambiguities. This does not appear to be a reasonable or generalizable solution to the issue, as it merely shifts the problem out of sight. Finally, most researchers neglect temporal aspects of spatial information, although the hippocampal formation (HF) is crucial for episodic memories (Scoville & Milner, 1957).

This letter is the first of a series to address these concerns. The series introduces a novel computational model suggesting that the entorhinal-hippocampal loop optimally stores and retrieves spatiotemporal sequences. In particular, the series suggests that GCs encode multiple spatial transitions of such sequences and investigates facets of the proposal either theoretically or using simulations. A schematic illustration of the proposed model is presented in Figure 1, and a description of its functional

levels is given in section 3. This letter introduces multitransition encoding with mathematical rigor and is concerned with hexagonal grid fields. Specifically, it focuses on the optimality of the hexagonal arrangement, not on learning these fields. Nonetheless, a recurrent network model is proposed to discuss how the theoretical results could be implemented biologically plausibly and to act as an outlook to the remainder of the series. The papers that follow will present a scale-space model that solves a behavioral issue of transition systems ideally when the scale increment between consecutive grid scales is $\sqrt{2}$, introduce and evaluate the biologically plausible model that self-organizes and learns grid modules based on dendritic computation, and use reinforcement learning (RF) to select an ideal trajectory in a behavioral model that is based on scale-space encodings of transitions.

2 Related Work

Long known to be essential for its formation and storage (Scoville & Milner, 1957), the HF was studied extensively in the context of episodic memories (Tulving, 1972; Jarrard, 1993; Buzsaki, 2015), as well as spatial information processing (O'Keefe & Dostrovsky, 1971; Morris et al., 1982; Hafting et al., 2005). Recently several studies have addressed sequence learning or transition encoding. They examined transitions and the stability of sequences during the formation of temporal memories in spiking neural networks (Hayashi & Igarashi, 2009; Hattori & Kobayashi, 2016) and proposed sophisticated models for the acquisition of episodic memories and the interaction of subareas within the HF (Cheng, 2013). Others proposed that spatial transition encoders are particularly useful during path planning operations and presented a biologically plausible model that was evaluated using a robotic platform in real-world scenarios (Cuperlier, Laroque, Gaussier, & Quoy, 2004; Cuperlier, Quoy, Giovannangeli, Gaussier, & Laroque, 2006). Later, this model was extended to select optimal trajectories from a number of candidate solutions using RF (Hirel, Gaussier, Quoy, & Banquet, 2010). However, none of these studies examined the optimality of transition encodings.

While many models use the integration of ego motion for the formation of hexagonal grid fields (Zilli, 2012), some researchers suggest that their origin lies in afferents from spatially modulated inputs. For instance, rate adaption was used successfully for the stable formation of hexagonal grid fields using PC-like presynaptic activity in Euclidean space (Kropff & Treves, 2008). Others used dendritic computation to cover a spatially modulated input space hexagonally and self-organization principles to arrange several GCs coherently to form proper grid modules (Kerdels & Peters, 2013). Unfortunately none of these studies addressed the concerns already mentioned.

3 Episodic Memories and Multitransition Systems

A recent study observed preplay of PCs to nearby target locations while animals were at rest (Pfeiffer & Foster, 2013). Presumably some form of mental travel or route selection, this sparked the following motivating idea. When navigating to a destination, an animal travels along a trajectory of intermediate places. To plan its trajectory, it requires accumulated knowledge about such locations, evidently memorized and processed in PCs (O'Keefe & Dostrovsky, 1971; Moser, Rowland, & Moser, 2015). However, and at least equally important, it needs insight into feasible (spatial) movements between them. In the basic case of minimizing distances toward targets, the animal also requires knowing approximately how far apart the places are. Without any a priori knowledge about this information, the animal needs to sample its surrounding and learn the relationship between places and, later, during recall, infer these distances using the accumulated data. The acquisition of such data may have happened during an exploration phase and should have happened in some optimal, but also general, manner to apply to arbitrary environments.

3.1 Model and Methods. The novel computational model to solve this task is depicted in Figure 1. It proposes that the entorhinal-hippocampal loop forms an interacting hierarchy of computations with the purpose of optimally storing and retrieving spatiotemporal sequences. Inputs (black arrows) from a suitable sensor space (e.g., boundary vector information, bottom row) or other spatially modulated neurons directly project to PCs that learn to represent locations and rewards (second row; receptive fields indicated as blue circles). Furthermore, movements or transitions between locations are memorized and retrieved in two different layers of the hierarchy. One computational layer (top row) records episodic memories of actually performed sequences—for instance, that they can be replayed in order. Another layer (third row) learns the relationship between spatially close locations. While learning temporal adjacency is supposed to require only interactions between PCs and transition encoders (arrows between top and second layers), acquiring knowledge about spatial transitions requires projections from the sensory representation as well as afferents from the PC layer to bind the appropriate spatially close PCs (arrows from bottom row to third and arrows between the second and third rows).

The remainder of this letter examines the logic and memory consumption of transition storage in this model from a mathematical point of view. The axiomatic system it introduces uses symbols to represent spatial locations and transitions to model movements from one symbol to another, both well known in computer science from automata theory, labeled transition systems, or Markov processes. Note, however, that the sensory input space (see the bottom row of Figure 1) is not explicitly modeled. In fact, any location is assumed to induce a unique sensory representation. The
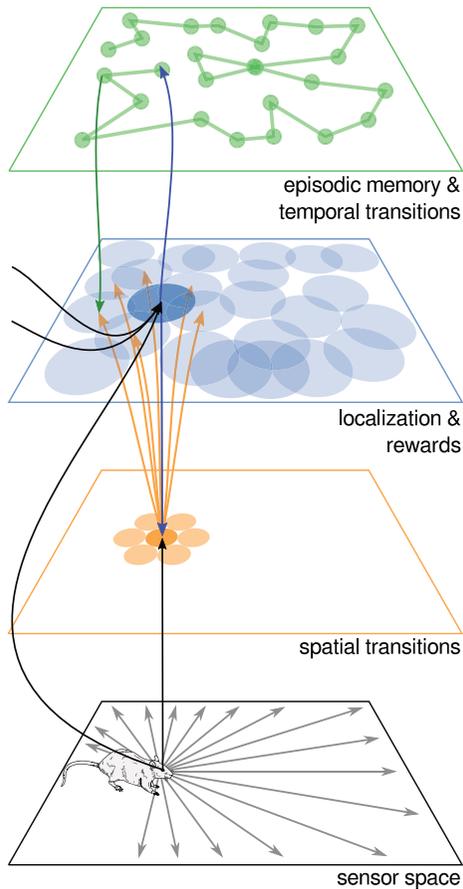
Functional Levels of Computation

Figure 1: Overview of the proposed model and functional levels of computation. PCs (second row) receive projections (black arrows) from a suitable sensory space (bottom row), and other afferents, such as the prefrontal cortex (not modeled). GCs (third row), learn spatial transitions between places based on coactivity of spatial afferents and PCs and recurrent projections to PCs. In addition, episodic memories are stored in a separate computational layer that learns temporal transitions (top row).

consequences as well as the plausibility of these abstractions and the simplification of the input space are discussed in section 4.

The deliberately abstract notion that is used for the analysis that follows allows treating goal-directed navigation as a particular instance of a general

memorization task: formation of episodic memories. The results presumably apply beyond spatial navigation or the entorhinal-hippocampal loop. Furthermore, it enables reasoning about the algorithmic level of the computation independent of the physical realization. Throughout this letter, several brief examples and discussions facilitate understanding the notation and logical analysis, both of them inspired by communicating sequential processes (CSP; Hoare, 1978), the analysis of time in distributed computing systems (Lamport, 1978), and the analysis of causality in theoretical computer science (Halpern, 2015).

3.2 Symbols, Alphabets, and Sequences. Consider an animal that moves across three rooms. The trajectory of the animal can be described by the sequence of symbols $A, B, C \ldots$, each representing one room. The meaning of a symbol is not predetermined, though; rather, it depends on the system being analyzed. For instance, symbols could also represent the event of perception of each corresponding room, particular views of a room or objects within a room, or other modalities as long as they are distinguishable. Moreover, symbols can describe various other forms of sequences, such as the production of a particular series of sounds or steps to perform a certain action. Although the examples in this letter use spatial navigation, the formal system generalizes to other applications.

The entirety of symbols forms an alphabet and their consecutive ordering a sequence, both captured in the following:

Definition 1 (*alphabet and sequence*). *An alphabet Σ is a finite set of symbols. A sequence (or word) is an ordered tuple of symbols $\sigma_i \in \Sigma$, that is, $(\sigma_0, \sigma_1, \ldots) = w \in \Sigma^+$, where $^+$ is the Kleene plus operator.*

Thereby a trajectory of an animal is described by a sequence of symbols, just as motivated in the introductory example. However, immediate repetition of a single symbol is disallowed, formally specified as follows:

Axiom 1 (*nonstationarity*). *A sequence is locally nonstationary if any two successive symbols σ_i and σ_{i+1} are distinguishable, that is, $\sigma_i \neq \sigma_{i+1}$.*

This constraint is inspired by neural dynamics. Specifically, the refractory period of neurons prohibits a continuous representation of a state by a single neuron during short timescales. Also, it is behaviorally relevant for the generation of a sequence. Consider an animal that tries to reach a certain goal location under the pressure of a nearby predator; it needs to find a sequence where symbols correspond to locations. If the animal were to recall a sequence that contains repetitions of locations, it would likely come to a halt at a repeated symbol and fall prey to the predator.

However, the axiom does not limit general capabilities. Two consecutive but distinct symbols of a sequence can have the same associated meaning, depending on the sequence that needs to be encoded. For instance, the perception of a certain room or location within a room in the case of spatial

navigation or a certain actuator state in the case of motor commands could be encoded in two consecutive symbols. Generally the associated meaning of a symbol is independent of the symbol itself.

Moreover, this constraint does not introduce explicit information about real time or prevent repetition of a symbol within a sequence at a later point. In fact, the axiom requires only that any two consecutive symbols are different. Consider an animal that stays in a room for a longer period of time and records its movement in terms of distinguishable locations. Without any additional information or an extension to explicitly incorporate real time, the recorded sequence contains only symbols of consecutive places that the animal's perceptual system can differentiate, regardless of when the change happened.

The directional ordering of a sequence is expressed using the arrow notation $\rightarrow$, which ignores time. In fact, the temporal order of evaluation needs to be stated explicitly, as will be shown further below. Thereby, symbols and transitions form propositions. Consider the example $A \rightarrow B$, which means that the symbol B causally follows after symbol A—in other words, if A is true, it follows that B is also true. Conversely, if A is false, so is B. Hence they form a chain of causality. In addition to $\rightarrow$, the arrow $\rightsquigarrow$ exists; for example, $A \rightsquigarrow C$ means that there exists a path from A to C that bridges $n \geq 0$ intermediate symbols. The negations of the notation are $\nrightarrow$ and $\not\rightsquigarrow$. Finding a path to a target requires the following constraint, though:

Axiom 2 (coherency). *Let $w = \sigma_i, i \in \{0, \ldots, N\}$ be a sequence of N symbols. w is coherent if and only if $\sigma_i \rightarrow \sigma_{i+1}, \forall i \leq N - 1$.*

Coherency is necessary for goal-direction navigation. Consider an animal that intends to travel to a remote target. In terms of the basic idea for this letter, it has to plan a trajectory without any significant gaps. Otherwise, it will get stuck or lost, and it may express undefined behavior or displacement activity. When the animal is not navigating to a specific goal, it is assumed that novel symbols are acquired for future planning operations by explorative movements. The animal's task is thus either to find a valid sequence to its target or acquire more knowledge.

This is not to be confused with definitions of automata in computer science. As defined above, a coherent sequence has symbols that deterministically follow one after the other. However, multiple sequences between two symbols may exist at the same time or different sequences may be generated at different times. It is therefore possible to specify, for instance, a nondeterministic automaton or Markov process that accepts or generates a coherent sequence, respectively.

Definition 2 (validity). *A sequence w is valid or acceptable if it is both nonstationary and coherent.*

Using these notations and axioms, goal-directed navigation from a start A to a goal C is a program that expands the path $A \rightsquigarrow C$ into any valid

sequence $A \to \sigma_1 \to \cdots \to \sigma_{N-1} \to C$, if it exists. The next sections describe how to expand a path, give several examples, and address the memory requirements for storing transitions.

3.3 On Universal Multitransition Systems. The arrow notation specifies relations between two symbols. Consider the transition $A \to B$, which can also be written as the tuple (A, B). The concept of transitions is well known—for instance, from reinforcement learning (RL; Sutton & Barto, 1998 or computer science (Van Benthem & Bergstra, 1994; Thomas, 2006). There, it is usually denoted as a transition function mapping one state to another given a set of actions R—$\tau : \Sigma \times R \to \Sigma$ (Sutton & Barto, 1998). This concept will now be extended to allow simultaneous encoding of multiple feasible transitions. The motivation for the extension will be stated further below.

Definition 3 *(transition system, set, bundle, and point). A multitransition system (MTS) $\mathcal{M}$ is the pair*

$$\mathcal{M} = (\mathcal{P}(\Sigma), \Pi) \tag{3.1}$$

where $\mathcal{P}(\Sigma)$ is the power set of Σ. Set $\Omega \in \mathcal{P}(\Sigma)$ is called a configuration of $\mathcal{M}$. All symbols $\sigma_i \in \Omega$ are considered to be true.

Set Π is called a transition set and contains sets π_i, called transition bundles. In turn, a transition bundle π_i is a set of transitions $\tau_i^k : \Sigma \to \Sigma$, called kth transition point of π_i.

Indices will be dropped if they are clear from the context. In addition, the following terminology and notation will be used:

1. A transition τ from $A \in \Sigma$ to $B \in \Sigma$ can be written (A, B) or $(A \to B)$.
2. $\tau = (A \to B)$ is *defined* for A and *leads to* B, written $A \prec \tau$ and $\tau \succ B$, respectively. The notation is transitive to bundles and sets: $A \prec \pi \Leftrightarrow \exists \tau \in \pi, A \prec \tau$, and $\pi \succ B \Leftrightarrow \exists \tau \in \pi, \tau \succ B$, respectively.
3. A bundle π forms a tuple (S, T) with start and target symbols $S = \{\sigma | \sigma \prec \tau, \tau \in \pi\}$ and $T = \{\sigma | \tau \succ \sigma, \tau \in \pi\}$, respectively.
4. If a transition bundle π_i is true, then so are all contained transitions $\tau_i^k \in \pi_i$.

Transitions form propositional terms that are independent of symbols. Consider the symbol A and the transition $(A \to B)$—both of them propositions—in the expression $A \wedge (A \to B)$. If A is true, it can be deduced logically that B is also true, written $A \wedge (A \to B) \Rightarrow B$. Here, $\wedge$ is the *logical and* operator, and A forms a precondition for the transition $(A \to B)$. As A is true, the precondition is met and so the transition is also true. Hence, B is the conclusion of the entire term.

Since order of evaluation is not specified during logical deduction, sequential evaluation of transitions is made explicit as follows:

Definition 4 (*transition evaluation*). *A configuration* $\Omega \in \mathcal{P}(\Sigma)$ *of an MTS* $\mathcal{M}$ *is evaluated according to the functions*

$$F_{\mathcal{M}} : \Omega, \Pi \mapsto \cup_i f_{\mathcal{M}}(\Omega, \pi_i \in \Pi), \tag{3.2}$$

$$f_{\mathcal{M}} : \Omega, \pi \mapsto \{\sigma_l | \sigma_k \in \Omega, \sigma_k \prec \pi, \sigma_l \text{ is true in } \pi\}. \tag{3.3}$$

The following section gives examples for transition evaluation using these definitions and discusses how they can be implemented in principle in a neural network.

3.4 Interpretations and Implementations of Universal MTS. An MTS $\mathcal{M}$ can be interpreted as a state machine that allows multiple states to be coactive, because the union in equation 3.2 goes over all transition bundles contained in Π. Hence, evaluation of a configuration yields a set of symbols that are true given all symbols of that configuration. Moreover, the function $F_{\mathcal{M}}(\Omega, \Pi)$ allows recursive usage for evaluation until a target symbol is reached. Consider the following example with four symbols, A, B, C, D, where A is the start and C the target. The initial configuration of $\mathcal{M}$ is $\Omega = \{A\}$. Furthermore, the following transition set, bundles, and points are defined:

$$\Pi = \{\pi_0 = \{\tau_0^0, \tau_0^1\}, \pi_1 = \{\tau_1^0, \tau_1^1\}\}, \tag{3.4}$$

$$\tau_0^0 = (A, B) \qquad \tau_0^1 = (A, D), \tag{3.5}$$

$$\tau_1^0 = (B, C) \qquad \tau_1^1 = (D, C). \tag{3.6}$$

It requires two recursive evaluations until the target is found: $C \in F_{\mathcal{M}}(F_{\mathcal{M}}(\Omega, \Pi), \Pi)$. This is similar to the evaluation of transition functions in RL (Sutton & Barto, 1998). However, the set notation allows the superposition of several symbols at the same time.

A visual depiction of the parallel evaluation and superposition of states is presented in Figure 2. The figure shows a different, more complex example and uses a directed graph to depict symbols and transitions. It shows symbols as circular nodes, and transition bundles with a square shape. In the figure, each symbol is associated with precisely one transition bundle, which in turn leads to other symbols. Each such pair is highlighted with a dashed box, and the reason for their pairwise occurrence will become evident in the next section. Given a starting symbol s, indicated as a black node in the left-most panel of Figure 2, the transition evaluation as defined in equation 3.2 allows multiple symbols to consecutively activate. This process repeats recursively until the target symbol t is in the set of active symbols. As illustrated in Figure 2, this yields a propagating wave of active symbols toward the target symbol.

The example presented in Figure 2 resembles the parallel execution of a breadth-first search in a directed graph and exposes a relationship to
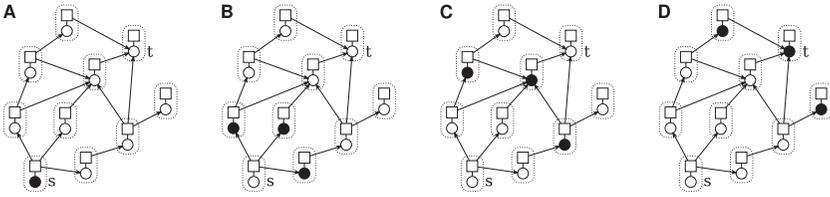
Figure 2: Recursive transition evaluation in a graph. Each symbol (circular node) is associated with a transition bundle (square node), from which other symbols can be reached. Given an initial symbol s (A), the transition evaluation as defined in equation 3.2 activates all subsequent symbols simultaneously (B). Repeatedly applied (C), this will ultimately activate the target symbol t (D) if and only if there is a path from s to t. Essentially, this is a partial parallel breadth-first search in a directed graph without the selection of a shortest path.

message-passing algorithms such as belief propagation in factor graphs. Note that a specific sequence is neither selected in the figure, nor is such a technique presented, as this would require some form of reward mechanism, which is beyond the scope of this letter. In principle, however, this could be implemented with well-known algorithms such as Dijkstra's algorithm or the Bellman-Ford algorithm. The latter was used in Hirel et al. (2010), who also proposed that the HF stores transitions and presented a biologically plausible implementation thereof.

The visual example of Figure 2 helps to derive possible implementations of transition encoding and evaluation, as each node can be interpreted as an individual (artificial) neuron. Such an implementation requires at least the following two modules: one associative memory M_Σ, which stores symbols, and another memory, M_Π, which acquires transitions between these symbols, depicted in the left-most panel of Figure 3. Due to its task of activating a symbol given excitatory afferents that are possibly corrupted by noise and to maintain the activity of the activated symbol until a transition occurs, a likely and biologically plausible implementation of M_Σ is a neural autoassociative memory with winner-take-all (WTA) dynamics (Palm, 1980; Palm, Schwenker, Sommer, & Strey, 1993). In contrast, M_Π is required to associate with consecutive symbols and therefore can be implemented as a neural heteroassociative memory. Recent evidence is in favor of these types of associative memories as existing in the HF (Cutsuridis & Wennekers, 2009; Le Duigou, Simonnet, Telenczuk, Fricker, & Miles, 2014; Mishra, Kim, Guzman, & Jonas, 2016).

The other panels of Figure 3 depict possible wiring diagrams of individual neurons of the two memories M_Σ and M_Π. Note that the diagrams ignore any inhibitory interneurons, assume that each connection is subject to a temporal delay such as axonal transmission, and depict only a few connections to reduce the complexity of the drawings. In each panel,
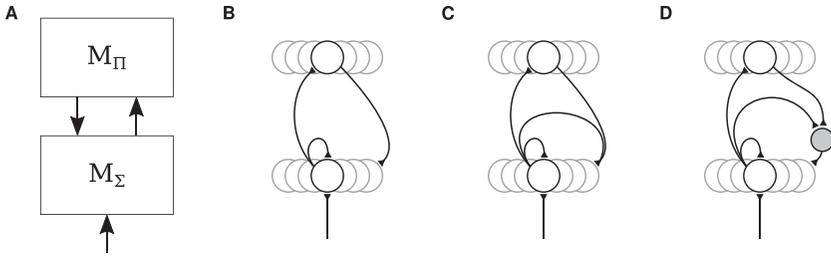
Figure 3: Examples of possible neural implementation schematics for the universal MTS. (A) Memory M_Σ stores symbols, whereas memory M_Π acquires transitions between symbols. They can be implemented biologically plausibly as auto- and heteroassociative memories, respectively. Direct coupling of individual neurons of the two memories (B) appears less likely than indirect coupling via a heterosynaptic connection (C) or interneuron (D), as it does not capture that transition evaluation requires both the preceding symbol as well as the transition to be true for a succeeding symbol to activate.

neurons of M_Σ require local recurrent excitation to form an autoassociative memory.

The schematics differ in the way the recurrent connectivity from M_Π to M_Σ is implemented. Direct coupling, as depicted in panel B, appears to be unlikely as such a system would ignore the start of a transition. More likely candidates are shown in panel C, which uses a heterosynaptic connectivity (indicated as two lines converging on a singular triangular endpoint), or in panel D, which uses an interneuron to integrate the state of the symbol for which a transition is defined as well as the transition. Recall that for a succeeding symbol to activate, both the symbol for which a transition is defined and the transition itself need to be logically true. Hence, the computation performed by the heterosynaptic connectivity or the interneuron is the *logical and*, which was used to evaluate transitions and symbols as propositional terms in section 3.3. In a spiking neural network implementation, this would require the spike times of both symbol and transition neurons to appear in a suitable temporal integration window to either fire the interneuron or activate the next symbol using the heterosynaptic connection. The study of necessary temporal dynamics as well as a concrete implementation of such a network will be left for future work.

In the general case, receptive fields of neurons in M_Σ depend on external excitatory afferents and are subject to the meaning associated with symbols. Neurons in M_Π, however, coactivate with neurons of M_Σ and thus inherit the receptive fields of M_Σ. Consider the example where neurons of M_Σ represent places; that is, they form PCs. Their receptive fields are driven by spatially modulated input and generate distinct place fields. Although

neurons of M_Π do not receive spatially modulated input, they will express place fields due to their coactivation with PCs of M_Σ.

Networks of the form illustrated in Figure 3 have already been reported in the literature, often to model the behavior of synfire chains (Abeles, 1991). A particularly appealing network model thereof was presented in Wennekers and Palm (2009), as it is not only structurally similar to the proposed implementation depicted in panel A of the figure. It was also used to generate syntactic sequences.

One important question that needs to be answered is which of the depicted connectivity schemes appears in the HF. Le Duigou et al. (2014) observed that local recurrent excitation of pyramidal cells in cornu ammonis 3 (CA3) is weak, whereas interneuron excitation appears to be quite effective. Hence, their finding is in favor of the implementation depicted in Figure 3D. Nevertheless, the variant depicted in Figure 3C is also possible, and additional studies are required to further verify if, and if so, which type of implementation is present in local microcircuits of the HF.

It also remains to investigate how many neurons are required to store either symbols or, more important, transitions of a universal MTS. This is the focus of the next section.

3.5 Encoding Capacity of the Universal MTS. The definition of π introduced a bundling trick. Transition bundling provides several benefits when analyzing the computational logic and storage requirements of an MTS, especially in the light of neural encodings. Consider the following thought experiment. Suppose that the generation of a bundle (e.g., a neuron) is energetically expensive; however, the addition of a transition point (e.g., a dendritic spine) to an existing bundle is comparably cheap. To avoid evolutionary pressure (Niven & Laughlin, 2008), the goal is thus to minimize the overall cost. This corresponds to maximizing the number of transition points while minimizing the number of bundles. As will be shown now, it is not possible to merge arbitrary transition points in one bundle without violating the axioms already introduced.

Theorem 1. *Let $\sigma \in \Sigma$, $\mathcal{M}$ an MTS on the alphabet Σ, Π the corresponding transition set, and $\pi = (S, T)$ a transition bundle. $\mathcal{M}$ generates valid sequences if and only if the following conditions hold:*

1. *$\sigma_k \prec \pi \Rightarrow \pi \nsucc \sigma_k$,*
2. *$\pi \succ \sigma_l \Rightarrow \sigma_l \nprec \pi$.*

The sets of symbols S, T for a $\pi = (S, T)$ must be mutually exclusive: $S \cap T = \emptyset$.

Proof. (1) From axiom 1, it follows immediately that any transition π that is defined for σ_k and leads to σ_k violates the nonstationarity condition. (2) Without loss of generality, consider the three symbols $\sigma_0, \sigma_1, \sigma_2 \in \Sigma$ and $\sigma_0 \to \sigma_1 \to \sigma_2$ but $\sigma_0 \nrightarrow \sigma_2$. This yields the transition points $\tau_0 = (\sigma_0, \sigma_1)$ and

$\tau_1 = (\sigma_1, \sigma_2)$. Assume further that τ_0 and τ_1 are bundled in π and that σ_0 and π are true. It follows that $\sigma_0 \wedge \tau_0 \Rightarrow \sigma_1$. However, $\sigma_1 \wedge \tau_1 \Rightarrow \sigma_2$ and thus $\sigma_0 \wedge \pi \Rightarrow \sigma_2$. This contradicts the assumption and violates the coherency constraint.   □

Definition 5 (minimality, universality). *An MTS $\mathcal{M}$ is minimal if there exists only one π_i for any σ_k: $\sigma_k \prec \pi_i \Rightarrow \sigma_k \not\prec \pi_j$ for any $j \neq i$. In a universal $\mathcal{M}$, any arbitrary transition between two symbols σ_k, σ_l is possible.*

Corollary 1. *The input set S_i of a transition bundle π_i is singleton for a minimal and universal $\mathcal{M}$.*

Proof. $\sigma_k \prec \pi_i$ and $\pi_i \succ \sigma_l, \forall l \neq k$. According to theorem 1, $\sigma_l \not\prec \pi_i$, $\forall l \neq k$.   □

Corollary 2. *Let Σ be an alphabet of size M and Π a transition set of N transition bundles $\pi_i = \{S_i, T_i\}$ for a minimal universal $\mathcal{M}$ in which a transition between any two symbols is feasible. Then $M = N$.*

The corollary can be proved by reduction to the graph-coloring problem. In this problem, each node of a graph is assigned a color such that no two neighboring nodes share the same color (Cormen, Leiserson, Rivest, & Stein, 2009). The number of required colors is called the chromatic number of the graph.

Proof. Construct the graph $\mathcal{G}$ of $\mathcal{M}$ in which each transition is represented by a node and any symbol by a directed edge. $\mathcal{G}$ is a complete digraph; that is, each pair of nodes is connected by a pair of directed edges. By merging any such pair of directed edges, $\mathcal{G}$ can be reduced to a simple complete graph.

According to theorem 1, $S_i \cap T_i = \emptyset$ for any π_i. Therefore, only those transitions can be bundled that are not connected by an edge in $\mathcal{G}$. The number of independent nodes in $\mathcal{G}$ is equivalent to the chromatic number of the graph, which is equal to the number of nodes in a complete graph (Cormen et al., 2009).   □

Figure 4B shows an example of a transition graph of a sequence of four symbols after reduction to the simple complete graph. Each edge is annotated with two symbols—the symbols by which the transition can be reached in the complete digraph.

3.6 Multitransition Systems in Euclidean Space. The space that is constructed by symbols δ_i and transitions τ_j above is the discrete topological space with the induced discrete metric. However, the world in which animals reside is not discrete, and, more important, arbitrary jumps between any two locations are infeasible. In particular, the perceived environment is a complete metric space, the Euclidean. For brevity this will simply be called *metric space* from now on. Hence, an MTS $\mathcal{L}$ that encodes transitions in a metric space has different constraints than a universal MTS $\mathcal{M}$ does.
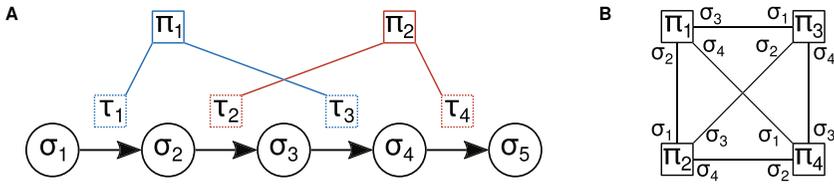
Figure 4: Sequence and transition graph example. (A) A sequence of five consecutive symbols $\sigma_1, \ldots, \sigma_5$. The transitions $\tau_1, \ldots, \tau_4$ can be bundled into π_1 and π_2 without violating any constraints. (B) The reduced transition graph for a fully connected universal MTS is a simple, complete graph, here shown for four transitions and four symbols.

Encoding transitions between locations in a metric space depends on the detection of two consecutive positions. The following analysis is based on the assumption that there exists a continuous signal that depends on and uniquely identifies each possible location of the animal. In terms of the Euclidean space M, this corresponds to locations $\mathbf{x} \in M$. Certainly an animal does not have access to coordinates; however, other stimuli are likely to provide the necessary information. For instance, geometrical information combined with head direction signals is sufficient to represent singular locations, as demonstrated in the boundary vector (BV) cell model (Barry et al., 2006).

According to definition 1, an alphabet is finite. This can be understood to correspond to a finite number of neurons that have to represent locations. However, the alphabet Δ of spatial symbols δ_i has to represent the continuous signals $\mathbf{x}$ of the input space M. Clearly, this corresponds to the well-known sampling theorem.

Definition 6 (*spatial symbol, enablement, and assignment*). *Let $\delta_i \in \Delta$ be spatial symbols according to a sampling process of a complete metric space $\mathcal{D} = (M, d)$. Each δ_i is centered at a $\mathbf{x}_i \in M$.*

A point $\mathbf{p} \in M$ enables δ_i if it is within the support of δ_i given by the open ball $B_{i,s}$ of radius r_s: $B_{i,s} = \{\mathbf{p} \in M | d(\mathbf{x}_i, \mathbf{p}) < r_s\}$.

The point $\mathbf{p}$ is assigned to the closest δ_i: δ_i for which $d(\mathbf{x}_i, \mathbf{p})$ is minimal. Given two adjacent symbols δ_i, δ_j, then $r_w = ||d(\mathbf{x}_i, \mathbf{x}_j)||/2$, describing a ball $B_{i,w}$ of radius r_w.

The definitions of enablement and assignment can be interpreted in the following way. The region in which a spatial symbol is enabled can be understood as its receptive field. In contrast, assignment identifies the closest symbol—for instance, as a result of a WTA mechanism. Due to the definition, spatial symbols are allowed to have overlapping receptive fields. Nevertheless, a single symbol is representative for any location, and transitions can be detected and learned when the winning symbol changes. Before

examining the optimal distribution of spatial transition bundles, it is necessary to determine the placement of symbols.

According to the Petersen-Middleton theorem (Petersen & Middleton, 1962), the ideal sampling strategy for two-dimensional continuous signals and therefore placement of spatial symbols δ_i is a hexagonal arrangement. From a different point of view, the sampling process can be understood as a solution to the problem of packing spheres with diameter r_w as densely as possible. The sphere-packing problem also yields a hexagonal lattice in the two-dimensional case (Conway, Sloane, & Bannai, 1987; Leech & Sloane, 1999).

Given such an ideal sampling process for spatial symbols, the optimal distribution of spatial transition bundles follows immediately.

Theorem 2. *Let $\mathcal{D} = (M, d)$ be a Euclidean space. Let $\mathcal{L} = (\mathcal{P}(\Delta), \Gamma)$ be a minimal transition system on $\mathcal{D}$ such that the countably finite alphabet Δ corresponds to the densest optimal covering with respect to r_w:*

1. *The number of transition bundles $\gamma_i \in \Gamma$ is constant.*
2. *The occurrence of any transition bundle γ_i is periodic.*

The theorem is proved by its corresponding graph-coloring problem, which was introduced above.

Proof. The densest arrangement of spatial symbols according to the Petersen-Middleton theorem is a hexagonal lattice (Petersen & Middleton, 1962). Furthermore, transitions between symbols are possible only between adjacent symbols. Consequently, the corresponding transition graph is not complete; only neighboring transitions are connected. Due to the hexagonal arrangement of symbols, the chromatic number of the resulting graph is 3, and the occurrence of colors is periodic.   □

The following section presents distributions of symbols in environments that are commonly used during rodent experiments and suggests a neural implementation for the dendritic computation.

3.7 An Online Method for Dense Packing of Spatial Symbols. Packing spheres into confined Euclidean spaces is a well-studied problem for which several algorithms exist (see Hifi & M'Hallah, 2009, for an overview). In most cases, however, a global optimization process is used to find the optimal packing. The optimal arrangement of symbols in two-dimensional (open) space according to the Petersen-Middleton theorem is depicted in Figure 5A. In addition, Figure 5B shows one solution of the graph-coloring problem applied to the symbols of Figure 5A. The occurrence of transition bundles follows the arrangement of spatial symbols and is hexagonally distributed as well as periodic in the optimal case.

An animal, however, does not appear to have access to global information or a global optimization procedure a priori. In contrast, a (near) optimal
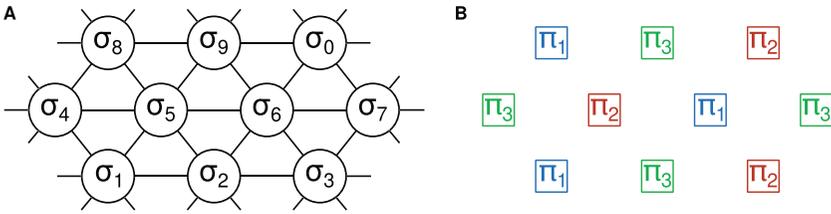
Figure 5: Densest spatial sampling and spatial transition example. (A) The optimal distribution of symbols is a hexagonal arrangement in the two-dimensional case (Petersen & Middleton, 1962). The example shows only undirected edges that represent bidirectional transitions for visibility. (B) The corresponding optimal distribution of transition bundles follows the distribution of symbols. However, bundles can be repeated periodically in a hexagonal fashion. Edges were omitted to improve clarity.

distribution of spatial symbols needs to be established while the animal is actively exploring an environment.

The behavior of spatial symbols was modeled and simulated as an N-body system with a moving agent as follows.[1] Based on definition 6, a spatial symbol was assumed to have circular extents and is centered at its preferred stimulus. Moreover, definition 6 specifies that there is a minimal distance between any two spatial symbols. Consequently, the following local interactions between symbols were used to achieve dense packing. When the simulated agent was at a certain location, the Euclidean distance between the agent and each symbol was evaluated. If there was no symbol within a minimal distance d_{min} to the agent (i.e., there was no sampling center that represented the current location appropriately), a novel symbol was generated and centered at the location. If, however, there was at least one symbol within d_{min}, these symbols interacted in the following manner. The closest symbol was moved toward the current location of the agent, while each symbol that was closer than d_{min} to any other symbol and not the winning symbol itself was pushed away slightly from the other symbol. In addition, symbols that were farther away from d_{min} but within a distance d_{max} to another symbol were attracted to the other symbol. Finally, the new position p_{new} of a symbol was determined by exponentially decaying its old position p_{old} and applying all symbol interactions:

$$p_{force} = p_{old} + \alpha v_{pull} - 5.0 \alpha v_{push}, \tag{3.7}$$

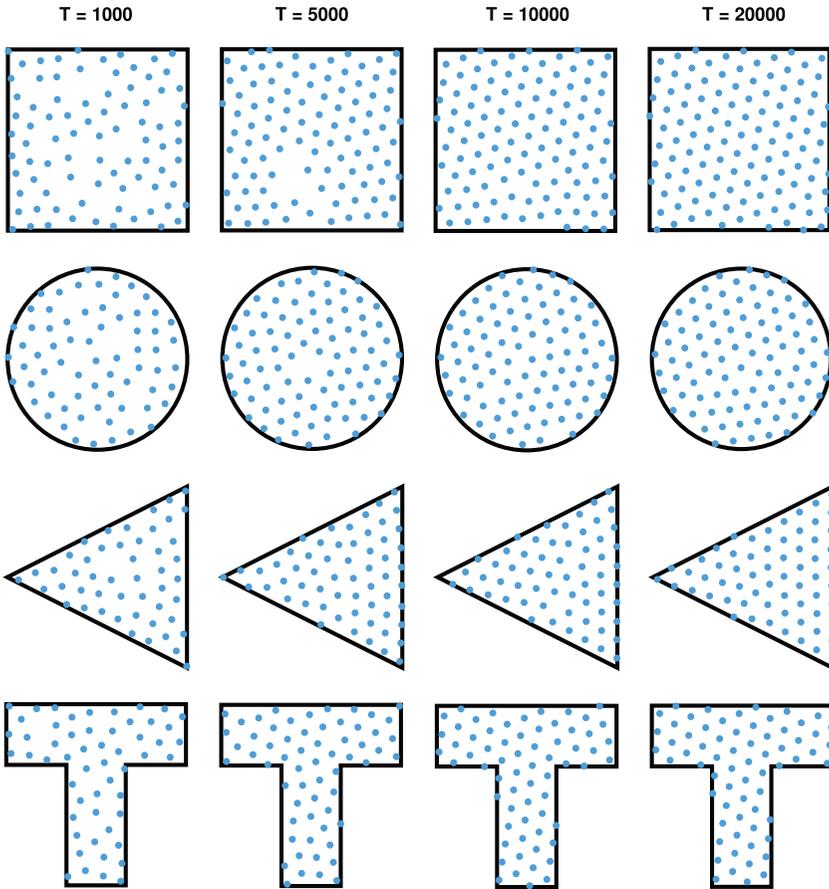$$p_{new} = \beta p_{old} + (1 - \beta) p_{force}, \tag{3.8}$$

Figure 6: Evolution of dense packing of spatial sampling centers in an N-body simulation of interacting symbols. A simulated agent explored several environments and generated novel sampling centers when necessary. The centers interacted in the way that centers that were too close repelled each other, whereas longer distances allowed mutual attraction. Over time (left to right, arbitrary temporal units), the hexagonality of the arrangement increased while fractures started to vanish.

where v_{push} and v_{pull} are the combined forces that repel or attract the symbol, respectively, and $\alpha = 0.02$ and $\beta = 0.8$ arbitrarily chosen.

   Figure 6 shows results of the N-body simulation to distribute sampling centers in several environments at different times after the simulations were started. In all simulations, $d_{\text{min}} = 0.2$, $d_{\text{max}} = 0.4$, and the environments were confined to be within $[-1, 1]^2$. Moreover, the movement statistics of

the agent were similar to those of real rodents: the movement speed was drawn from a gamma distribution, while the movement direction was selected from the current direction in combination with a Laplace distribution. However, details about the statistics were found not to be relevant for the presented results. The displayed temporal unit was chosen arbitrarily and does not reflect any real time. The figure shows that over time, the centers stabilize and increase the hexagonal symmetry and thereby also remove fractures within the arrangement (e.g., see lower part of the square environment at $T = 10,000$ and compare to $T = 20,000$). Note in particular the arrangement of centers in the square environment. The wall offset is reproducible in every simulation and increases the number of symbols packed within the environment. A similar wall-offset behavior was reported for real GCs by Stensola, Stensola, Moser, and Moser (2015), who attributed the effect to an increase of asymmetry with respect to the environment.

4 Discussion

A novel axiomatic system was used to investigate transition encoding in arbitrary and spatially confined sequences. Moreover, possible neural implementations were presented, and simulations showed how spatial symbols optimally arrange in two-dimensional environments.

   Although the model was presented in the context of spatial navigation, the results of both the universal MTS and the spatial MTS are general and possibly apply to other representations. In particular, any system in which transitions in arbitrary spaces need to be encoded suffer from the results obtained for the universal MTS. Systems in which the input can be mapped to a Euclidean space and where transitions should be bundled optimally express behavior like the spatial MTS. Observations of GCs in other brain regions that process sequences are therefore expected.

   In the following, the benefits of hexagonal packings and reasons, as well as implications for a separation of the representation, will be discussed. First, however, the results of the N-body simulation will be used to derive a neural model for a spatial MTS, which will then be integrated with one of the proposed computational models for a universal MTS of Figure 3.

   4.1 Proposed Neural Model of Spatiotemporal Sequence Encoding. Although not a biologically plausible neural network, the results of the N-body simulation can be used to guide the design of a suitable neuron model. Note that the system dynamics of the model that I propose will be described and evaluated in detail only in a following paper in the series. Nevertheless, it is included here to show how a neural network in principle can implement an MTS for spatiotemporal sequences.

   Recall the functional levels of the model presented in Figure 1 and described in section 3, where it was suggested that GCs learn transitions in spatiotemporal sequences and bind the appropriate PCs that are spatially

nearby. In other words, GCs need to learn about spatial relationships given suitable sensory information and convey this information to PCs to which they are recurrently connected. This not only learns spatial transitions but also decouples PCs from explicit information about spatial relationships. This latter point is particularly beneficial during recall and is discussed in section 4.4. To perform this task, GCs are proposed to learn a dense packing of spatial symbols in a suitable sensory space as part of their dendritic computation during exploration of an environment. Specifically, assume that GCs express multiple receptive fields that behave the same as the symbols of the N-body simulation with only a minor extension: the receptive fields represent transition information based on dense packing of symbols.

What do these receptive fields look like, and how can the dendritic computation self-organize appropriately and biologically plausibly within both a single cell and a network of GCs?

The proposed dendritic organization of a single GC is depicted in Figure 7A. It shows that individual branches of the dendritic tree of a GC express their own receptive fields. Moreover, receptive fields need to express certain dynamics to accommodate the requirements of transition encoding. Given a starting location, a transition bundle is not allowed to associate with other symbols in the immediate surrounding. On-center and off-surround dynamics, for example, in the form of a Mexican hat function, appear to be a straightforward solution where the on-center corresponds to the receptive field of the symbol for which the transition is defined and the off-surround region captures any possible symbol in the immediate neighborhood of that symbol. This means that given the sensory representation for a certain location, a GC that expresses a certain grid field needs to dissociate from sensory representations for locations that are in the immediate neighborhood. How can this form without supervision, considering the complex temporal dynamics of neurons? Most neurons, and in particular neurons that represent sensory information, express bell-shaped tuning curves (Jazayeri & Movshon, 2006; Butts & Goldman, 2006). Therefore, their spike times vary relative to their preferred stimulus. Dissociation of GCs from nearby spatial locations can thus be achieved based on relative spike times of GCs and sensory neurons, and suitable spike-timing-dependent plasticity (STDP). In other words, a GC that encodes the start of a transition will have fired before the spikes of neurons that encode for nearby states arrive. Hence, synaptic weights between such states and the GC will be depressed. An idealized illustration of this process is depicted Figure 8A. The figure shows tuning curves of several optimally arranged sensory neurons in the bottom row, possible spike times of the neurons given that the animal is at the preferred location of one sensory neuron. In addition, an inlay shows an asymmetric STDP tuning curve. A gray background indicates the temporal integration window for association.

Local recurrent connectivity with on-center and off-surround regions is common within continuous attractor neural network (CAN) models of
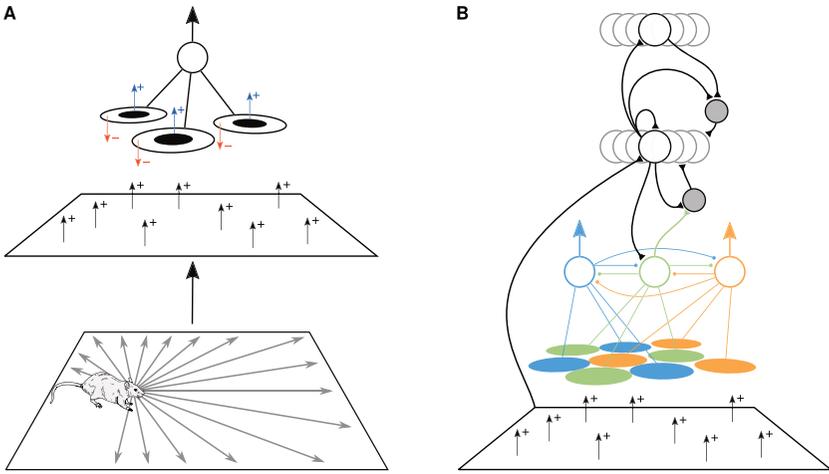
Figure 7: Proposed computational model for a GC and interactions with PCs. (A) The dendritic tree of a GC is formed by organizing on-center and off-surround receptive fields such that they closely pack transitional information based on a suitable sensory space. The on-center regions correspond to symbols for which transition bundles are defined, whereas the off-center regions are target areas where transitions lead to. (B) To minimize the number of GCs (colored circles), they are required to express a WTA mechanism, which can be implemented using exclusively inhibitory recurrent collateral. Moreover, they are recurrently connected to PCs (bottom row of black circles). Episodic memories are stored in episodic transition bundles (top row of black circles), and interneurons (filled gray circles) compute the *logical and* operation that is necessary for transition evaluation.

GCs (Burak & Fiete, 2009; Shipston-Sharman et al., 2016). In particular, a recently published model found that precisely the proposed form of local center-surround interactions leads to stable formation of grid fields (Weber & Sprekeler, 2018). In contrast to these models, however, the purpose of GCs is suggested not to be localization or integration of distances. Rather, a single neuron is suggested to encode as many transitions as efficiently as possible. Hence, a network of GCs assigned this task also needs to adhere to the constraints of MTS.

The local microcircuit of GCs needs to establish WTA dynamics to minimize the number of transition bundles and reduce the ambiguities of transitions. Fast local recurrent inhibition appears to be an ideal solution, as it avoids computational complexities and, in particular, temporal delays that would be the consequence of mechanisms that compare firing rates. Also, it is likely to align the responses of a network of GCs. Local inhibition is
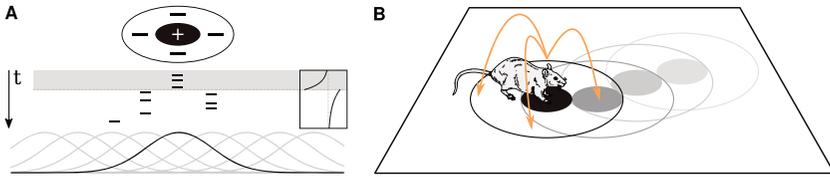
Figure 8: Receptive field formation and transition interpretation. (A) On-center and off-surround receptive fields may occur due to an asymmetric STDP window (small inlay). Sensory neurons (bottom row, one-dimensional bell-shaped tuning curves) that prefer the current location will spike early (middle row, shaded area of vertical spike plot) and before a postsynaptic spike, whereas neurons for nearby locations will arrive only after (middle row, white area of spike plot). The expected results are on-center and off-surround receptive fields (top row, illustrated for a two-dimensional receptive field). (B) Circular receptive fields for spatial transition bundles inform about the target area to which an animal can move with constant cost (orange arrows). Consecutive transitions can be used to form a sequence (gray receptive fields).

well supported by the findings of Couey et al. (2013), who found that the predominant interaction within the entorhinal cortex (EC) is via inhibitory interneurons.

Finally, GCs are suggested to interact with PCs similar to the way they do with episodic transition neurons. The proposed model with all local interactions is presented in Figure 7B, which uses interneurons to compute the *logical and* operations that are necessary during evaluation of transitions.

4.2 Requirements and Predictions for Neurons and Dendritic Trees. Following corollary 2, an implementation of a universal minimal $\mathcal{M}$ requires as many entities to store transition bundles as it has symbols. Due to the dependence on their associated symbols during learning or expansion of a path, transition neurons inherently coactivate with their symbols. This is also visible from the proposed neural model depicted in Figure 3. While symbols receive external afferents, transition encoders only receive feedforward drive from and recurrently project back to symbols. Hence, this could provide an explanation for the two types of PCs found in regions cornu ammonis 1 (CA1) and CA3: one population of PCs acts as spatial symbols, whereas the other encodes temporal transitions. A qualitative change to one population is therefore immediately reflected in the other. This could potentially lead to novel insights into PC remapping (Colgin et al., 2008; Solstad, Yousif, & Sejnowski, 2014), or when and how it is induced. In addition, the difference between place remapping and grid realignment (Fyhn, Hafting, Treves, Moser, & Moser, 2007) is expected to be a result of their independent input sampling processes and the suggested abstraction layer that GCs

provide. Due to the proposed separation of temporal and spatial transitions, place remapping, and thus recall of a different set of temporal sequences, can be performed independent of remapping spatial transitions.

To decorrelate from their target symbols (i.e., to fulfill theorem 1), transition neurons require individual receptive fields per branch and complex dendritic computations. Several recent studies believe that neurons can express this form of sophistication. For instance, dendritic spines were found to express individual structural plasticity (Bosch & Hayashi, 2012), as well as local synaptic plasticity (Segal, 2005; Cichon & Gan, 2015; Weber et al., 2016). Furthermore, dendrites were found to be capable of encoding multiple sensory stimuli (Varga, Jia, Sakmann, & Konnerth, 2011). In combination with the complex intrinsic organization of the EC (Canto, Wouterlood, & Witter, 2008; Witter, Doan, Jacobsen, Nilssen, & Ohara, 2017), it appears likely that GCs perform multiple distinct computations in and self-organization of their dendritic tree corresponding to the proposed multiple receptive fields for transition encoding. So far, it is unclear if the ideal sampling that is the basis of the proposed self-organization should also be reflected in the activity of PCs. If this were the case, PCs should express peak activities that are distributed hexagonally in the absence of other spatial cues.

The model assumes that sensory afferents from BV and HD cells lead to unique sensory representations from which spatial symbols were sampled. Therefore, these cell types directly influence the formation of symbols. In fact, sampling symbols from a sensory space that is spanned by BV and HD cells could explain the findings by Derdikman et al. (2009). They reported that GCs repeat their grid fields in every other corridor of a hairpin maze. The common features in every other corridor are the movement direction in which the animal was running, as well as the HD-dependent geometry of the corridor. Moreover, this could explain the results presented by Krupic, Bauza, Burton, Barry, and O'Keefe (2015) and Krupic, Bauza, Burton, and O'Keefe (2016), who discovered that the geometry of an environment influences the responses of GCs. It is important to note that because the N-body simulation sampled directly from Euclidean coordinates and not from a BV-HD-space, it did not yield any deformations or displacements of fields.

Although the optimality results for MTS were obtained for packing symbols and not entire transition encoders, the results apply to GCs in the following way. The on-center and off-surround receptive fields are ideally circular in spaces that provide unique sensory stimuli due to the Peterson-Middleton theorem. Therefore, their densest packing also follows the sphere packing problem. Furthermore, each part of the proposed neural model (i.e., individual dendrites and the entire neuron) corresponds to entities of MTS (i.e., transitions and bundles). Consequently, the optimality results are believed to be transitive to the proposed neuron model. However, care must be taken with respect to the finite capabilities of a dendritic tree, the sensory afferents, and the discretization of space.

4.3 Discrete spaces, Hexagons, and Sphere Packings. Using symbols in MTS discretizes the input space. Moreover, transitions were assumed to be binary above, that is, indicate only if a transition exists. However, it is conjectured that the obtained optimality results and the Peterson-Middleton theorem still apply even if symbols are not discrete and transitions are associated with a transition probability. Concretely, assume that symbols are represented by bell-shaped tuning curves, similar to optimal representations found in neural networks (Jazayeri & Movshon, 2006; Butts & Goldman, 2006). If these curves are well chosen, the activities of symbol neurons correspond to probabilities. In combination with transition probabilities, this is expected to result in definitions of MTS that are similar, if not identical, to Markov chains or message-passing algorithms. The latter is particularly evident in Figure 2, which shows a bipartite graph of symbols and transitions precisely in the way that factor graphs are commonly depicted and used in probabilistic graphical models (Kschischang, Frey, & Loeliger, 2006). Given symbol and transition probabilities, transition evaluation is thus expected to resemble belief propagation. To conclude this argument, it appears to be feasible to extend the presented discrete MTS to probabilistic MTS.

A notable feature of (discrete) spatial MTS is the minimal number of three transition bundles for continuous spaces. It is important to note that this applies only in the mathematical treatment, where a transition bundle may contain arbitrarily many transition points and symbols perfectly discretize the input space. Furthermore, this would require a perfect WTA mechanism to select the appropriate transition bundle. Instantaneous local inhibition is, however, highly unlikely in a real neural network. It is therefore expected that the amount of overlap between grid fields depends on the temporal delay until local recurrent inhibition acts, which is determined by the synaptic strength between spatial afferents to GCs and the mechanism of inhibition. Moreover, a real neuron is limited in the number of dendrites and synapses, is subject to noisy afferents, and can therefore realistically cover only a fraction of an entire input space. This leads to two important observations.

First, it is expected that the number of GCs found in the rodent EC depends on the dendritic capabilities and the size of their grid fields, and it follows a law that tries to cover the typical habitual space of a rat. Without exact numbers on available synapses and how the recurrent connectivity with PCs is implemented, it is difficult to predict numbers of expected GCs. Nonetheless, an assessment of expected numbers is presented in section 4.4.

Second, continued exploration of an environment is expected to increase the synaptic strength between presynaptic sensory information and GCs. Consequently, grid fields may be fuzzy and their overlap larger in the beginning of an exploration phase due to uncertainties encoded by low synaptic strengths. Given sufficient exploration and suitable plasticity, synaptic strengths will increase. In turn, this should reduce the time to spike of GCs and, consequently, the latency of local recurrent inhibition. It is therefore

predicted that grid fields separate more strongly over time due to faster local inhibition.

An intimately related question to the previous observations is why it is beneficial to encode multiple transitions within a single neuron instead of encoding each transition separately. First, neurons are energetically expensive (Niven & Laughlin, 2008). Minimizing their number appears to be a likely optimization problem that is solved by the brain in an effort to reduce energy consumption. Second, learning transitions requires plasticity suitable for timescales of spatial exploration. Although adult neurogenesis is reported for the hippocampus (Zhao, Deng, & Gage, 2008), only spike-based plasticity rules operate on timescales applicable to spatial navigation and exploration. Consider an animal that explored an environment once and now needs to trace its original trajectory back home. Because the time between exploration and retrieval may be only a few minutes apart, the entire process needs to rely on fast short-term memory. Third, and finally, transition bundles require less physical space. Knowledge of a feasible transition is little more than just a bit of information. If each transition required a separate neuron, the number of required neurons would explode. This is especially true for the universal MTS, where, given N symbols, the system would require N^2 transition neurons. It is also true, however, for a spatial MTS. Ideally, it would require only 6 transition neurons per symbol. Consider, however, the N-body simulation for the square maze of 2 m $\times$ 2 m and symbols of size 20 cm. The converged solution consists of 115 symbols and would require about 690 transition neurons. Extrapolating from these numbers to an area of the size 100 m $\times$ 100 m, which is well below roaming areas of rats (Harper & Rutherford, 2016), this would require 1,725,000 transition neurons. Bundling transitions reduces this number significantly. As mentioned above, the true number of required transition bundles is difficult to assess due to missing numbers for the synapse count.

Another important point to address is why dense hexagonal packing of receptive fields is beneficial for spatial navigation or spatiotemporal sequences in general. Consider the information that is encoded in circular receptive fields for transition encoding. It provides information about constant cost operations to interact with surrounding states. This is depicted in Figure 8B for the case of spatial navigation. The figure depicts that given its current location, a transition informs about the surrounding neighborhood with which the animal can interact directly (i.e., with constant cost because distances are uniform in all directions). This drastically simplifies algorithms that work on such data, as they do not have to consider corner cases or distinguish between certain configurations of symbols.

To make this argument concrete, consider an animal that explored an environment and wishes to find its way back home. Ignoring any other reward signals, the animal wants to plan the shortest trajectory to minimize energy consumption. On a flat surface, a basic task for the animal is thus to approximately compute distances between two arbitrary locations. To be

useful, however, the mechanism needs to work in any environment without a priori knowledge. Because symbols are circular and densely packed, an algorithm that operates on these data can expect that any two neighboring symbols are equidistant in sensory space. Conclusively, there is no mechanism required to distinguish between different neighboring symbols or to align symbols to the surrounding world. The observed alignment and shearing effects of grid fields with respect to the geometry of environments as reported in Stensola et al. (2015), Krupic, Bauza, Burton, Barry, et al. (2015) and Krupic, Bauza, Burton, & O'Keefe (2016) are, in fact, considered to be an artifact of densely packing spherical receptive fields in a suitable sensor space.

Now consider a different animal for which the packing is not hexagonal but, for instance, a square lattice. The fundamental difference is that the lattice loses the constant neighborhood property, and a transition would not inform about an area that is equidistant in any direction. This is because points across a diagonal of the square lattice are farther away than others. Any algorithm now also needs to know how to reach other symbols depending on their position on the lattice, for instance, via vertex or via edge. Again, this is a result of the nonconstant neighborhood. Finally, it is unclear to which external frame of reference the lattice should be aligned, if at all. Essentially, lattices other than the hexagonal introduce avoidable complexities for spherical receptive fields.

Conclusively, dense packing allows learning information about an environment in a general bottom-up fashion without a priori knowledge and without having to deal with corner cases. Later, this information can be used in top-down algorithms. It is also likely that the information that was acquired in this bottom-up fashion is used, for instance, to prune unnecessary information and learn abstractions during a later stage.

One may wonder if there are other representations that could be used to represent space. For instance, previous work showed how PCs can be used to triangulate exact localizations and how the hippocampus forms a topological map (Dabaghian, Memoli, Frank, & Carlsson, 2012; Dabaghian, Brandt, & Frank, 2014). Still, this method requires a process that learns distances or relations of points in the topological map. An alternative representation is to use the fewest neurons to encode the largest volume of space. This, as well as the previous encoding scheme, would require a distinct decoding mechanism to infer the exact location. An efficient implementation of either approach could use rank-order codes similar to the ones proposed for the visual cortex (Thorpe & Gautrais, 1998), that is, the relative time of spikes of PCs informs about the exact location. This decoding mechanism limits the distance between locations that can be represented, though, because temporal integration windows for decoding neurons are not arbitrarily long. Furthermore, it would require knowledge about distances to properly tune the spike times a priori. Another decoding scheme could infer the location based on rate coding. Spatial navigation, however, is an

operation that requires fast execution times—on the order of a fraction of a second—for instance, when finding the shortest path toward a safe location while at the risk of a predator. Encoding transitions explicitly reduces the retrieval time to a bounded factor between neighboring locations (i.e., the axonal transmission time) and yields acceptable times for short and medium distances. In addition, it provides the necessary information about spatial relationships for topological representations without having to deal with the geometry of the environment explicitly. Nevertheless, recursive expansion is an issue for long trajectories and will be addressed in the following paper of the series, which introduces a technique to optimally accelerate retrievals in MTS.

4.4 On the Separation of Spatial and Transitional Information. Sequence and transition learning in the HF was suggested previously (Hayashi & Igarashi, 2009; Hattori & Kobayashi, 2016). However, these studies ignored spatial information or GCs. Furthermore, spatial transitions were the focus of other studies with biologically plausible models of the HF (Cuperlier et al., 2004, 2006; Hirel et al., 2010). Nevertheless, these models did not differentiate between spatial and temporal transition systems, sequences were not defined rigorously, and optimality of transition encoding was not their concern.

Why, though, should there be a distinction between spatial and transitional information in the first place? Several observations are in favor for this. First, numerous failed (and thus unreported) experiments using spiking neuron models indicated that it is difficult to construct a network with only a single neuron population that is capable of both maintaining activity of a representation infinitely but also toggling state transitions arbitrarily. The necessary parameters for a stable network were biologically highly unlikely, especially when neurons were realistically noisy.

Second, the dendritic tree of a PC is certainly not infinite. Thus, the integration of presynaptic afferents is limited by the number of synapses. Given that PCs integrate a plethora of cues, for instance olfactory information (Zhang & Manahan-Vaughan, 2015), receive projections from the entorhinal cortex (Witter, 2007; Witter et al., 2017), and are directly or indirectly coupled with the pre frontal cortex (PFC) (Swanson & Kohler, 1986; Jay & Witter, 1991; Varela, Kumar, Yang, & Wilson, 2014; Ito, 2018), it seems unlikely that there are sufficiently many synapses left to also encode transitions. Certainly the number of synapses is also limited for GCs. In turn, this limitation allows estimating the number of expected GCs depending on the size of their grid fields. Because the number of reported synapses per neuron varies significantly in the literature, the following estimate should be taken with due care. Consider a neuron that has 15,000 synapses, a number on the lower bound of reported synapses for pyramidal neurons in rats (DeFelipe, Alonso-Nanclares, & Arellano, 2002; Markham & Greenough, 2004). Furthermore, assume that not every synapse associates with presynaptic

sensory states, but also, for instance, to recurrent projections from place cells and other neurons in the local microcircuit, and that more than a single synapse is required to drive the neuron to its spiking threshold. In the following estimate, it is therefore assumed that only one-fifth of the total synaptic capacity contributes to learning transitions. Hence, a single neuron can bind to up to 3000 individual presynaptic symbols. Now consider the N-body simulation, which packed 115 symbols with a diameter of 20 cm into the square maze, which extrapolates to 287,500 symbols for an area of size 100 m × 100 m. Despite the number of symbols, this requires only 96 neurons that perform transition bundling, which is several orders of magnitude smaller than the 1,725,000 neurons to store each transition individually. Clearly, the estimate lacks knowledge about other properties of presynaptic neurons, such as bursting behavior, which could trigger a postsynaptic spike with fewer synapses, about how many different states the entirety of presynaptic neurons can represent, or if the reportedly rich dendritic organization of cells in layer II of the EC (Lingenhohl & Finch, 1991) exhibits a significantly larger number of synapses. Still, numbers of expected transition neurons are an the order of a few hundred or, at most, a few thousand for realistic or even large environments, which not only suits the number of pyramidal and stellate cells found in the EC, but also that only a few of these cells express grid-like behavior (Gatome, Slomianka, Lipp, & Amrein, 2010). Moreover, the dependency of the number of neurons on the size of their fields could explain the small number of GCs with large grid fields (Stensola et al., 2012).

Third, the architecture of the HF appears to be a combination of auto- and hetero-associative memories (McNaughton & Morris, 1987; Káli & Dayan, 2000; Papp, Witter, & Treves, 2007; Le Duigou et al., 2014). While the first type can be used to maintain and recall memories even from noise inputs (Palm, 1980), the latter is ideal to store state transitions. In fact, a combination of the two was already used to learn sequences via Hebbian plasticity (Wennekers & Palm, 2009).

Fourth, a separation increases fault tolerance and provides computational benefits. If a transition neuron vanishes, the spatial knowledge is retained and vice versa. Furthermore, both PCs and GCs are suggested to independently acquire their representations due to afferents, which carry spatial information as depicted in Figure 7B. As discussed above, while PCs are thought to learn to identify singular locations for both temporal and spatial purposes and thus directly correspond to both temporal and spatial symbols, GCs are believed to not only associate with their presynaptic spatial afferents. Moreover, they associate with coactive PCs. This has the benefit that GCs learn transitions not only in their spatial input space but also in the symbolic space established by PCs. When operating with reduced presynaptic inputs (e.g., in total darkness), the accumulated errors in the GC representation are expected to be reset using strong presynaptic stimuli that override the afferents from PCs, similar to the realignment model

proposed by Mulas et al. (2016). Due to this proposed coactivity learning, sensory activation in presynaptic neurons is not required during recall and path planning, as it only requires recursive activation of GCs and PCs. Another benefit of separating spatial and transitional representations is that they can vary independently from each other. This form of abstraction layer is widely used in computer science due to its power and known as the bridge or mediator pattern (Gamma, 1995). More important, though, this predicts two distinct modes of operation: learning and recall. During learning, GCs require sufficient drive from PCs as well as afferents from upstream neurons that represent spatial information. During recall, they can rely only on activation from PCs. Conclusively, GCs or the local microcircuit within the EC are expected to expose a mechanism to toggle their mode between either state. I expect these states to be represented heterosynaptically as *logical and* or *logical or* operations, respectively, both of which can be implemented easily in neural networks (Koch, 2004) or via interneurons. Early data suggest that GCs indeed expose at least these two modes of operation (R.G. Morris, private communication, October 6, 2016).

Fifth, and finally, PCs mature earlier during postnatal development than GCs do (Langston et al., 2010; Wills, Cacucci, Burgess, & O'Keefe, 2010). The data suggest that GCs appear as soon as rats start exploring the space around them. This clearly indicates that they perform computations that are relevant only after places can be identified. In particular, temporal sequences, stored in a universal minimal $\mathcal{M}$, are considered to be behaviorally more important than spatial sequences to preweanling rats.

4.5 Current Limitations and Possible Extensions. The introduced MTS restricts the behavioral capabilities of animals. In its current form, an animal equipped with a system similar to the one presented in Figure 7 can exactly recall only previously explored trajectories. The reason is that appropriate places need to be visited consecutively to learn transitions, and any additional transition cannot be acquired without another exploration phase. Furthermore, the system cannot find any shortcuts between places that are beyond the distance of two spatial symbols. For instance, assume that an animal walks on a U-formed trajectory in an environment without any walls, where the start and target locations are at the start and the end of the U, respectively. If the two locations are too far apart, the system is unable to recognize that there is a shortcut between the two locations and merely follows the episodic memory. One solution is to use probabilistic symbols and transitions instead of discretized spaces. Then there may be certain nonzero activity of symbols that are far away, depending on the tails of the probability distribution. This may also lead to the discovery of some shortcuts and more realistic trajectories than the exact succession of discrete symbols. Another solution to this problem is to introduce transition encoders that can recognize places over longer distances. This solution

solves another issue with transition systems—questionable run-times when evaluating trajectories. It, as well as the extension to probabilistic representations, will be presented in detail in the following paper in the series.

Another shortcoming is the assumption of a process that uniquely identifies any location. However, this simplification allowed assessing transition encoding theoretically in an idealized case without considering precise neural dynamics and will be extended to more realistic input spaces in the future. In particular, the model presented in Barry et al. (2006) already showed that BVs cells contain sufficient information to uniquely encode spatial locations. The evaluated environments were mostly of size and shape that are similar to real-world experiments: square or circular with no or only few obstacles. A spatial sampling process that is inherent in dendritic computations of GCs suggests that their fields depend on the uniqueness of the perceived stimuli, though. Recent studies that showed that GC firing strongly depends on the geometry of the environment are clearly in favor of this assumption (Krupic et al., 2015, 2016).

5 Conclusion and Outlook

This letter proposed that GCs optimally encode transitions in Euclidean space. For this purpose, an axiomatic system was introduced to examine the logic and memory consumption of transition encoding in sequences. Furthermore, a novel bundling trick was presented that allows analysis of entities that are capable of encoding multiple transitions at the same time. Finally, the results of the theoretical derivation were discussed in detail, shortcomings of the analysis were addressed, and future work was pointed out. For instance, transition bundling was argued to be performed by dendritic computation and local spatial sampling. In turn, this allowed making predictions and explaining several recent observations of real GCs.

This letter is part of a series that proposes transition coding as the core functionality of GCs. The following work will address multiple scales of transitions and demonstrate why a scale increment of $\sqrt{2}$ is optimal and present a biologically plausible model of dendritic self-organization for transition bundling.

References

Abeles, M. (1991). *Corticonics: Neural circuits of the cerebral cortex*. Cambridge: Cambridge University Press.

Barry, C., Lever, C., Hayman, R., Hartley, T., Burton, S., O'Keefe, J., . . . Burgess, N. (2006). The boundary vector cell model of place cell firing and spatial memory. *Rev. Neurosci.*, *17*(1–2), 71–97.

Bosch, M., & Hayashi, Y. (2012). Structural plasticity of dendritic spines. *Curr. Opin. Neurobiol.*, *22*(3), 383–388.

Burak, Y., & Fiete, I. R. (2009). Accurate path integration in continuous attractor network models of grid cells. *PLoS Computational Biology*, *5*(2), 1–16. doi:10.1371 /journal.pcbi.1000291

Burgess, N., Barry, C., & O'Keefe, J. (2007). An oscillatory interference model of grid cell firing. *Hippocampus*, *17*(9), 801–812.

Butts, D. A., & Goldman, M. S. (2006). Tuning curves, neuronal variability, and sensory coding. *PLoS Biol.*, *4*(4), e92.

Buzsaki, G. (2015). Hippocampal sharp wave-ripple: A cognitive biomarker for episodic memory and planning. *Hippocampus*, *25*(10), 1073–1188.

Canto, C. B., Wouterlood, F. G., & Witter, M. P. (2008). What does the anatomical organization of the entorhinal cortex tell us? *Neural Plasticity*, *2008*, 1–18. doi:10.1155/2008/381243

Chen, L. L., Lin, L.-H., Green, E. J., Barnes, C. A., & McNaughton, B. L. (1994). Head-direction cells in the rat posterior cortex. *Experimental Brain Research*, *101*(1), 8–23. doi:10.1007/BF00243212

Cheng, S. (2013). The crisp theory of hippocampal function in episodic memory. *Frontiers in Neural Circuits*, *7*, 88. doi:10.3389/fncir.2013.00088

Cichon, J., & Gan, W.-B. (2015). Branch-specific dendritic CA2+ spikes cause persistent synaptic plasticity. *Nature*, *520*(7546), 180–185. http://dx.doi.org/10.1038 /nature14251

Colgin, L. L., Moser, E. I., & Moser, M. B. (2008). Understanding memory through hippocampal remapping. *Trends Neurosci.*, *31*(9), 469–477.

Conway, J. H., Sloane, N. J. A., & Bannai, E. (1987). *Sphere-packings, lattices, and groups*. New York: Springer-Verlag.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to algorithms* (3rd ed.). Cambridge, MA: MIT Press.

Couey, J. J., Witoelar, A., Zhang, S. J., Zheng, K., Ye, J., Dunn, B., . . . Witter, M. P. (2013). Recurrent inhibitory circuitry as a mechanism for grid formation. *Nat. Neurosci.*, *16*(3), 318–324.

Cuperlier, N., Laroque, P., Gaussier, P., & Quoy, M. (2004). Planning and navigation strategies using transition cells and neural fields. In *Proc. of the 2004 Conference of Artificial Intelligence and Soft Computing/International Association of Science and Technology for Development*. http://publi-etis.ensea.fr/2004/CLGQ04

Cuperlier, N., Quoy, M., Giovannangeli, C., Gaussier, P., & Laroque, P. (2006). Transition cells for navigation and planning in an unknown environment. In S. Nolfi, G. Baldassarre, R. Calabretta, J. C. T. Hallam, D. Marocco, J.-A. Meyer, . . . D. Parisi (Eds.), *From animals to animats 9: Proceedings of the 9th International Conference*

on Simulation of Adaptive Behavior (pp. 286–297). Berlin: Springer. doi:10.1007/11840541_24

Cutsuridis, V., & Wennekers, T. (2009). Hippocampus, microcircuits and associative memory. *Neural Networks*, *22*(8), 1120–1128. doi:https://doi.org/10.1016/j.neunet.2009.07.009

Dabaghian, Y., Brandt, V. L., & Frank, L. M. (2014). Reconceiving the hippocampal map as a topological template. *eLife*, *3*, e03476. doi:10.7554/eLife.03476

Dabaghian, Y., Memoli, F., Frank, L., & Carlsson, G. (2012). A topological paradigm for hippocampal spatial map formation using persistent homology. *PLoS Comput. Biol.*, *8*(8), e1002581.

DeFelipe, J., Alonso-Nanclares, L., & Arellano, J. I. (2002). Microstructure of the neocortex: Comparative aspects. *J. Neurocytol.*, *31*(3–5), 299–316.

Derdikman, D., Whitlock, J. R., Tsao, A., Fyhn, M., Hafting, T., Moser, M.-B., & Moser, E. I. (2009). Fragmentation of grid cell maps in a multicompartment environment. *Nat. Neurosci.*, *12*(10), 1325–1332. doi:10.1038/nn.2396

Fuhs, M. C., & Touretzky, D. S. (2006). A spin glass model of path integration in rat medial entorhinal cortex. *J. Neurosci.*, *26*(16), 4266–4276.

Fyhn, M., Hafting, T., Treves, A., Moser, M.-B., & Moser, E. I. (2007). Hippocampal remapping and grid realignment in entorhinal cortex. *Nature*, *446*(7132), 190–194. doi:10.1038/nature05601

Fyhn, M., Solstad, T., & Hafting, T. (2008). Entorhinal grid cells and the neural basis of navigation. In S. Mizumori (Ed.), *Hippocampal place fields* (pp. 237–252). New York: Oxford University Press.

Gamma, E. (1995). *Design patterns: Elements of reusable object-oriented software*. Reading, MA: Addison-Wesley.

Gatome, C. W., Slomianka, L., Lipp, H. P., & Amrein, I. (2010). Number estimates of neuronal phenotypes in layer II of the medial entorhinal cortex of rat and mouse. *Neuroscience*, *170*(1), 156–165.

Giocomo, L. M., Moser, M.-B., & Moser, E. I. (2011). Computational models of grid cells. *Neuron*, *71*(4), 589–603. doi:http://dx.doi.org/10.1016/j.neuron.2011.07.023

Gorchetchnikov, A., & Grossberg, S. (2007). Space, time and learning in the hippocampus: How fine spatial and temporal scales are expanded into population codes for behavioral control. *Neural Netw.*, *20*(2), 182–193.

Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature*, *436*(7052), 801–806. doi:10.1038/nature03721

Halpern, J. Y. (2015). *A modification of the Halpern-Pearl definition of causality*. CoRR. abs/1505.00162

Harper, G. A., & Rutherford, M. (2016). Home range and population density of black rats (*Rattus rattus*) on a seabird island: A case for a marine subsidised effect? *New Zealand Journal of Ecology*, *40*(2), 219–228. doi:10.20417/nzjecol.40.25

Hattori, M., & Kobayashi, Y. (2016). A hippocampal model for episodic memory using neurogenesis and asymmetric STDP. In *Proceedings of the 2016 International Joint Conference on Neural Networks* (pp. 5189–5193). Piscataway, NJ: IEEE. doi:10.1109/IJCNN.2016.7727885

Hayashi, H., & Igarashi, J. (2009). LTD windows of the STDP learning rule and synaptic connections having a large transmission delay enable robust sequence learning amid background noise. *Cogn. Neurodyn.*, *3*(2), 119–130.

Hifi, M., & M'Hallah, R. (2009). A literature review on circle and sphere packing problems: Models and methodologies. *Advances in Operations Research*, *2009*, 1–22. doi:10.1155/2009/150624

Hirel, J., Gaussier, P., Quoy, M., & Banquet, J.-P. (2010). Why and how hippocampal transition cells can be used in reinforcement learning, In S. Doncieux, B. Girard, A. Guillot, J. Hallam, J.-A. Meyer, & J.-B. Mouret (Eds.). In *From Animals to Animats 11: Proceedings of the 11th International Conference on Simulation of Adaptive Behavior* (pp. 359–369). Berlin: Springer. doi:10.1007/978-3-642-15193-4_34

Hoare, C. A. R. (1978). Communicating sequential processes. *Commun. ACM*, *21*(8), 666–677. doi:10.1145/359576.359585

Ito, H. T. (2018). Prefrontal-hippocampal interactions for spatial navigation. *Neuroscience Research*, *129*, 2–7. doi:https://doi.org/10.1016/j.neures.2017.04.016

Jarrard, L. E. (1993). On the role of the hippocampus in learning and memory in the rat. *Behav. Neural Biol.*, *60*(1), 9–26.

Jay, T. M., & Witter, M. P. (1991). Distribution of hippocampal CA1 and subicular efferents in the prefrontal cortex of the rat studied by means of anterograde transport of *Phaseolus vulgaris-leucoagglutinin*. *J. Comp. Neurol.*, *313*(4), 574–586.

Jazayeri, M., & Movshon, J. A. (2006). Optimal representation of sensory information by neural populations. *Nat. Neurosci.*, *9*(5), 690–696. doi:10.1038/nm1691

Káli, S., & Dayan, P. (2000). The involvement of recurrent connections in area CA3 in establishing the properties of place fields: A model. *Journal of Neuroscience*, *20*(19), 7463–7477. http://www.jneurosci.org/content/20/19/7463

Kerdels, J., & Peters, G. (2013). A computational model of grid cells based on dendritic self-organized learning. In *Proceedings of the 5th International Joint Conference on Computational Intelligence* (pp. 420–429). Setúbal, Portugal: SciTe Press. doi:10.5220/0004658804200429

Koch, C. (2004). *Biophysics of Computation: Information processing in single neurons*. New York: Oxford University Press.

Kropff, E., & Treves, A. (2008). The emergence of grid cells: Intelligent design or just adaptation. *Hippocampus*, *18*(12), 1256–1269. doi:10.1002/hipo.20520

Krupic, J., Bauza, M., Burton, S., Barry, C., & O'Keefe, J. (2015). Grid cell symmetry is shaped by environmental geometry. *Nature*, *518*(7538), 232–235. http://dx.doi.org/10.1038/nature14153

Krupic, J., Bauza, M., Burton, S., & O'Keefe, J. (2016). Framing the grid: Effect of boundaries on grid cells and navigation. *Journal of Physiology*, *594*(22), 6489–6499. doi:10.1113/JP270607

Kschischang, F. R., Frey, B. J., & Loeliger, H.-A. (2006). Factor graphs and the sum-product algorithm. *IEEE Trans. Inf. Theor.*, *47*(2), 498–519. doi:10.1109/18.910572

Lamport, L. (1978). Time, clocks, and the ordering of events in a distributed system. *Commun. ACM*, *21*(7), 558–565. doi:10.1145/359545.359563

Langston, R. F., Ainge, J. A., Couey, J. J., Canto, C. B., Bjerknes, T. L., Witter, M. P., . . . Moser, M. B. (2010). Development of the spatial representation system in the rat. *Science*, *328*(5985), 1576–1580.

Le Duigou, C., Simonnet, J., Telenczuk, M. T., Fricker, D., & Miles, R. (2014). Recurrent synapses and circuits in the CA3 region of the hippocampus: An associative network. *Front. Cell Neurosci.*, *7*, 262.

Leech, J., & Sloane, N. J. A. (1999). Sphere packing and error-correcting codes. In J. H. Conway & N. J. A. Sloane (Eds.), *Sphere packings, lattices and groups* (pp. 136–156). New York: Springer.

Lingenhohl, K., & Finch, D. M. (1991). Morphological characterization of rat entorhinal neurons in vivo: Soma-dendritic structure and axonal domains. *Exp. Brain Res.*, *84*(1), 57–74.

Markham, J. A., & Greenough, W. T. (2004). Experience-driven brain plasticity: Beyond the synapse. *Neuron Glia Biol.*, *1*(4), 351–363.

Mathis, A., Herz, A. V. M., & Stemmler, M. (2012). Optimal population codes for space: Grid cells outperform place cells. *Neural Computation*, *24*(9), 2280–2317. doi:10.1162/NECO_a_00319

McNaughton, B. L., & Morris, R. G. M. (1987). Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends in Neurosciences*, *10*(10), 408–415. doi:10.1016/0166-2236(87)90011-7

Mishra, R. K., Kim, S., Guzman, S. J., & Jonas, P. (2016). Symmetric spike timing–dependent plasticity at CA3-CA3 synapses optimizes storage and recall in autoassociative networks. *Nature Communications*, *7*. http://dx.doi.org/10.1038/ncomms11552

Morris, R. G., Garrud, P., Rawlins, J. N., & O'Keefe, J. (1982). Place navigation impaired in rats with hippocampal lesions. *Nature*, *297*(5868), 681–683.

Moser, E. I., Kropff, E., & Moser, M. B. (2008). Place cells, grid cells, and the brain's spatial representation system. *Annu. Rev. Neurosci.*, *31*, 69–89.

Moser, E. I., & Moser, M. B. (2008). A metric for space. *Hippocampus*, *18*(12), 1142–1156.

Moser, E. I., Roudi, Y., Witter, M. P., Kentros, C., Bonhoeffer, T., & Moser, M.-B. (2014). Grid cells and cortical representation. *Nat. Rev. Neurosci*, *15*(7), 466–481. http://dx.doi.org/10.1038/nrm3766

Moser, M. B., Rowland, D. C., & Moser, E. I. (2015). Place cells, grid cells, and memory. *Cold Spring Harb. Perspect. Biol.*, *7*(2), a021808.

Mulas, M., Waniek, N., & Conradt, J. (2016). Hebbian plasticity realigns grid cell activity with external sensory cues in continuous attractor models. *Front. Comput. Neurosci.*, *10*, 13. 26924979. doi:10.3389/fncom.2016.00013

Niven, J. E., & Laughlin, S. B. (2008). Energy limitation as a selective pressure on the evolution of sensory systems. *J. Exp. Biol.*, *211*(Pt. 11), 1792–1804.

O'Keefe, J. (1979). A review of the hippocampal place cells. *Prog. Neurobiol.*, *13*(4), 419–439.

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, *34*(1), 171–175. doi:http://dx.doi.org/10.1016/0006-8993(71)90358-1

Palm, G. (1980). On associative memory. *Biological Cybernetics*, *36*(1), 19–31. doi:10.1007/BF00337019

Palm, G., Schwenker, F., Sommer, F., & Strey, A. (1993). Neural associative memories. *Biological Cybernetics*, *36*, 19–31.

Papp, G., Witter, M. P., & Treves, A. (2007). The CA3 network as a memory store for spatial representations. *Learn. Mem.*, *14*(11), 732–744.

Petersen, D. P., & Middleton, D. (1962). Sampling and reconstruction of wave-number-limited functions in n-dimensional Euclidean spaces. *Information and Control*, *5*(4), 279–323. doi:http://dx.doi.org/10.1016/S0019-9958(62)90633-2

Pfeiffer, B. E., & Foster, D. J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. *Nature*, *497*(7447), 74–79.

Preston-Ferrer, P., Coletta, S., Frey, M., & Burgalossi, A. (2016). Anatomical organization of presubicular head-direction circuits. *eLife*, *5*, e14592. doi:10.7554/eLife.14592

Ranck, J. B. (1984). Head-direction cells in the deep cell layers of dorsal presubiculum in freely moving rats. *Society for Neuroscience Abstracts*, *10*.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *J. Neurol. Neurosurg. Psychiatry*, *20*(1), 11–21. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC497229/

Segal, M. (2005). Dendritic spines and long-term plasticity. *Nat. Rev. Neurosci.*, *6*(4), 277–284. doi:10.1038/nrm1649

Shipston-Sharman, O., Solanka, L., & Nolan, M. F. (2016). Continuous attractor network models of grid cell firing based on excitatory-inhibitory interactions. *J. Physiol. (Lond.)*, *594*(22), 6547–6557.

Solstad, T., Moser, E. I., & Einevoll, G. T. (2006). From grid cells to place cells: A mathematical model. *Hippocampus*, *16*(12), 1026–1031.

Solstad, T., Yousif, H. N., & Sejnowski, T. J. (2014). Place cell rate remapping by CA3 recurrent collaterals. *PLoS Computational Biology*, *10*(6), 1–10. doi:10.1371/journal.pcbi.1003648

Sreenivasan, S., & Fiete, I. (2011). Grid cells generate an analog error-correcting code for singularly precise neural computation. *Nat. Neurosci.*, *14*(10), 1330–1337. http://dx.doi.org/10.1038/nm.2901

Stemmler, M., Mathis, A., & Herz, A. V. M. (2015). Connecting multiple spatial scales to decode the population activity of grid cells. *Science Advances*, *1*(11). doi:10.1126/science.1500816

Stensola, H., Stensola, T., Froland, K., Moser, M.-B., & Moser, E. I. (2012). The entorhinal grid map is discretized. *Nature*, *492*(7427), 72–78. doi:10.1038/nature11649

Stensola, T., Stensola, H., Moser, M.-B., & Moser, E. I. (2015). Shearing-induced asymmetry in entorhinal grid cells. *Nature*, *518*(7538), 207–212. http://dx.doi.org/10.1038/nature14151

Stepanyuk, A. (2015). Self-organization of grid fields under supervision of place cells in a neuron model with associative plasticity. *Biologically Inspired Cognitive Architectures*, *13*, 48–62. doi:http://dx.doi.org/10.1016/j.bica.2015.06.006

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

Swanson, L. W., & Kohler, C. (1986). Anatomical evidence for direct projections from the entorhinal area to the entire cortical mantle in the rat. *J. Neurosci.*, *6*(10), 3010–3023.

Thomas, W. (2006). Automata theory and infinite transition systems. In *Preproceedings of EMS Summer School CANT 2006*. Institut de Mathématique, Université de Liège.

Thorpe, S., & Gautrais, J. (1998). Rank order coding. In J. M. Bower (Ed.), *Computational neuroscience: Trends in research*, *1998* (pp. 113–118). Boston: Springer. doi:10.1007/978-1-4615-4831-7_19

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*(4), 189–208. doi:10.1037/h0061626

Tulving, E. (1972). Episodic and semantic memory 1. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 382–404). London: Academic Press.

Van Benthem, J., & Bergstra, J. (1994). Logic of transition systems. *Journal of Logic, Language and Information*, *3*(4), 247–283. doi:10.1007/BF01160018

Varela, C., Kumar, S., Yang, J. Y., & Wilson, M. A. (2014). Anatomical substrates for direct interactions between hippocampus, medial prefrontal cortex, and the thalamic nucleus reuniens. *Brain Struct. Funct.*, *219*(3), 911–929.

Varga, Z., Jia, H., Sakmann, B., & Konnerth, A. (2011). Dendritic coding of multiple sensory inputs in single cortical neurons in vivo. *Proceedings of the National Academy of Sciences*, *108*(37), 15420–15425. doi:10.1073/pnas.1112355108

Weber, J. P., Andrasfalvy, B. K., Polito, M., Mago, A., Ujfalussy, B. B., & Makara, J. K. (2016). Location-dependent synaptic plasticity rules by dendritic spine cooperativity. *Nat. Commun.*, *7*, 11380.

Weber, S. N., & Sprekeler, H. (2018). Learning place cells, grid cells and invariances with excitatory and inhibitory plasticity. *eLife*, *7*, e34560. doi:10.7554/eLife.34560

Wennekers, T., & Palm, G. (2009). Syntactic sequencing in Hebbian cell assemblies. *Cogn. Neurodyn*, *3*(4), 429–441. 9095[PII]. doi:1007/s11571-009-9095-z

Wills, T. J., Cacucci, F., Burgess, N., & O'Keefe, J. (2010). Development of the hippocampal cognitive map in preweanling rats. *Science*, *328*(5985), 1573–1576.

Witter, M. P. (2007). Intrinsic and extrinsic wiring of CA3: Indications for connectional heterogeneity. *Learn. Mem.*, *14*(11), 705–713.

Witter, M. P., Doan, T. P., Jacobsen, B., Nilssen, E. S., & Ohara, S. (2017). Architecture of the entorhinal cortex: A review of entorhinal anatomy in rodents with some comparative notes. *Front. Syst. Neurosci.*, *11*, 46.

Zhang, S., & Manahan-Vaughan, D. (2015). Spatial olfactory learning contributes to place field formation in the hippocampus. *Cerebral Cortex*, *25*(2), 423. doi:10.1093/cercor/bht239

Zhao, C., Deng, W., & Gage, F. H. (2008). Mechanisms and functional implications of adult neurogenesis. *Cell*, *132*(4), 645–660.

Zilli, E. (2012). Models of grid cell spatial firing published 2005–2011. *Frontiers in Neural Circuits*, *6*, 16. doi:10.3389/fncir.2012.00016