Master's Degree in Physics of Complex Systems

Capacity of the Ant Navigation System

Author: Maria Chiara Montorfano

Supervisors: Andrea Pagnani, Rémi Monasson

Academic Year: 2024–2025

Contents

1 Introduction

Insects such as ants, bees, and flies are capable of navigating complex environments, learning from experience, and exhibiting adaptive behaviors, despite having a compact nervous system with orders of magnitude fewer neurons than vertebrates [1]. Understanding the principles underlying these capabilities is a major goal in computational neuroscience and could inform the design of efficient artificial systems. Among the insect brain structures, the mushroom body (MB) stands out as a central learning hub. It is widely recognized for its role in olfactory memory but has also been implicated in higher-order tasks such as visual navigation, context-dependent decision-making, and reinforcement learning.

Initially studied in the context of olfaction in Drosophila melanogaster [2, 3], the MB has since been shown to play a broader role in learning in different settings and tasks. In ants and bees, for example, the MB could be essential for visual navigation [4, 5, 6]. These insects encode panoramic visual scenes while foraging and later use them to return to their nest, relying on familiarity-driven route recognition. While visual navigation in ants has long been a subject of behavioral study, only recently have explicit hypotheses about the underlying neural implementation been put forward, with growing evidence pointing to the MB as a key substrate for visual memory in navigation tasks [7, 8]. Complementary computational models demonstrate that the known architecture of the MB is in principle sufficient to support this function, further strengthening the link between anatomy and behavior. Quantifying the memory capacity is, in general, a notoriously difficult theoretical and experimental problem. However, recent behavioral studies on visually navigating ants provide a rare opportunity to anchor this question in ecological reality, by offering measurable estimates of how far ants can travel using route memory alone.

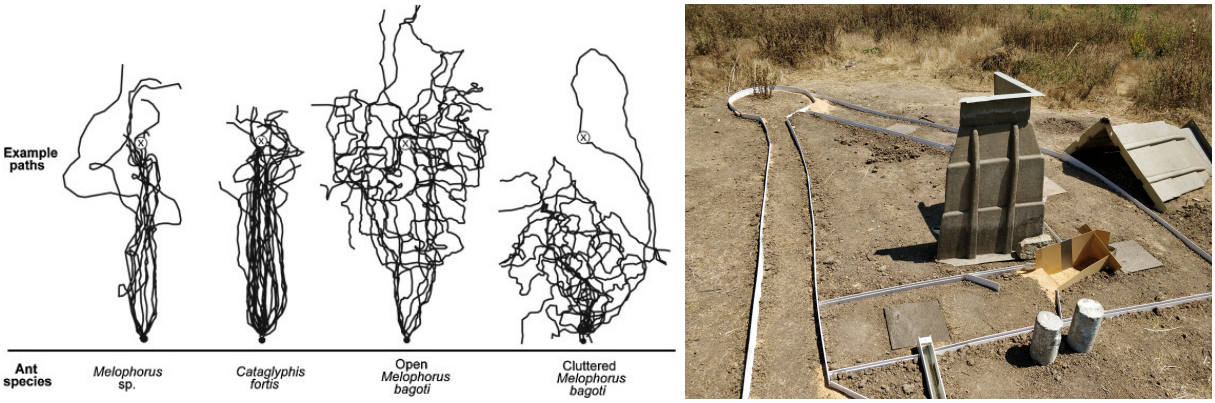

Figure 1: Examples from behavioral studies on ant navigation. On the left examples of paths performed by different ant species from [9]. On the right an experimental setup for ant navigation in a controlled natural environment from `https://antnavigation.wordpress.com/people/`.

The anatomical structure of the MB is highly conserved across insect species and is characterized by a distinctive feedforward organization. Projection Neurons (PNs) relay preprocessed sensory information, such as visual or olfactory input, from earlier brain regions to the mushroom body. This information is conveyed to a large population of Kenyon Cells (KCs). Each KC receives input from a small, random subset of PNs via sparse synaptic connections. Because of their high activation thresholds, only a small proportion of KCs respond to any given input, leading to a sparse and decorrelated representation of sensory stimuli [2]. These KC activation patterns are then integrated by a smaller set of Mushroom Body Output Neurons (MBONs), which ultimately influence behavior. Learning occurs primarily at the synapses between KCs and MBONs and is modulated by neuromodulatory inputs, most notably dopaminergic signals that convey reinforcement [3, 10].

The mushroom body is especially interesting from a computational perspective because it constitutes a biologically inspired example of a neural architecture capable of learning with only three functional layers and synaptic plasticity localized only in the third layer. This simplicity provides an accessible yet

powerful model to study learning and memory in neural systems. This architectural setup supports a number of computationally attractive properties. The sparsity of KC activation maximizes the memory capacity of the network by minimizing interference among stored patterns. The randomness in PN-to-KC connectivity promotes decorrelation, which is crucial for discrimination. Finally, the restriction of plasticity to the final layer simplifies the learning mechanism and reduces the number of degrees of freedom that must be tuned, thus facilitating theoretical modeling.

While these features have made MB-inspired networks popular in computational models, most studies have focused on tasks involving unstructured or uncorrelated input statistics. Natural visual scenes, though, are known to exhibit scale-invariant statistics, including long-range correlations and power-law distributions in their spatial frequency spectra [11, 12]. It is not clear how these nontrivial statistical properties influence the behavior of sensory systems and impact the performance of learning networks like the MB. Furthermore, despite the approachability and interesting features of this biological neural network, a theoretical framework for the functioning of the Mushroom Body is still lacking. There are very few analytical results connecting the statistical structure of sensory inputs to the KC activation, particularly under realistic conditions. Most current understanding is based on numerical simulations, and even those are often restricted to simplified input models.

The goal of this project is to close this gap by constructing a theoretical statistical framework that describes how different input statistics affect the dynamics and functionality of mushroom body inspired circuits. We specifically aim to analyze how spatially and temporally correlated inputs modulate the KC activity, and how this modulation influences learning and memory processes. Particular attention is devoted to identifying the maximum memory length and estimating the storage capacity of the insect-inspired navigation system, in order to understand how efficiently such circuits can retain information.

Our approach combines both simulations and analytical tools from stochastic process theory. In particular, we interpret the activation of KCs as a threshold-crossing event in a stochastic trajectory. Given a high-dimensional input vector processed through sparse, random connections, the activity of a KC can be modeled as a noisy signal approaching a threshold. When inputs evolve over time, either due to movement through a visual environment or temporal structure in sensory input, this becomes a first passage time (FPT) problem: the time at which a fluctuating process first exceeds a threshold.

This connection enables a deeper understanding of how correlation properties in the input affect the KC activity. By comparing these theoretical predictions with simulations on MB-like networks, we can validate the analytical framework and explore how the learned representations at the KC-MBON interface are affected by input statistics.

In summary, this work proposes a principled theoretical approach to understanding how the statistical structure of sensory inputs modulates learning and memory in sparse, plastic neural circuits. Using the insect mushroom body as a model, we combine ideas from computational neuroscience, statistical physics, and stochastic processes to derive predictive insights about neural coding and learning under realistic conditions. Our findings not only aim to advance the understanding of insect intelligence but also to offer inspiration for the design of compact, efficient artificial learning systems [1].

2 Methods

The primary objective of the project was to characterize how different input statistics influence the output dynamics of a sparse binary neural network and to elucidate the mechanisms underlying learning in this architecture. A particular focus was placed on determining the maximum memory capacity of the network under online learning conditions, where synaptic weights are updated in real time as new inputs are presented. To achieve this, we combined analytical approaches with numerical simulations. The analytical results provided theoretical insights into the behavior of the system, while the simulations served to validate and refine these predictions, ensuring consistency between both methodologies.

2.1 The Neural Network

We first present the architecture of the neural network model used in our study, which is inspired by the anatomical and functional organization of the insect Mushroom Body and it is well-established in literature [5, 6]. Sensory information is processed through the network across three different layers. Learning is mediated by synaptic plasticity, the biological mechanism by which the strength of connections between neurons changes in response to activity and experience.

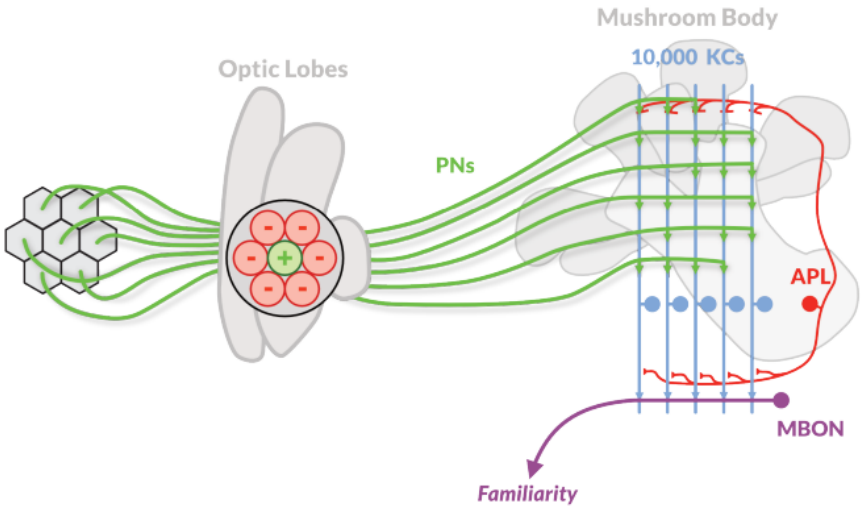

Figure 2: Schematic representation of the Mushroom Body circuit architecture, adapted from [6]. Visual input is collected by ommatidia and processed by the optic lobes (left), which extract relevant features and send the information via Projection Neurons (PNs) to the Kenyon Cells (KCs). Each green line represents a PN forming sparse synaptic connections with a subset of the 10^4 KCs (blue). The APL neuron (red) provides global inhibition, ensuring that only a small fraction of KCs fire in response to a given stimulus. These activated KCs then project to a single MBON (purple), whose output represents the familiarity of the current visual scene.

The three-layered structure of the Mushroom Body involves:

- **Projection Neurons (PNs)** that convey sensory information from sensory organs;
- **Kenyon Cells (KCs)** that form a high-dimensional sparse representation of sensory inputs;
- **Mushroom Body Output Neuron (MBON)** that classify sensory patterns based on learned associations.

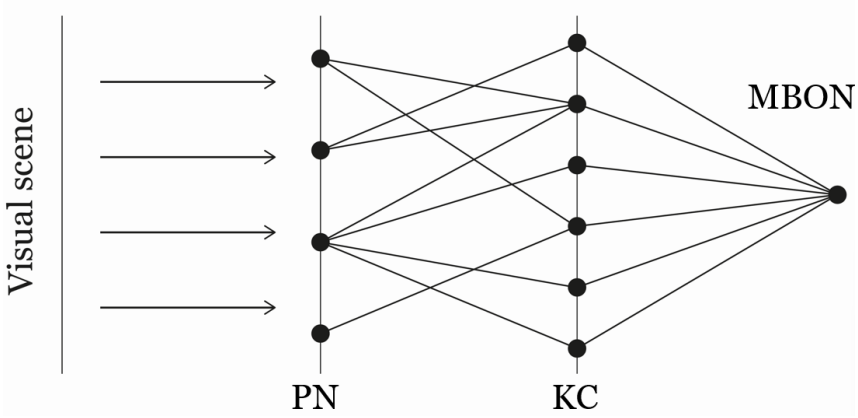

Figure 3: Simplified diagram of the three-layer feedforward structure of the Mushroom Body neural network. Each arrow represents information flow from the left to the right. Visual scenes are encoded as input signals processed by Projection Neurons (PNs), shown as the first vertical layer. Each PN samples a portion of the input and transmits it to a small, random subset of Kenyon Cells (KCs), forming a sparse and distributed code in the second layer. Finally, all KCs project to a single Mushroom Body Output Neuron (MBON).

Each PN receives input from an ommatidium (visual photoreceptor unit), which samples a portion of the panoramic visual scene. The visual signal from the ommatidium i is denoted by S_i and is a function of the ant's position and angular orientation.

A biologically plausible preprocessing step in the network is lateral inhibition [13, 14], a mechanism that enhances spatial contrast and extracts spatial features such as edges from the visual field, known to be critical for navigation [14, 15]. In this operation, each Projection Neuron (PN) receives input from its corresponding photoreceptor, but this signal is reduced by averaging the activity of its neighboring photoreceptors. Mathematically, the response of PN_i is given by:

$$\mathrm{PN}_i = S_i - \frac{1}{|\mathrm{nn}|} \sum_{i' \in \mathrm{nn}(i)} S_{i'} \tag{1}$$

where $\mathrm{nn}(i)$ denotes the set of neighboring photoreceptors, and $|\mathrm{nn}(i)|$ is the number of such neighbors.

Each KC receives input from a small, random subset of PNs. The activity of the KC is determined via a thresholded sum:

$$\mathrm{KC}_j = H\left(\sum_i V_{ji}\,\mathrm{PN}_i - T\right) \tag{2}$$

Here, V_{ji} is a sparse random binary matrix indicating the presence or absence of a synaptic connection between the i^{th} PN and the j^{th} KC; T is a global inhibition threshold regulated by the APL neuron while H is the Heaviside function defined as:

$$H(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x > 0 \end{cases} \tag{3}$$

Biologically, the threshold T is controlled by the APL neuron (Anterior Paired Lateral neuron), which receives excitatory input from all Kenyon Cells and sends back a global inhibitory signal to all of them. This feedback loop ensures that only a small fraction of the most strongly activated KCs fires. This

mechanism supports sparse coding by creating a competitive dynamic among KCs and is crucial for efficient and non-redundant memory storage in the Mushroom Body [6]. Ensuring that only a small fraction of KCs fire for any given input enforces a sparse code that maximizes memory capacity and reduces overlap between representations.

Initially, all KC-to-MBON synapses W_j are set to 1, and the MBON activity is:

$$\text{MBON} = H\left(\sum_j W_j \,\text{KC}_j - T'\right) \tag{4}$$

The plasticity of the KC-to-MBON synapses constitutes the core learning mechanism of the network. When a scene is learned as "familiar", the active KCs at that time have their synapses depressed, so W_j is set to 0, effectively silencing their contribution to the MBON. This synaptic depression is a biologically plausible form of long-term plasticity observed in insects [3, 6].

Thus, when the ant re-encounters a known view, the same KCs fire, but their output is blocked at the MBON, resulting in zero activity. The MBON output therefore encodes a familiarity signal, where a familiar scene corresponds to an output MBON = 0 and a novel scene corresponds to MBON = 1. As more scenes are learned, the network stores memory traces by systematically silencing specific KC-to-MBON pathways, creating a sparse and distributed memory code.

With increasing numbers of learned scenes, more synapses are set to zero, increasing the risk of "false familiarity" where a novel scene produces zero output in the MBON. It is interesting to address at which point of the navigation every novel scene would be perceived as familiar to the neural network. To understand it we want to study the availability of the plastic connections W_j.

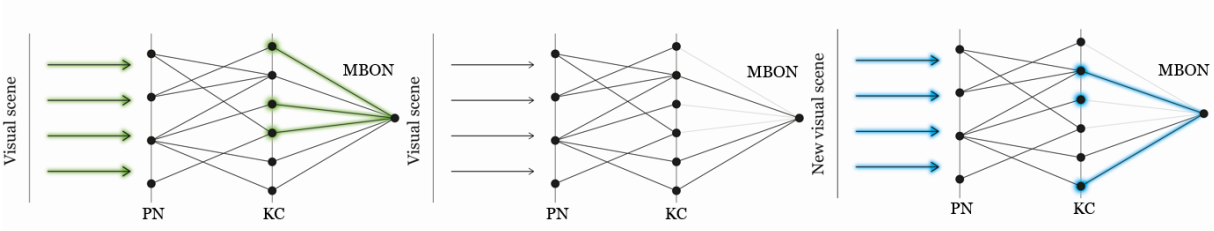

Figure 4: Schematic rendering of the mushroom body learning process through KC-MBON plasticity. Once the first visual scene is presented all the KC-MBON connections are available, all the activated KCs contribute to the MBON final output. Those activated connections are then suppressed for all the navigation-learning. Once a new visual scene is presented to the neural network, it activates another combination of KCs are activated, some may correspond to active KC-MBON connections while other may not. These last ones do not contribute to the final output.

This model is implemented as a sparse binary neural network in PyTorch. Visual scenes are first processed with a lateral inhibition filter that enhances contrast. The result is then projected onto a large population of Kenyon cells via a sparse, random connection matrix, where each element is independently set to 1 with low probability p, mimicking the anatomical PN–KC connectivity. Each KC activates only if its input exceeds a threshold, producing a sparse code. The MBON sums the active KC outputs, weighted by plastic synapses that are depressed (set to zero) when a scene is learned. This simulates familiarity detection through synaptic silencing. Simulations were conducted while varying key parameters, such as network size, connectivity probability and activation threshold to study how these factors affect memory capacity and generalization. Additionally, two other strategies were explored to control KC activity levels: one based on dynamically tuning the activation threshold through an iterative update rule, and another based on selecting only the top K most active KCs in each trial.

2.2 Online learning - neural activations as a function of time

Online learning refers to a learning paradigm in which synaptic updates occur incrementally, in real time, as each new input is presented to the network. This contrasts with batch learning, where updates are made after processing the entire dataset. Online learning is particularly relevant for biological agents like insects in continuous learning tasks such as navigation, where limited neurons and connections must adapt to changing environments. To model this scenario, we explicitly introduce time as a discrete variable and reinterpret the behavior of the neural network as a dynamical process. At each time step t, the agent encounters a visual input corresponding to a specific location and orientation in space. This input is processed by the network, giving rise to a time-dependent cascade of neural activity. We therefore study the responses of the neurons as a function of time t: $\mathbf{PN}(t)$, $\mathbf{KC}(t)$, and MBON(t) where $\mathbf{PN}(t)$, $\mathbf{KC}(t)$ represent respectively the vector of the Projection Neurons and the vector of the Kenyon Cells. This temporal formulation allows us to analyze the evolution of neural activity and synaptic plasticity across successive inputs, laying the groundwork for a theoretical understanding of how familiarity detection emerges over time in response to correlated sensory scenes.

In order to build a theoretical framework for the functioning of Mushroom Body networks, we focus on the mechanisms underlying the Kenyon cells activation. Recalling the mechanism of the KC activation in (2), we can define the quantity:

$$a_j(t) = \sum_i V_{ji}\,\mathrm{PN}_i(t)$$

The process of KC activation and synaptic suppression can be then interpreted as a first passage time (FPT) problem: each KC-to-MBON synapse is suppressed the first time the corresponding $a(t)$ exceeds a critical threshold T. During navigation, due to spatial correlations in the landscape the animal moves in, a visual scene presented to the network is correlated to the ones presented before, inducing temporal correlations in the activation patterns of KCs. So this correlation in space leads to an autocorrelation in time in the quantity $a(t)$. To take into account the autocorrelated nature of the problem, we model $a(t)$ as an autoregressive (AR) process, where $a(t)$ depends linearly on its previous values plus a stochastic innovation term. This framework allows us to analytically describe the evolution of KC activation over time in the presence of correlated inputs, and to compute the probability that $a(t)$ reaches the threshold for the KC activation at a given time. The use of AR processes provides a bridge between the empirical dynamics observed in simulations and the theoretical analysis. This theoretical approach enables us to quantify how structured inputs affect the long-term behavior of the network and its ability to discriminate between familiar and new visual stimuli.

2.2.1 Autocorrelations

In order to compute the autocorrelations $\langle a_j(t) \cdot a_j(t+\tau)\rangle_t$ we write $\mathbf{PN}(t)$ and, as a consequence, $a_j(t)$ as a function of the flattened visual input $\mathbf{S}(t)$, resulting in

$$a_j(t) = \mathbf{V_j}^T \hat{\Delta}\mathbf{S}(t)$$

where the $\hat{\Delta}$ encodes the linear inhibition operation. The vector $\mathbf{W_j}$, row j of the sparse connection matrix $\hat{W}$, denotes the synaptic weights from the Kenyon Cell j to the input layer. In our model, $\mathbf{W_j}$ is a sparse binary vector of dimension equal to the number of PNs, with ones indicating the subset of inputs that KC j receives. The autocorrelations then result in the following expression:

$$\langle a_j(t) \cdot a_j(t+\tau)\rangle_t = \langle \mathbf{V_j}^T \hat{\Delta}\mathbf{S}(t)\mathbf{S}^T(t+\tau)\hat{\Delta}^T\mathbf{V_j}\rangle_t \tag{5}$$

Let us define:

$$\hat{Z}(\tau) = \langle \mathbf{S}(t)\mathbf{S}^T(t+\tau)\rangle_t$$

the correlation matrix of the visual inputs at lag τ. This quantifies how each component $S_i(t)$ of the visual input is correlated to each component $S_k(t+\tau)$ at a later time. It therefore encodes both spatial and temporal correlations across the sensory stream experienced during navigation. Since the sparse matrix $\hat{W}$ which connects the first and second layers is quenched and does not depend on time, the autocorrelation becomes:

$$\langle a_j(t) \cdot a_j(t+\tau)\rangle_t = \mathbf{V_j}^T \hat{\Delta}\hat{Z}(\tau)\hat{\Delta}^T\mathbf{V_j} \tag{6}$$

The computation of $\mathbf{Z}(\tau)$ highly relies on the statistics of the visual inputs presented to the neural network. For example, in the case of independent inputs, for $\tau = 0$ it results $\hat{Z}(\tau = 0) = \mathbb{I}$ the Identity matrix while for $\tau > 0$ it results $\hat{Z}(\tau) = \mathbf{0}$ the zero matrix.
From this we can normalize and then compute the expected values of the autocorrelations over the sparseness of the connectivity matrix $\hat{W}$, finding a correlation function which we will call $C(\tau)$.

$$C(\tau) = \frac{\mathbb{E}_{\hat{V}}\left[\mathbf{V_j}^T \hat{\Delta}\hat{Z}(\tau)\hat{\Delta}^T\mathbf{V_j}\right]}{\mathbb{E}_{\hat{V}}\left[\mathbf{V_j}^T \hat{\Delta}\hat{Z}(0)\hat{\Delta}^T\mathbf{V_j}\right]} \tag{7}$$

Since the elements of $\hat{W}$ are random variables $W_{ij} \in \{0, 1\}$ resulting from a Bernoulli distribution with sparseness given by the probability q, we can compute:

$$\begin{aligned}
\mathbb{E}_{\hat{V}}\left[\mathbf{V_j}^T \hat{\Delta}\hat{Z}(\tau)\hat{\Delta}^T\mathbf{V_j}\right] &= \mathbb{E}_{\hat{V}}\left[\sum_{iklh} V_{ji}\Delta_{ik}Z_{kl}(\tau)\Delta_{lh}V_{jh}\right] \\
&= \mathbb{E}_{\hat{V}}\left[\sum_{ikl} V_{ji}\Delta_{ik}Z_{kl}(\tau)\Delta_{li}V_{ji} + \sum_{\substack{iklh \\ i\neq h}} V_{ji}\Delta_{ik}Z_{kl}(\tau)\Delta_{lh}V_{jh}\right] \\
&= \sum_{ikl} \mathbb{E}_{\hat{V}}\left[V_{ji}^2\right]\Delta_{ik}Z_{kl}(\tau)\Delta_{li} + \sum_{\substack{iklh \\ i\neq h}} \mathbb{E}_{\hat{V}}\left[V_{ji}V_{jh}\right]\Delta_{ik}Z_{kl}(\tau)\Delta_{lh} \\
&= q \cdot \mathrm{Tr}(\hat{\Delta}\hat{Z}(\tau)\hat{\Delta}^T) + q^2 \cdot \sum_{\substack{iklh \\ i\neq h}} \Delta_{ik}Z_{kl}(\tau)\Delta_{lh}
\end{aligned} \tag{8}$$

2.2.2 Autoregressive Processes

Once the autocorrelations of the Kenyon Cell (KC) activations are computed, before applying the thresholding function, we can model their temporal dynamics using an autoregressive (AR) process.

An autoregressive process of order p, denoted AR(p), assumes that the current value of the signal $a(t)$ depends linearly on its p previous values, plus a stochastic noise term:

$$a(t) = \sum_{k=1}^{p} \varphi_k \, a(t-k) + \eta(t), \tag{9}$$

where φ_k are the autoregressive coefficients to be estimated, and $\eta(t)$ is a white noise term with zero mean. The parameter p determines how many past time steps influence the current activity, reflecting the temporal extent of the process's memory.

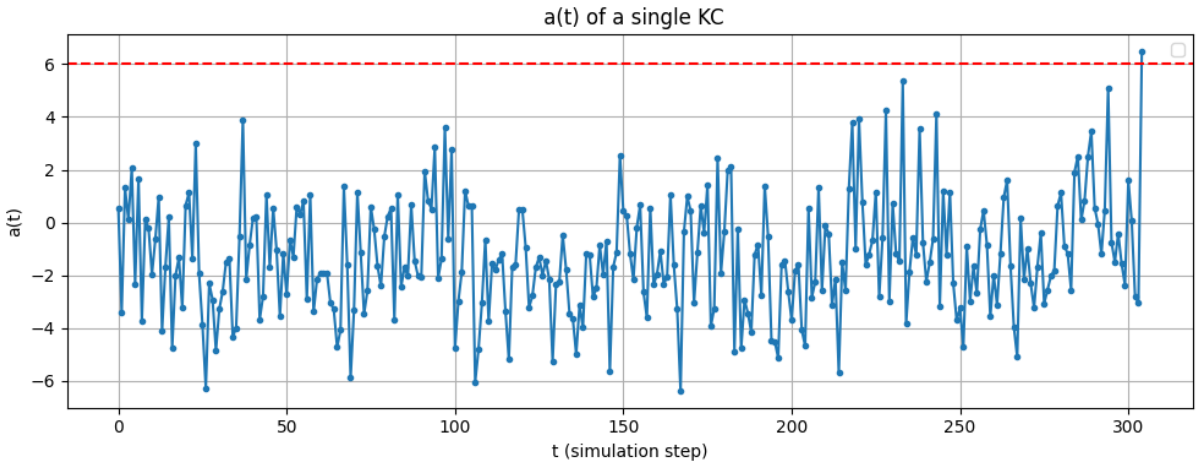

Figure 5: Example of the trajectory of a single KC activation value a(t) during a simulation. Once it reaches the threshold T (represented as the red line), the corresponding connection is set to 0.

Given the autocorrelation function

$$C(\tau) = \frac{\langle a(t)\, a(t-\tau)\rangle_t}{\langle a(t)^2\rangle_t}$$

derived as described in the previous section, the AR process coefficients φ_k can be estimated in the following way.

We multiply both sides of the AR equation by $a(t-k)$ and take the expectation:

$$\langle a(t)\, a(t-k)\rangle = \left\langle \left(\sum_{j=1}^{p} \varphi_j\, a(t-j) + \eta(t) \right) a(t-k) \right\rangle \tag{10}$$

$$= \sum_{j=1}^{p} \varphi_j\, \langle a(t-j)\, a(t-k)\rangle + \langle \eta(t)\, a(t-k)\rangle. \tag{11}$$

Since $\eta(t)$ is uncorrelated with $a(t-k)$ for $k \geq 1$, the second term vanishes:

$$\langle a(t)\, a(t-k)\rangle = \sum_{j=1}^{p} \varphi_j\, \langle a(t-j)\, a(t-k)\rangle. \tag{12}$$

Normalizing by dividing both sides by $\langle a(t)^2\rangle$ we get

$$C(\tau) = \sum_{k=1}^{p} \varphi_k\, C(\tau-k), \qquad \text{for } \tau = 1, \ldots, p. \tag{13}$$

which in matrix form reads as:

$$\begin{bmatrix} C(0) & C(1) & \cdots & C(p-1) \\ C(1) & C(0) & \cdots & C(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ C(p-1) & C(p-2) & \cdots & C(0) \end{bmatrix} \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_p \end{bmatrix} = \begin{bmatrix} C(1) \\ C(2) \\ \vdots \\ C(p) \end{bmatrix}. \tag{14}$$

Once the coefficients φ_k are obtained, the variance of the white noise $\eta(t)$ can be computed as:

$$\langle \eta_t^2 \rangle = \langle a(t)^2 \rangle - \sum_{k=1}^{p} \varphi_k^2 \langle a(t-k)^2 \rangle, \tag{15}$$

or, equivalently, using the identity:

$$\langle \eta_t^2 \rangle = \left(1 - \sum_{k=1}^{p} \varphi_k^2\right) \langle a(t)^2 \rangle, \tag{16}$$

where $\langle a(t)^2 \rangle$ is a known quantity. This provides a self-consistent way to estimate the noise strength in the AR model from the measured autocorrelations.

2.2.3 First Passage Time - Analytical Solution

The key quantity to analyze in the dynamics of the Mushroom Body neural network is the activity of the MBON at time t, which identifies a scene as familiar or unfamiliar, assuming the role of motor control. As described in the previous sections, a Kenyon Cell j generate a non-zero output only if its activation crosses the firing threshold. Once this occurs, the corresponding synaptic weight W_j that connects the KC to the Output Neuron is suppressed, preventing further contribution to the MBON at successive times. Hence, the MBON signal at time t arises exclusively from those KCs whose activation exceeds the threshold for the first time at t.

To quantify the expected MBON output, we are then interested in the expected number of KCs whose input activation $a_j(t)$ crosses the threshold T at time t for the first time. This is equivalent to computing the probability distribution of the first-passage time (FPT) of the stochastic processes $a_j(t)$, modeled here as an autoregressive process.

The first-passage time (FPT) is a fundamental quantity in the analysis of stochastic processes, representing the time it takes for a process to reach a given threshold for the first time. For autoregressive (AR) processes, the computation of the FPT distribution is generally challenging. In the case of an AR process of order one (AR(1)), exact analytical results can be obtained using different tools[16]. However, for higher-order processes, no exact solution is available. As a result, the FPT distribution is typically studied through numerical simulations, recursive algorithms, or approximate methods, such as those based on discretized Markov chains or Monte Carlo approaches [17, 18].

We present here the derivation of the first-passage time distribution for a first-order autoregressive (AR(1)) process[17], adopted in our study as a first approximation to model the dynamics of activation in the Mushroom Body.

Considering a discrete-time autoregressive process of order 1 (AR(1)):

$$a(t+1) = \varphi a(t) + \eta(t+1), \tag{17}$$

where $\eta(t) \sim \mathcal{N}(0, \sigma^2)$ are i.i.d. Gaussian innovations, and the process is initiated at $a(0) = a_0 \in \mathbb{R}$.

Let $T \in \mathbb{R}$ be the threshold. The first-passage time (FPT) is defined as:

$$\tau = \inf\{t > 0 \mid a(t) \geq T\}. \tag{18}$$

Define $q_t = \mathbb{P}(\tau = t)$. Then:

$$q_1 = \mathbb{P}(a(1) \geq T \mid a_0), \tag{19}$$
$$q_2 = \mathbb{P}(a(1) < T, a(2) \geq T \mid a_0), \tag{20}$$
$$q_3 = \mathbb{P}(a(1) < T, a(2) < T, a(3) \geq T \mid a_0), \tag{21}$$

and so on.

We also have:

$$\mathbb{P}(a(t) \geq T \mid a_0) = \sum_{s=1}^{t} q_s \cdot \mathbb{P}(a(t) \geq T \mid \tau = s), \tag{22}$$

where the conditional term is defined as:

$$\mathbb{P}(a(t) \geq T \mid \tau = s) = \frac{1}{\mathbb{P}(a(s) \geq T \mid a_0)} \int_T^\infty \mathbb{P}(a(s) = z \mid a_0) \cdot \mathbb{P}(a(t) \geq T \mid a(s) = z)\, dz. \tag{23}$$

From the AR(1) dynamics, we obtain:

$$\mathbb{E}(a(t+k) \mid a(t)) = \kappa \cdot \frac{1-\rho^k}{1-\rho} + \rho^k y_t, \tag{24}$$
$$\mathrm{Var}(a(t+k) \mid a(t)) = \sigma^2 \sum_{j=0}^{k-1} \rho^{2j} = \sigma^2 \cdot \frac{1-\rho^{2k}}{1-\rho^2}. \tag{25}$$

Hence, the transition probability is:

$$\mathbb{P}(a(t+k) \geq T \mid a(t)) = \Phi\left(\frac{\mathbb{E}(a(t+k) \mid a(t)) - T}{\sqrt{\mathrm{Var}(a(t+k) \mid a(t))}}\right), \tag{26}$$

where Φ is the standard normal CDF.

Let A_k be the vector of tail probabilities:

$$A_k = \begin{bmatrix} \mathbb{P}(a(1) \geq T) \\ \mathbb{P}(a(2) \geq T) \\ \vdots \\ \mathbb{P}(a(k) \geq T) \end{bmatrix},$$

and define the lower triangular matrix Θ_k with entries:

$$(\Theta_k)_{ij} = \begin{cases} 1 & \text{if } i = j, \\ \mathbb{P}(a(i) \geq T \mid \tau = j) & \text{if } j < i, \\ 0 & \text{if } j > i. \end{cases}$$

Then the FPT distribution vector $\vec{q} = (q_1, \ldots, q_k)^T$ satisfies the linear system:

$$A_k = \Theta_k \cdot \vec{q}, \qquad \Rightarrow \qquad \vec{q} = \Theta_k^{-1} \cdot A_k. \tag{27}$$

This approach, relying on conditional expectations and matrix inversion,is convenient for numerical implementation and allows to find exact results for the AR(1) process.

2.2.4 Stationary distribution

A fundamental property of AR processes is the existence of a stationary distribution under certain conditions. In particular, for the process to be weakly stationary, the roots of the characteristic polynomial

$$1 - \sum_{k=1}^{p} \varphi_k z^k = 0$$

must lie outside the unit circle in the complex plane. When this condition is satisfied, the process admits a unique stationary distribution with finite moments. For example, In the AR(1) case, stationarity requires $|\varphi| < 1$, and the process then converges to a Gaussian stationary distribution:

$$a(t) \sim \mathcal{N}\left(0,\, \frac{\sigma^2}{1-\varphi^2}\right). \tag{28}$$

This allows us to assess whether the KC activation processes converge to an stationary distribution and, moreover, it provides a way to verify the self-consistency of the modeling. This distribution plays a crucial role in the estimation of first passage times. this stationary distribution provides a well-defined prior for initializing the process $a_j(t)$ for each KC, and allows us to compute the FPT statistics without worrying about transient effects from arbitrary initial conditions. This is especially important for computing long-time MBON behavior and estimating the number of still-active synapses. The knowledge of the stationary density function allows us to compute statistical averages over the long-term behavior of the process.

2.3 Natural Images Statistics and Gaussian Random Field

Our goal was to generate synthetic visual landscapes that replicate the spatial statistics of real-world environments. To this end, we first analyzed the spatial structure of natural images by studying their power spectra in the Fourier domain. It is well established that natural scenes exhibit scale-invariant statistics, and their power spectra typically follow a $1/|\mathbf{k}|^2$ decay with a regularization at low frequencies. Classical studies such as [11, 12] showed that a commonly adopted model for the power spectrum is of the form

$$P(\mathbf{k}) = \frac{1}{|\mathbf{k}|^2 + \xi^{-2}},$$

where ξ is the correlation length of the image. This form yields exponentially decaying correlations in real space and is widely used to generate structured yet naturalistic input statistics. Although such spectra capture the general decay of spatial correlations, they often assume isotropy and therefore fail to reproduce the directional structure commonly found in natural scenes. For instance, natural landscapes present a strong horizontal and vertical feature. To incorporate this directional structure, we define the spatial frequency vector $\mathbf{k} = (k_x, k_y)$, where k_x and k_y are the horizontal and vertical spatial frequencies, respectively. To account for this anisotropy, we considered a more general power spectrum of the form:

$$P(\mathbf{k}) = \frac{1}{|k_x|^\alpha + |k_y|^\beta + m}, \tag{29}$$

where α and β control the rate of spectral decay along the x and y directions, respectively, and m encodes the information about the correlation length.

The validity of this theoretical model was evaluated by analyzing the spectral properties of real panoramic visual scenes, such as those used in insect navigation studies. We confronted this theoretical spectrum

with the power spectra of panoramic images used in ant navigation studies such as the one presented in Figure 6. Relying on this recent simulation by [6], and analyzing the power spectrum of those images used, we were able to confirm that such a spectral profile is a good approximation for the visual environments navigated by insects.

In particular, to validate the anisotropic nature of the input, we analyzed the decay of the empirical power spectrum along different directions, along f_x, f_y, and the diagonal 45° direction (Figure 7). In this figure, the log-log plot reveals that the slope of the power-law decay differs along each axis, providing clear evidence of anisotropic scaling properties. The different slopes observed in these directions demonstrate the anisotropic nature of the input images. These findings justify our choice of the anisotropic form for $P(\mathbf{k})$ and establish the relevance of the synthetic stimuli used in our simulations.

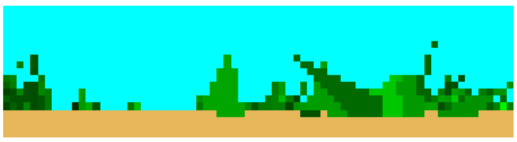

Figure 6: Example of a simulated ant view, adapted from [6]. The panoramic image is obtained by simulating what an ant sees at ground level in a naturalistic environment, capturing spatial features such as skyline boundaries and vertical elements that are critical for navigation.

To generate random fields with this prescribed power spectrum, we adopted the method described in [19]. We first sample a white noise field $\varphi(\mathbf{x})$ with unit amplitude and no spatial correlations, i.e., $\langle \varphi(\mathbf{x})\varphi(\mathbf{y})\rangle = \delta^d(\mathbf{x}-\mathbf{y})$. We then compute its Fourier transform $\widehat{\varphi}(\mathbf{k})$ and apply a frequency-domain filter:

$$\widehat{R}(\mathbf{k}) = P^{1/2}(\mathbf{k})\,\widehat{\varphi}(\mathbf{k}).$$

Finally, we perform an inverse Fourier transform to obtain the real-space field:

$$R(\mathbf{x}) = \int d^d\tilde{k}\, e^{i\mathbf{k}\cdot\mathbf{x}}\widehat{R}(\mathbf{k}).$$

This procedure ensures that the generated field $R(\mathbf{x})$ exhibits spatial correlations consistent with the desired spectrum $P(\mathbf{k})$, making it a biologically plausible proxy for panoramic visual input.

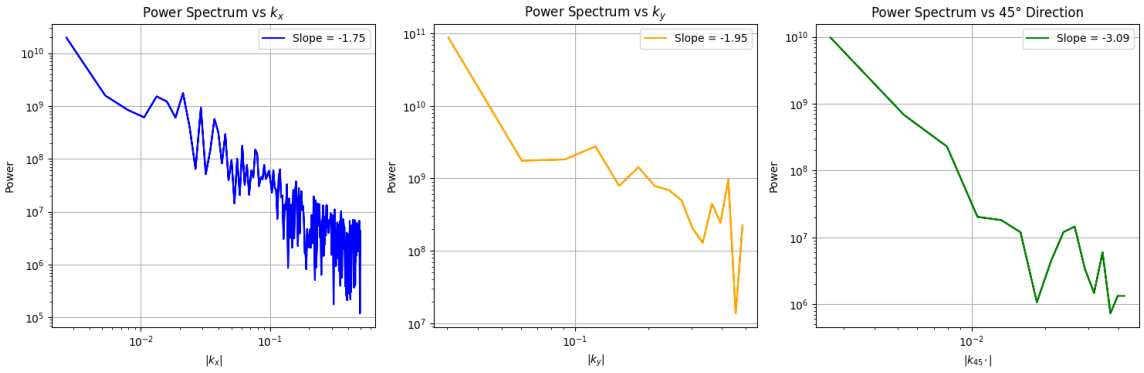

Figure 7: Directional power spectrum analysis of a panoramic image. The log-log plots show the power spectral density as a function of spatial frequency along the horizontal (k_x), vertical (k_y), and diagonal (k_{45°) directions. The distinct slopes confirm the presence of anisotropy in the image statistics.

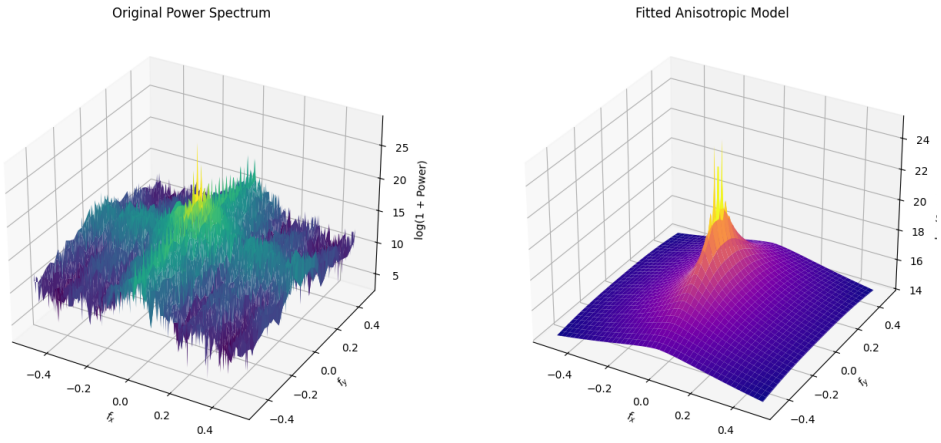

Figure 8: Example of the two-dimensional power spectrum of a panoramic scene along with a fitted power spectrum using the anisotropic model.

3 Results

We applied the theoretical framework described in the previous sections to a variety of input statistics, in order to explore how different stimulus properties influence the neural dynamics of the model. For each case, we computed analytical predictions based on the derived formalism and systematically compared them with the outcomes of numerical simulations. This validation step allows us to assess the accuracy and robustness of the theoretical approach.

3.1 Random Inputs

We simulated a simplified Mushroom Body neural network as described in section 2.1 using artificially generated stimuli composed of Gaussian i.i.d. values. Various network parameters were systematically explored, including the connection probability q of the random binary matrix linking Projection Neurons (PNs) to Kenyon Cells (KCs). In all subsequent simulations, the connectivity was fixed such that each KC received input from 25 randomly selected PNs. The number of KCs was always set to at least 10000, and the activation threshold was carefully chosen to ensure that only about $0.5\% - 1\%$ of KCs responded to each stimulus. These values were selected based on biological evidence reported in the literature and are consistent with parameters used in recent ant navigation simulations and studies [6, 5].

Two alternative mechanisms were implemented to generate sparse KC activations: either selecting the top-K most active KCs per input, or applying a fixed scalar threshold to determine KC activation. The binary activations of KCs were then linearly summed at the level of the MBON, and synaptic suppression was applied by setting to zero the weights corresponding to activated KCs. In order to quantify how familiarity signals degraded with time, the MBON threshold was set to zero, allowing us to track the cumulative loss of available synaptic weights. We monitored both the evolution of the MBON output and the decay in the number of functional KC-MBON synapses across simulation steps.

Since the stimuli are i.i.d. and lack temporal correlations, the dynamics do not follow an autoregressive process. Instead, the activity pattern at each time step can be modeled as an independent random vector. Recalling the form of the autocorrelations 6, we have:

$$\hat{Z}(\tau = 0) = \mathbb{I}, \qquad \hat{Z}(\tau > 0) = \mathbf{0}.$$

In this simplified case, where inputs are drawn from a standard normal distribution $\mathcal{N}(0, 1)$, the distribution of the KC input sums remains Gaussian. The variance of the summed input to each KC is given by:

$$\sigma^2 = 25 \cdot \frac{|\mathrm{nn}| + 1}{|\mathrm{nn}|},$$

where $|\mathrm{nn}|$ is the number of neighboring pixels used in the linear inhibition step, and the factor accounts for the added inhibitory contribution. However, a correction must be applied because border pixels do not have the same number of neighbors as those in the central region. In particular, if we use a square input where each pixel typically has 8 neighbors, we still divide by 8 in the linear inhibition step, even though edge pixels have only 5 neighbors and corner pixels have just 3. Therefore, this non-uniformity must be considered when computing the input variance. The same result can be obtained using the formalism described in section 6 as $\mathbb{E}_{\hat{W}}\left[\mathbf{W_j}^T \hat{\Delta} \hat{Z}(0) \hat{\Delta}^T \mathbf{W_j}\right]$

We derived theoretical predictions for the decay of the MBON output over time. The key quantity is the probability P that a given KC exceeds the activation threshold for a random input. Under the independence assumption, the probability that a KC remains unused after t steps is $(1-P)^t$. Considering

$\mathcal{O}(t)$ the number of connections that contribute to the MBON signal, which is equal to $\sum_j W_j(t)\,\mathrm{KC}_j(t)$, we can write:

$$\mathcal{O}(t) = N_{\mathrm{KC}} \cdot P \cdot (1-P)^t \tag{30}$$

where N_{KC} is the total number of Kenyon Cells (second layer neurons) in the implementation of the neural network. We also considered the $top-K$ implementation, where a fixed number K of KCs is activated at each step. In this case, each KC has a uniform chance of $P = \frac{K}{N_{\mathrm{KC}}}$ of being selected. The probability that a given KC survives up to time t is $(1-\frac{K}{N_{\mathrm{KC}}})^t$, so the expected MBON signal is:

$$\mathcal{O}(t) = K \cdot \left(1 - \frac{K}{N_{\mathrm{KC}}}\right)^t \tag{31}$$

This corresponds to a Bernoulli process with success probability K/N_{KC} at each step.

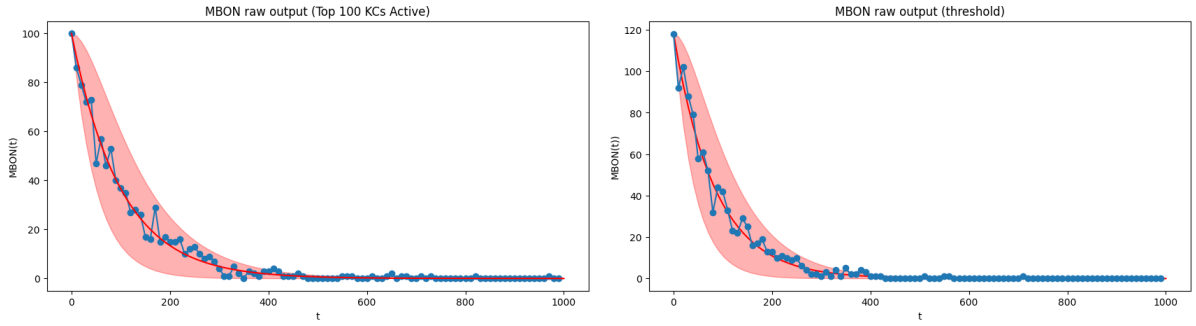

Figure 9: Raw MBON output both in the top-K active KCs (with K = 100) and the threshold implementation (with threshold T=12)

Finally, the expected number of remaining available KCs (whose synapse is still active) at time t can be computed as:

$$N_{KC}(t) = N_{\mathrm{KC}} - \sum_{t'=0}^{t} \mathcal{O}(t') \tag{32}$$

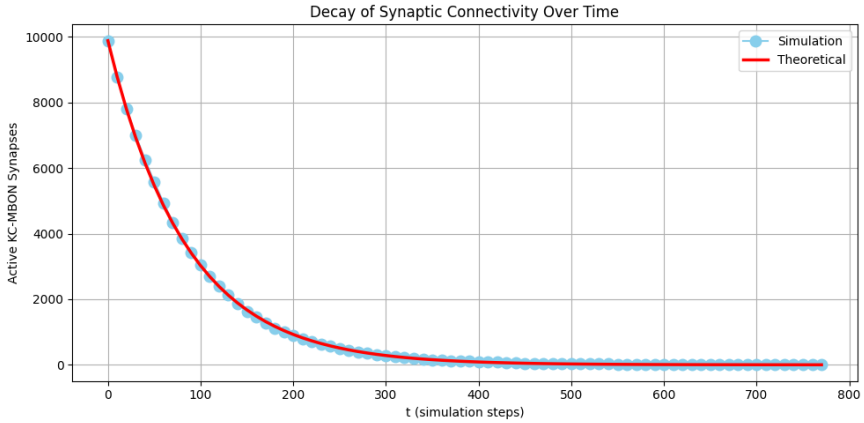

Figure 10: Availability of KC-MBON connections as a function of simulation steps t

As it is shown in figures 9 and 10, those theoretical curves shows good agreement with numerical simulations.

3.2 2D landscape

To explore how stimulus correlations affect memory saturation in our model, we constructed a two-dimensional landscape composed of Gaussian i.i.d. noise. This synthetic environment is represented as a large $10^4 \times 10^4$ matrix, where each pixel encodes a random intensity value. At each simulation step, an $N \times N$ image patch is extracted from the landscape based on the current position. Movement is implemented as discrete jumps across the landscape, allowing the agent to sample different but spatially structured inputs over time. While the pixel values are locally uncorrelated, the spatial continuity of motion introduces structured overlap between consecutive images, effectively creating a block-correlated input sequence. By navigating this visual environment, we can quantify the impact of structured input correlations on the internal dynamics of the neural network. In particular, we are interested in how the Linear Inhibition step shapes the temporal statistics of Kenyon Cell (KC) activation patterns.

In this case, the correlation between inputs is encoded by the matrix $\mathbf{Z}(\tau)$, which takes the form:

$$\hat{Z}(\tau) = \left[\begin{array}{c|c} \mathbf{0} & \mathbb{I} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right] \quad \text{with } \hat{Z}(\tau) \in \mathbb{R}^{(N^2)\times(N^2)} \text{ and } \mathbb{I} \in \mathbb{R}^{(N^2-N\cdot\tau)\times(N^2-N\cdot\tau)}$$

$$Z_{ij}(\tau) = \delta_{i+N\cdot\tau,\, j}$$

As discussed in Eq. 6,

$$\mathbb{E}_{\hat{W}}\left[\mathbf{W_j}^T \hat{\Delta} \hat{Z}(0) \hat{\Delta}^T \mathbf{W_j}\right] = q \cdot \mathrm{Tr}(\hat{\Delta}\hat{Z}(\tau)\hat{\Delta}^T) + q^2 \cdot \sum_{\substack{iklh \\ i\neq h}} \Delta_{ik} Z_{kl}(\tau) \Delta_{lh}$$

The trace encodes the direct autocorrelation of the same Projection Neuron (PN) across time, shaped by the Linear Inhibition step. Therefore, this term is non-zero only for $\tau = 0$, 1, and 2. The off-diagonal summation instead captures the possibility that different PNs receive the same portion of the landscape at different times. This term is non-zero only for $\tau < N$, within the overlap length between image patches.

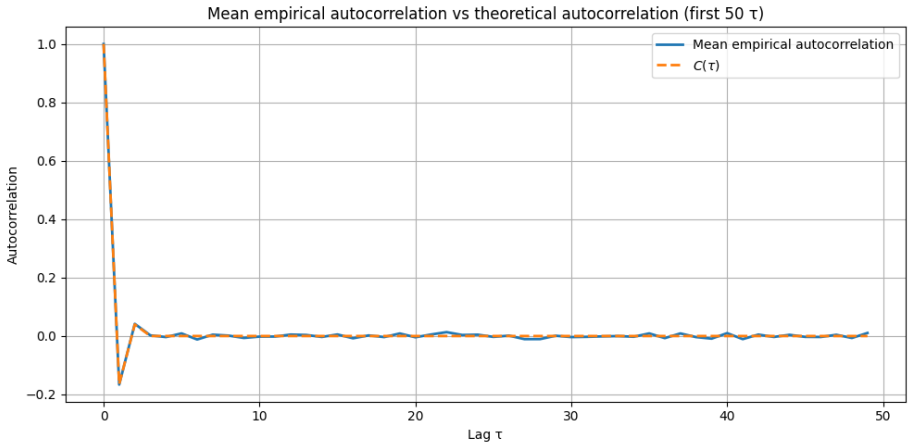

Figure 11: Comparison between the simulated mean autocorrelation as a function of lag τ and the theoretical prediction.

These autocorrelation effects have been verified and show excellent agreement with simulations.

The trace and the off-diagonal sum are of comparable magnitude, which is due to the sparsity of both $\hat{Z}$ and $\hat{\Delta}$ (each row of $\hat{\Delta}$ has only 9 non-zero entries). Given that $q = \frac{25}{N^2}$, the first term is of order $O(1)$ with respect to N, while the second term is of order $O(1/N^2)$ and can therefore be neglected. This validates the use of this uncorrelated landscape to isolate the effects of the Linear Inhibition step and the internal structure of the neural network. Even though the external inputs are uncorrelated, correlations emerge due to the network architecture. Thus, the activity $a(t)$ of each KC can be modeled as an autoregressive process of order 2.

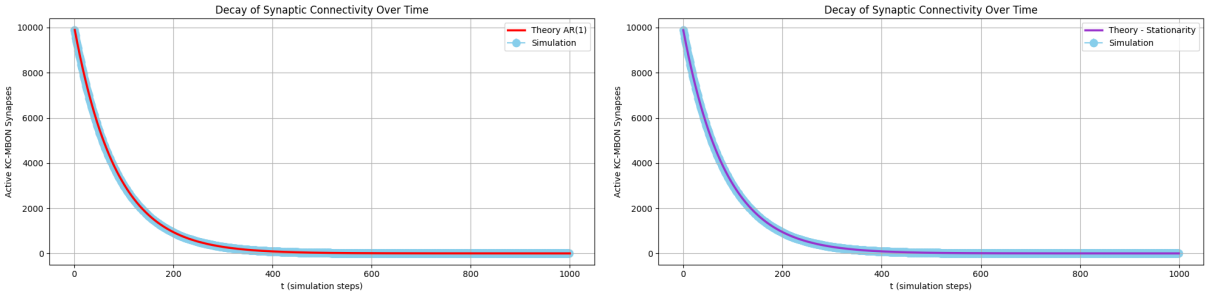

Figure 12: Comparison between simulation results and theoretical predictions for the availability of KC-MBON connections over simulation steps t. On the left, mean result of 10 simulations (threshold $T = 12$) compared with the theoretical curve obtained using the AR(1) approximation. On the right, the same simulations compared with the prediction from the stationarity approximation. Both theoretical models show good agreement with the simulation data.

To proceed with this modeling, we adopted the procedure previously described to fit autoregressive coefficients. However, since the second term is of order 10^{-2} and the first of order 10^{-1}, as a first approximation we employed the exact solution of the first-order autoregressive process to study its behavior. We then also used the stationary approach as described in section 2.2.4, deriving the probability P of $a(t)$ to reach the threshold and finding the observables $\mathcal{O}(t)$ and $N_{KC}(t)$ as done in the uncorrelated case in equations 30 and 32. The results from both methods show similar performance, as the contribution from the second term is small, confirming the robustness of the approximation, and both of them show very good agreement with simulations as it is shown in Figure 12.

3.3 2D natural statistics landscape

To assess the impact of realistic spatial correlations on the neural dynamics, we generated synthetic visual environments whose statistics emulate those of natural images. Using the method described in Section 2.3, we constructed two-dimensional Gaussian random fields, mimicking the structure of natural scenes encountered by ants during navigation.

As in the previous case, the agent explores these landscapes by extracting overlapping $N \times N$ image patches along a linear trajectory. As before, this generates temporally correlated input sequences for the network. However, in contrast to the uncorrelated case, the spatial correlations embedded in the landscape result in long-range temporal correlations in the input. This implies that the observed autocorrelations are no longer solely a consequence of patch overlap, but are now intrinsically present in the stimulus statistics.

The effect of these correlations on the network dynamics is illustrated in Figure 13, where we compare the predicted and simulated autocorrelation functions. We observe significantly longer memory traces in the Kenyon Cell (KC) activity compared to the white noise case.

Even though the correlated landscape is normalized to have the same pixel signal S_i variance as the uncorrelated one, its internal structure leads to reduced local variance after Linear Inhibition is applied.

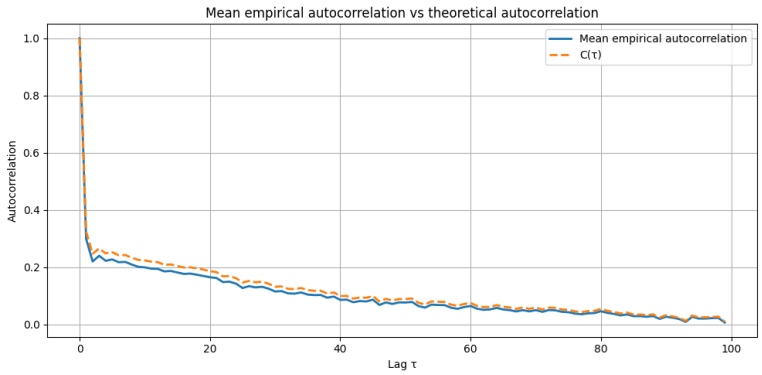

Figure 13: Comparison between the simulated and predicted mean autocorrelation as a function of lag τ in the correlated landscape generated with natural statistics.

This is because the pixel values are locally redundant, and subtracting nearby values (as done in the inhibition step) cancels out much of the signal.

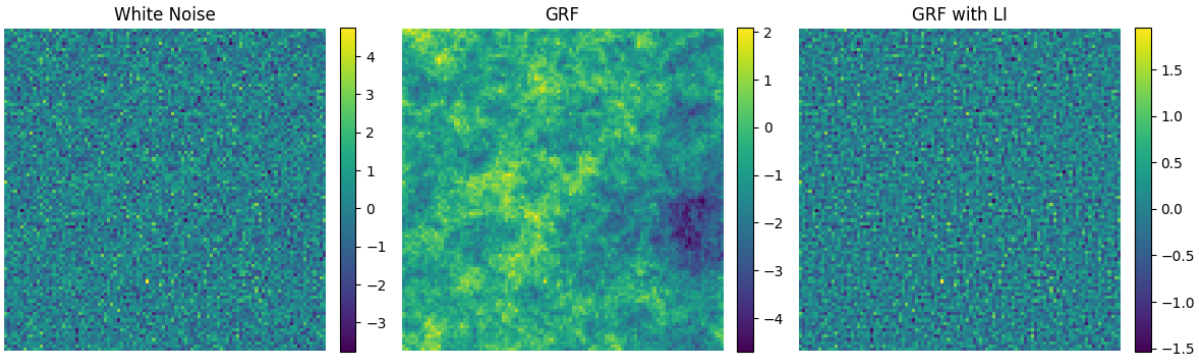

Figure 14: Visual inputs comparison. Left: white noise landscape. Center: spatially correlated landscape generated via Gaussian Random Field. Right: landscape after Linear Inhibition. The inhibition decorrelates and reduces the variance of the input.

As a result, KCs require a lower activation threshold to respond. Figure 14 shows how this variance reduction affects the output $\mathcal{O}(t)$ under the same network architecture, including the threshold. We compared the same network dynamics under both uncorrelated and correlated conditions, keeping the threshold fixed. As shown in Figure 15, uncorrelated inputs lead to higher average activity which causes earlier memory saturation.

As in the previous section, we mapped the KC activity $a(t)$ to an autoregressive process, fitting its coefficients using the method described earlier. In this case, however, the correlation structure of the input is significantly more persistent in time, due to the structured nature of the landscape. The correlation matrix $\mathbf{Z}(\tau)$ was therefore computed numerically by averaging over multiple trajectories sampled from the landscape. It this case we have long range correlations and as a consequence, the first-order exact solution derived previously is no longer applicable. Instead, we adopted the stationary approach, computing the transition probability P and following the same procedure as illustrated before. While this method does not match the simulation results as accurately as in the uncorrelated case, it still provides a useful indicator of the qualitative behavior of the network under temporally correlated stimuli. The slight discrepancy between theory and simulation may stem from a mild overestimation of the effective correlation, which in turn leads to a slight overestimation of the variance of the equilibrium distribution.

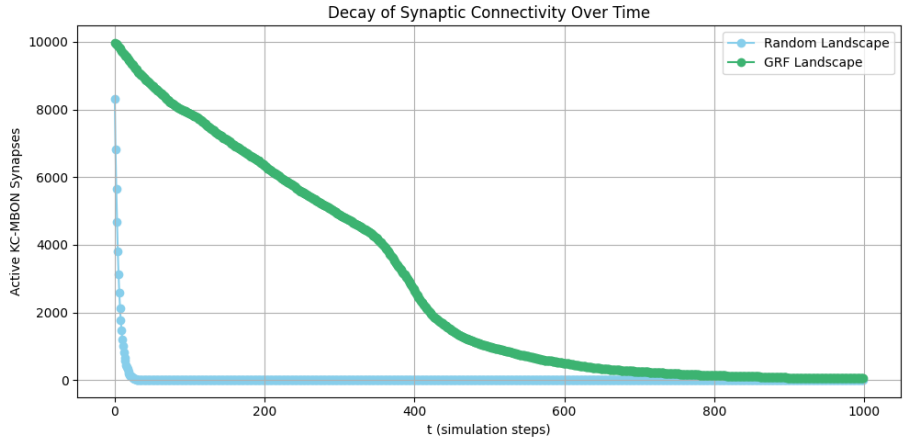

Figure 15: Comparison between the performance of the same neural network (Under the form of KC-MBON synapse availability) in both the uncorrelated and the correlated landscape. We can see that in the uncorrelated landscape the connections are suppressed much faster in comparison to the correlated landscape even though the variance of the initial input is the same.

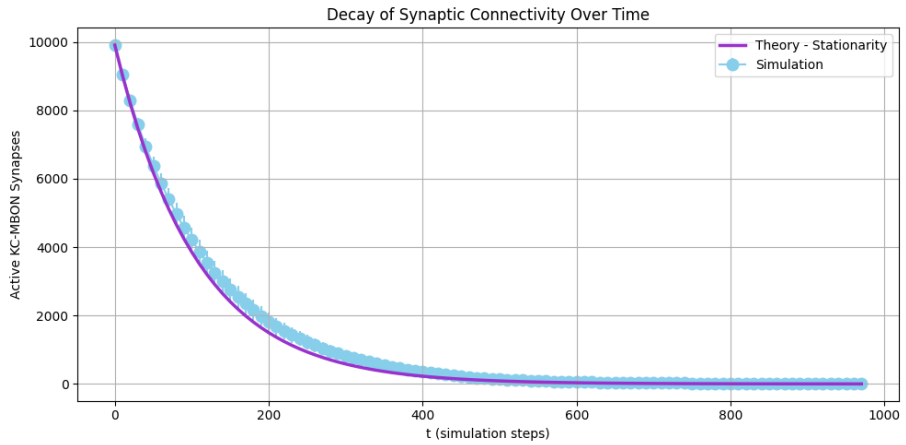

Figure 16: Mean result of 10 simulations (threshold $T = 5$) compared with the prediction from the stationarity approximation.

3.4 Memory saturation

Behavioral experiments have shown that visually navigating ants are capable of storing route memories over ecologically relevant distances. This offers a rare opportunity to ground theoretical studies of memory capacity in measurable data. To interpret our simulation results in terms of real navigation behavior, we associate each simulation step with a spatial distance. This mapping depends on the walking speed of the ant and the time interval between successive memory updates, which in turn is governed by the rate of synaptic suppression. From the theoretical curves, we can infer the point at which no novel stimulus will cause a response in the MBON, when memory is completely saturated. This gives us an estimate of $L_{\max}$, the maximal distance of a navigational path that the current network architecture is capable of storing.

For dopaminergic modulation, the biological process that regulates synaptic suppression, direct measurements of suppression rates in ant navigation are currently lacking. Most of our understanding stems from studies in Drosophila Melanogaster and honeybees signals [20] but in ants the underlying neuromodulatory dynamics are less well characterized [4, 5]. Analogously to findings in Drosophila, suppression or facilitation of Kenyon Cell–MBON synapses by dopaminergic neurons can occur on fast timescales, typically within hundreds of milliseconds to a few seconds [21, 22]. These insights suggest that dopaminergic suppression in ants could plausibly operate on similar timescales.

Considering a timescale for synaptic suppression of $T = 1\ s$ and considering an average ant velocity of $v = 1,2\ cm/s$, which has been observed during learning walks of the species Myrmecia Croslandi ants [23], we can get an estimate of the maximum length of the path. In fact, considering that a simulation step correspond to $v \cdot T$ and that any scene would be perceived as unfamiliar once all the KC-MBON connections are set to 0, recognizing this moment as memory saturation and deriving it from the analysis in Section 3.3, we can find an estimate of the maximum length path $L_{max} \simeq 12\ m$. This result aligns with values reported in the same behavioral experiments from which the walking speed was derived and with other studies on the same species[23, 24] where maximum paths of $10 - 15\ m$ were observed. This agreement supports the robustness of the model and confirms the central role of the mushroom body in ant navigation, as well as the relevance of the theoretical framework employed.

4 Conclusion

In this work, we have constructed a theoretical and computational framework to analyze how the statistical properties of sensory inputs modulate learning and memory in Mushroom Body (MB) inspired neural networks. By drawing direct connections between visual input statistics, network architecture, and dynamical behavior, we demonstrated how sparse plastic networks can store and retrieve structured sensory information over extended temporal horizons. A central contribution of our study is the reinterpretation of KC activation dynamics as a first passage time (FPT) problem, providing a powerful analytical lens to study memory saturation in networks governed by synaptic depression. This probabilistic perspective allowed us to derive predictions for memory capacity and synaptic availability under different classes of inputs. In particular, we showed that under uncorrelated input statistics, the dynamics of KC activation follow simple probabilistic rules that are well captured by autoregressive processes. These simplified dynamics allowed us to derive tractable expressions for quantities such as the number of active synapses and the MBON output as functions of time, offering analytical insight into the learning dynamics of the MB. Building on this, we introduced a biologically grounded interpretation of memory saturation by linking simulation time steps to real spatial distances based on the ant's walking speed and the estimated timescale of dopaminergic synaptic suppression. This enabled us to derive a realistic estimate of the maximum path length $L_{\max}$ that the network can store before full memory saturation occurs, yielding a value of approximately 12 meters. This prediction is consistent with experimental observations of trajectories in Myrmecia Croslandi ants, thus supporting the biological plausibility of the model.

When input correlations were introduced, we observed longer memory traces in the network dynamics for the same architecture. Our results also underscore the critical role of the Linear Inhibition step, not only as a contrast-enhancing preprocessing mechanism, but as a key transformation that reshapes the internal statistics of the input and helps regulate network sparsity. This operation introduces nontrivial correlations across Projection Neurons (PNs), which are then inherited by the Kenyon Cells (KCs) and influence their threshold crossing statistics. In particular, we showed that even for inputs that are locally uncorrelated, the architecture and inhibition dynamics introduce internal temporal structure that can be well described by low-order AR processes. Importantly, we found that the memory capacity of the system depends on several key parameters: the number of Kenyon Cells (KCs), which is biologically constrained; the activation threshold, which mimics the effect of APL-mediated inhibition; and the statistical structure of the visual landscape being explored. Understanding how this threshold is biologically implemented, and whether it remains fixed or adapts to the environment, could provide insights into whether learning occurs not only at the KC-MBON synapses but also in the inhibitory circuitry itself.

Beyond its biological relevance, our work highlights the power of combining statistical physics tools with neural modeling to study learning in structured environments. The simplicity of the MB architecture, with its sparse connectivity and localized plasticity, makes it ideal for developing tractable theories of learning and memory. Looking forward, this framework opens multiple avenues for further exploration. From a theoretical perspective, one could extend the analytical treatment to different input statistics. More generally, our approach suggests that incorporating the statistical structure of natural inputs into neural modeling is not only biologically realistic, but also essential for understanding how real-world environments shape the learning capabilities of neural systems.

In summary, the Mushroom Body represents an interesting intersection of biological realism and analytical tractability. By bridging simulation, theory, and biological constraints, we have shown that key properties of memory and familiarity detection in insect-inspired neural systems can be rigorously characterized.

References

[1] Martin Egelhaaf et al. “Spatial vision in insects is facilitated by shaping the dynamics of visual input through behavioral action”. In: *Frontiers in neural circuits* 6.108 (2012).

[2] J. Perez-Orive. “Oscillations and Sparsening of Odor Representations in the Mushroom Body”. In: *Science* 297 (2002), pp. 359–365.

[3] Yoshinori Aso et al. “Mushroom body output neurons encode valence and guide memory-based action selection in Drosophila”. In: *eLife* 3 (2014), e04580.

[4] Cornelia Buehlmann et al. “Mushroom Bodies Are Required for Learned Visual Navigation, but Not for Innate Visual Behavior, in Ants”. In: *Current Biology* 30 (2020), 3438–3443.e2.

[5] Paul Ardin et al. “Using an Insect Mushroom Body Circuit to Encode Route Memory in Complex Natural Environments”. In: *PLoS Computational Biology* 12.2 (2016), e1004683.

[6] Florent Le Moël and Antoine Wystrach. “Vision is not olfaction: impact on the insect Mushroom Bodies connectivity”. In: *bioRxiv* (2024).

[7] Bart Baddeley et al. “A Model of Ant Route Navigation Driven by Scene Familiarity”. In: *PLoS Computational Biology* 8.1 (2012), e1002336.

[8] Florent Le Moël and Antoine Wystrach. “Opponent processes in visual memories: A model of attraction and repulsion in navigating insects’ mushroom bodies”. In: *PLoS Computational Biology* 16.2 (2020), e1007631.

[9] Ken Cheng et al. “Beginnings of a synthetic approach to desert ant navigation”. In: *Behavioural Processes* 102 (2014). A Synthetic Approach to Comparative Cognition: A Tribute to the Contributions of Alan C. Kamil, pp. 51–61. ISSN: 0376-6357.

[10] Mehrab N. Modi, Yichun Shuai, and Glenn C. Turner. “The Drosophila Mushroom Body: From Architecture to Algorithm in a Learning Circuit”. In: *Annual Review of Neuroscience* 43 (2020), pp. 465–484.

[11] David J Field. “Relations between the statistics of natural images and the response properties of cortical cells”. In: *Journal of the Optical Society of America A* 4.12 (1987), pp. 2379–2394.

[12] Daniel L Ruderman and William Bialek. “Statistics of natural images: Scaling in the woods”. In: *Physical Review Letters* 73.6 (1994), pp. 814–817.

[13] N. J. Strausfeld and J. A. Campos-Ortega. “Vision in Insects: Pathways Possibly Underlying Neural Adaptation and Lateral Inhibition”. In: *Science* 195.4281 (1977), pp. 894–897.

[14] Douglas M Bolzon, Karin Nordström, and David C O’Carroll. “Local and large-range inhibition in feature detection”. In: *The Journal of Neuroscience* 29.42 (2009), pp. 14143–14150.

[15] Xin Liu. “Inhibition SNN: unveiling the efficacy of various lateral inhibition learning in image pattern recognition”. In: *Discover Applied Sciences* 6.11 (2024), p. 611.

[16] Alexander Novikov and Nino Kordzakhia. “Martingales and first passage times of AR(1) sequences”. In: *Stochastics: An International Journal of Probability and Stochastic Processes* 80.2-3 (2008), pp. 197–210.

[17] Marcin Jaskowski and Dick van Dijk. *First-Passage-Time in Discrete Time*. Econometric Institute Research Papers. Jan. 2015.

[18] Elvira Di Nardo. *On the first passage time for autoregressive processes*. Jan. 2008.

[19] Garrett Goon. *Gaussian Fields*. `https://garrettgoon.com/gaussian-fields/`. Accessed: 2025-06-06. 2020.

[20] Martin Schwaerzel and et al. “Dopamine and octopamine differentiate between aversive and appetitive olfactory memories in Drosophila”. In: *Journal of Neuroscience* 23.33 (2003), pp. 10495–10502.

[21] Annie Handler et al. “Distinct Dopamine Receptor Pathways Underlie the Temporal Sensitivity of Associative Learning”. In: *Cell* 178 (June 2019).

[22] Kristina V. Dylla and Elizabeth J. Hong. “Mapping Odor to Action: (Dopaminergic) Timing Is Everything”. In: *Cell* 178.1 (2019), pp. 5–7. ISSN: 0092-8674.

[23] Piyankarie Jayatilaka et al. “The choreography of learning walks in the Australian jack jumper ant Myrmecia croslandi”. In: *Journal of Experimental Biology* 221.20 (2018), jeb185306.

[24] Ajay Narendra, Sarah Gourmaud, and Jochen Zeil. “Mapping the navigational knowledge of individually foraging ants, *Myrmecia croslandi*”. In: *Proceedings of the Royal Society B: Biological Sciences* 280.1765 (2013), p. 20130683.