PREDICTION AND GENERALISATION OVER DIRECTED ACTIONS BY GRID CELLS

Anonymous authors
Paper under double-blind review

ABSTRACT

Knowing how the effects of directed actions generalise to new situations (e.g. moving North, South, East and West, or turning left, right, etc.) is key to rapid generalisation of action to new situations. Markovian tasks can be characterised by a state space and a transition matrix and recent work has proposed that neural grid codes provide an efficient representation of the state space, as eigenvectors of a transition matrix reflecting diffusion across states, that allows efficient prediction of future state distributions. Here we address the ability to predict the effects of directed actions across new tasks, irrespective of differences in local transition structure, to allow fast generalisation and form a "sense of direction". We extend the eigenbasis prediction model, utilising tools from Fourier analysis, to prediction over arbitrary translation-invariant directed transition structures (i.e. displacement and diffusion), showing that a single set of eigenvectors can support predictions over arbitrary directed actions via action-specific eigenvalues. We show the equivalence between the generalised prediction framework and traditional models of grid cells firing driven by self-motion to perform path integration using oscillatory interference models (via Fourier components as velocity-controlled oscillators) or continuous attractor networks (via analysis of the update dynamics). We thus provide a unifying framework for the role of the grid system in predictive planning, sense of direction and path integration. The resulting model allows a single grid-like representation to support efficient prediction over directed transitions in spatial and non-spatial tasks, supporting generalisable inference over directed actions across different tasks.

1 INTRODUCTION

A "cognitive map" encodes relations between objects and supports flexible planning (Tolman [40]), with hippocampal place cells and entorhinal cortical grid cells thought to instantiate such a map (O'Keefe and Dostrovsky [33]; Hafting et al. [21]). Each place cell fires when the animal is near a specific location, whereas each grid cell fires periodically when the animal enters a number of locations arranged in a triangular grid across the environment. Together, this system could support representation and flexible planning in state spaces where common representational structure is preserved across states and tasks, affording generalisation and inference, e.g., in spatial navigation where Euclidean transition rules are ubiquitous (Whittington et al. [43]).

Recent work suggests that place cell firing provides a local representation of state occupancy, while grid cells comprise an eigenbasis of place cell firing covariance (Dordek et al. [16]; Stachenfeld et al. [39]; Sorscher et al. [38]; Kropff and Treves [27]). Accordingly, grid cell firing patterns could be learned as eigenvectors of a symmetric (diffusive) transition matrix over state space, providing a basis set enabling prediction of occupancy distributions over future states. The "intuitive planning" operates by replacing multiplication of state representations by the transition matrix with multiplication of each basis vector by the corresponding eigenvalue (Baram et al. [2]; Corneil and Gerstner [14]). Thus a distribution over state space represented as a weighted sum of eigenvectors can be updated by re-weighting each eigenvector by its eigenvalue to predict future state occupancy.

Fast prediction and inference given different actions across environments is vital and economical for the survival of animals, and intuitive planning, in its original form, supports such ability under a single transition structures, most often corresponding to symmetrical diffusion. Here we assume

a common translation-invariant transition structure underlying multiple specific tasks which break translation invariance in different ways (e.g., Euclidean spatial environments with periodic boundary conditions containing different patterns of obstacles) and generalise the prediction models with grid codes to incorporate directed transitions based on Fourier analysis. We demonstrate the utility of the proposed theoretical framework by showing its ability for performing prediction and path integration in a number of simple yet illustrative examples. We define a "sense of direction" quantity, i.e., which proportions of directed actions would best lead to the goal, based on the underlying transition structure ignoring local obstacles, and can be computed based on the "Fourier modes" cells under our formulation, which we propose as a computational role for the medial entorhinal grid cells. We then show, on the theory-level, how different classes of computational models for grid cells could be unified, under the assumption that the grid cells firing result from superpositions of inputs from neurons with Fourier-mode-like firing patterns. Hence demonstrating that the proposed model acts as a unifying theoretical framework for understanding grid cell firing.

2 "INTUITIVE PLANNING" WITH SYMMETRICAL TRANSITIONS

Intuitive planning represents the occupancy distribution over the state space as a weighted sum of the (real) eigenvectors of the (symmetric) transition matrix, so that the effect of one step of the transition dynamics on the distribution can be predicted by reweighting each of the eigenvectors by the corresponding eigenvalue. And this generalises to calculating the cumulative effect of discounted future transitions (Baram et al. [2]).

Specifically, consider a transition matrix, $T \in \mathbb{R}^{N\times N}$, $T_{ss'} = \mathbb{P}(s_{t+1} = s'|s_t = s)$ where s_t encodes the state at time t and N is the number of states. Then, T^n is the n-step transition matrix, and has the same set of eigenvectors as T. Specifically, the eigendecomposition of T and T^n are:

$$T = Q\Lambda Q^{-1}, \qquad T^n = Q\Lambda^n Q^{-1} \tag{1}$$

where each column of the matrix Q is an eigenvector of T and $\Lambda = diag(\sigma_P(T))$, where $\sigma_P(T)$ is the set of eigenvalues of T. Similarly, any polynomial in T, $p(T)$, shares the same set of eigenvectors as T and the set of eigenvalues $\sigma_P(p(T)) = p(\sigma_P(T))$. Hence:

$$\sum_{k=0}^{\infty}(\gamma T)^k = (I - \gamma T)^{-1} = Q\text{diag}(\mathbf{w})Q^{-1}, \quad \text{where } \mathbf{w} = \left\{\frac{1}{1-\gamma\lambda}, \text{ for } \lambda \in \sigma_P(T)\right\} \tag{2}$$

The resolvent form (eq. 2) is an infinite discounted summation of transitions, which under a policy and transition structure corresponding to diffusion, is equivalent to the successor representation (SR, Fig. 1E) with discounting factor γ (Dayan [15]; Stachenfeld et al. [39]). See Mahadevan and Maggioni [30] for a related spectral approach using Fourier decomposition of T for estimating the value function. The SR has been shown to be useful for navigation via gradient ascent of the future probability of occupying the target state, and has a linear relationship with the true underlying Euclidean distances in spatial tasks (hence "intuitive planning", see Fig. 1 and Fig. 2D-E).

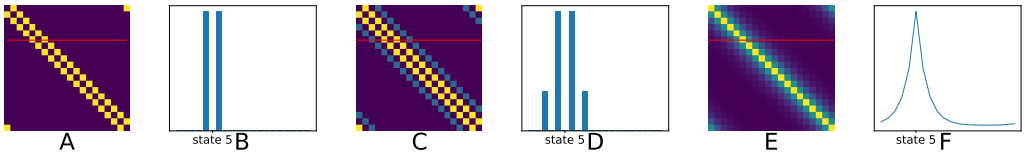

Figure 1: **Demonstration of intuitive planning on a diffusive transition task on a 1D ring track**. **A**: Example transition matrix; **B**: $\mathbb{P}(s_{t+1} = s'|s_t = 5)$; **C, D**: same are shown for T^3, showing predicted distribution over the next three time steps; **E** the resolvent form/SR (eq. 2) computed from the eigenbasis of the transition matrix; **F**: SR values for state 5, which can used for navigation.

The eigenvectors of the diffusion transition matrix generally show grid-like patterns, suggesting a close relationship to grid cells. However, intuitive planning cannot predict the effects of directed actions or drift (i.e., variable asymmetric transition structure), which is useful in itself and also corresponds to the presumed role of grid cells in path integration. Moreover, predictions over different directed actions would require different eigendecompositions, hence incurring high computational costs that

undermines its biological plausibility. In Section 3 we unify the prediction and path integration approaches by exploiting translation invariant symmetries to generalise across actions, using a single common eigenbasis and cheaply calculated updates via action-dependent eigenvalues.

3 FLEXIBLE PLANNING WITH DIRECTED TRANSITIONS

Updating state representations to predict the consequences of arbitrary directed actions is an important ability of mobile animals, known as path integration and thought to depend on grid cells (McNaughton et al. [31]). To generalise to directed transition structures, we consider the transition dynamics corresponding to translation (drift) and Gaussian diffusion with arbitrary variance (including 0, equivalent to plain translation). Our assumption that the transition structure is translation invariant (implying periodic boundary conditions), leads to circulant transition matrices.

Consider a 2D rectangular environment with length L and width W where each state is a node of the unit square grid, then the transition matrix can be represented by $\mathbf{T} \in \mathbb{R}^{LW \times LW}$, with each row the vectorisation ($\mathbf{vec}(\cdot)$) of the matrix of transition probabilities starting from the specified location, i.e., $\mathbf{T}[i,:] = \mathbf{vec}[\mathbb{P}(s(t+1)|s(t)=i)]$, where $\mathbf{T}$ is constructed by considering the 2D state space as a 1D vector and concatenating the rows ($i = xL + y$ for $(x, y) \in [0, W-1] \times [0, L-1]$), see Fig. 2A.

The transition matrix is circulant due to the translation invariance of the transition structure (see Appendix Prop. A.1), and takes the following form:

$$\mathbf{T} = \begin{bmatrix} T_0 & T_{LW-1} & \cdots & T_2 & T_1 \\ T_1 & T_0 & T_{LW-1} & \cdots & T_2 \\ \vdots & T_1 & T_0 & \ddots & \vdots \\ T_{LW-2} & \cdots & \ddots & \ddots & T_{LW-1} \\ T_{LW-1} & T_{LW-2} & \cdots & T_1 & T_0 \end{bmatrix} \tag{3}$$

The normalised eigenvectors of the circulant matrix $\mathbf{T} \in \mathbb{R}^{N \times N}$ ($N = LW$) are the vectors of powers of Nth roots of unity (the Fourier modes):

$$\mathbf{q}_k = \frac{1}{\sqrt{N}} \left[1, \omega_k, \omega_k^2, \cdots, \omega_k^{N-1}\right]^T \tag{4}$$

where $\omega_k = \exp(\frac{2\pi i}{N} k)$, for $k = 0, \ldots, N-1$, and $i = \sqrt{-1}$. Hence the matrix of eigenvectors (as the columns), $F = (\mathbf{q}_0, \mathbf{q}_1, \ldots, \mathbf{q}_{N-1})$, is just the (inverse) discrete Fourier transform matrix (Bracewell [4]), where $F_{kj} = \omega_j^k$ for $0 \leq k, j \leq N-1$. The Fourier modes projected back onto the $L \times W$ 2D spatial domain are plane waves, as shown in Fig. 2G, with wavevector determined by the value of k that specifies the direction and spatial frequency of each plane wave (see Appendix B). We can immediately compute the corresponding eigenvalues for the eigenvectors in eq. 4 (equivalent to taking the discrete Fourier transform (DFT) of the first row (or column) of T, see Bracewell [4]):

$$\lambda_m = \sum_{j=0}^{N-1} T_j \omega_j^m, \quad \text{for } m = 0, \ldots, N-1 \tag{5}$$

where $\{T_0, \ldots, T_{N-1}\}$ are the N unique elements that fully specifies the circulant matrix T (eq. 3).

We can then utilise tools from Fourier analysis for efficient updating of the eigenvalues whilst leaving the universal eigenbasis unaffected. For a transition matrix $\mathbf{T}^{\mathbf{v}}$ corresponding to an arbitrary action (translation velocity) $\mathbf{v} = (v_x, v_y)$, each row of $\mathbf{T}^{\mathbf{v}}$ is again a circulant, but shifted version of the corresponding row vector of the symmetric transition matrix corresponding to zero drift velocity, $\mathbf{T}^0$. Specifically, the first rows of the two matrices are related as follows:

$$\mathbf{T}^0(k + v_x L + v_y) = \mathbf{T}^v(k), \quad \text{for } k = 0, \ldots, N-1 \tag{6}$$

Given the eigenvalues for $\mathbf{T}^0$, $\Lambda^0 = \left[\lambda_0^0, \lambda_1^0, \ldots, \lambda_{N-1}^0\right] \in \mathbb{C}^N$ (via the DFT of the first row of $\mathbf{T}^0$, eq. 5), we can immediately derive the eigenvalues of $\mathbf{T}^{\mathbf{v}}$, $\Lambda^{\mathbf{v}}$, via a one-step update based on the Fourier shift theorem (Bracewell [4]) without recomputing the eigendecomposition:

$$\Lambda^{\mathbf{v}}[k] = \exp\left(\frac{2\pi i}{N}(v_x L + v_y)k\right)\Lambda^0[k] \quad \text{for } k = 0, \ldots, N-1 \quad \text{for arbitrary } \mathbf{v} \tag{7}$$

$$\text{i.e., } \Lambda^{\mathbf{v}} = \Phi_{\delta^{\mathbf{v}}}\Lambda^0, \quad \Phi_{\delta(\mathbf{v})} = \left[1, \omega_{\delta(\mathbf{v})}, \omega_{\delta(\mathbf{v})}^2, \ldots, \omega_{\delta(\mathbf{v})}^{N-1}\right], \quad \text{where } \delta(\mathbf{v}) = v_x L + v_y \tag{8}$$

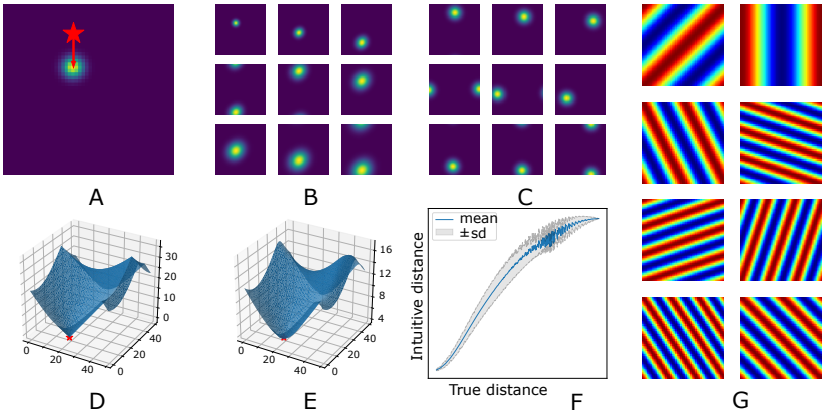

Figure 2: **Demonstration of our method in a 2D 50×50 environment with periodic boundary conditions**. **A**: Transition matrix $\mathbb{P}(x_{t+1} = (i', j')|x_t = (i, j))$ starting from a randomly chosen state ($(8, 10)$; red star) with drift velocity 10 units southward and Gaussian diffusion (red arrow); **B**: Usage of the eigenbasis (eq. 4) and our analysis (eq. 8) for predicting the distribution over future states given the transition structure given in A, showing successive changes to state occupancy; **C**: Application of our model to translation-only transition structures (drift velocity $v = (3, 3)$); **(D-E)**: Ground-truth shortest distance to state $(8, 20)$ (red star; D) as a function of the starting location, and distance estimated under the intuitive planning framework (E) using plain diffusion; **F**: Estimated distance measure (E) v.s. the corresponding ground-truth distance (D) over all pairs of states; **G**: Examples of 2D Fourier modes (real parts shown, see Appendix Fig. 7 for the top 100 eigenvectors).

Hence allows path integration by reweighting the common set of eigenvectors at each timestep by the updated eigenvalues corresponding to the current drift velocity (eq. 8). Note that additionally, T^0 can include diffusion, thus reweighting by the eigenvalues of the diffusive transition matrix also allows tracking of increasing uncertainty.

Utilising the fixed eigenbasis (eq. 4) and the respective eigenvalues (eq. 8) for arbitrary transition structures, we can make efficient prediction for the distribution of future state occupancy with respect to arbitrary action (see Figs. 2B-C). Adding translation to the translation-invariant transition matrix does not change the set of eigenvectors - allowing one set of eigenvectors (Fourier modes) to support prediction for actions in all directions (or plain diffusion), hence prediction of effects of directed actions can be efficiently generalised across environments.

We define a "sense of direction", as the angle of the transitions (or the linear combinations of the available actions in a non-spatial setting) that maximise the future probability of reaching the target state given an initial state, which is modelled by the SR matrix.

$$\theta^* = \arg\max_{\theta} \sum_j \frac{\exp[2\pi i(\mathbf{s}_G - \mathbf{s}_0) \cdot \mathbf{k}_j]}{1 - \gamma D_j \exp[2\pi i \mathbf{v}_\theta \cdot \mathbf{k}_j]} \tag{9}$$

where γ is the discounting factor, $D_j, j = 1, \cdots, LW$ are the eigenvalues for the symmetric diffusion transition matrix, $\mathbf{k}_j, j = 1, \ldots, LW$ are the wavevectors for the j-th Fourier components, $\mathbf{s}_0, \mathbf{s}_G$ are the coordinates of the start and goal states, and $\mathbf{v}_\theta = (v\cos(\theta), v\sin(\theta))$ represents the velocity (with speed v and head direction θ). We see that the "sense of direction" supports generalisation of predictions of effects of actions across all environments with the same translation-invariant transition structure, i.e., such predicted effects ignore any differences in local structure. See Appendix B for the derivation of eq. 9. Note that here we assume that the goal state $\mathbf{s}_G$ is known a priori, e.g., we consider a problem where the animal is navigating towards a previously visited location.

We propose that a computational role for the neural grid codes: generating a "sense of direction" (capturing the transition structure of the state space, ignoring the obstacles and boundaries) that reflects a global sense of orientation which allows generalisation to completely new environments, via utilising a Fourier basis for a larger toroidal pseudo space.

The proposed model can be applied to flexible planning under arbitrary drift velocity as demonstrated in Fig. 3 (A-E). An agent is trying to navigate towards a goal state in a windy grid world. The navigation is performed by following the ascending "gradient" of the SR for occupancy of the target state (the resolvent metric, eq. 2). The SR computed from the transition matrix including the effects of diffusion and wind (Fig. 3 B) based on our analysis (eq. 8) leads straight to the target (Fig. 3 C).

Given the analytical expression of the SR (eq. 2), we could efficiently adjust the SR matrix to accommodate local changes in the state space, e.g., insertion of a barrier, using the Woodbury inversion formula to update the parts of the SR matrix affected by the local obstacles (see Appendix A.3 for derivations); and again in this case, the agent correctly adjusts for the wind as well as taking the shortest path around the inserted wall (Fig. 3 D-E).

We note, however, that the proposed model is also able to solve tasks without periodic boundary conditions, by considering the original task state space, S_0, being embedded into a larger, periodically bounded pseudo state space S_p, at least twice as large in each dimension as S_0 (Fig. 3 F). We again follow the previous procedures, utilising the Fourier modes, this time computed on S_p, to perform predictions in S_0 (Fig. 3 F-G), and the performance is unaffected. Note that under such formulation, the underlying transition structures can be applied to environments with both periodic and non-periodic boundary conditions - allowing sense of direction planning in either case.

We can also use our model for path integration (Fig. 3 H, see Section 4) in S_0, by taking velocity inputs (from any path in the grid world) to update the state occupancy distribution (eq. 8). We observe that the path integration performance is strongly sensitive with respect to self-motion noise, with perfect path integration when the uncertainty is low up to 1000 time steps (the discretisation of state space means that uncertainty below 0.075 has no effect), and monotonically increasing path integration error when the uncertainty is higher, where the uncertainty represents the degree of diffusion caused by self-motion noise, in addition to translation.

The derived analytical expression for the sense of direction can be retrieved via a lookup table when the state space is small and discrete, whereas in large or continuous state spaces, it can be found either via optimisation algorithms (the problem is non-convex by its periodic nature, but there exists a unique global optimum), or by a neural network modelling the non-linear mapping between the quantity to be optimised over in eq. 9 and the sense of direction, see Bush et al. [12] for neural network approaches to finding goal directions from grid representations.

To demonstrate the utility of state representation by eigenvectors weighted by action-dependent eigenvalues, we adapted a deep-Q network (DQN, Mnih et al. [32]) to allow it to learn such representations (Fig. 4A), applied to the CartPole task (Barto et al. [3]). The basic DQN takes 4 continuous state variables as input (cart position and velocity, pole angle and angular velocity) and learns to output Q values for actions 0 or 1 (cart movement left or right) via a temporal difference learning rule . The gc-DQN has 2 additional networks: the first takes as inputs the n low-frequency Fourier modes over the discretised state space (with $N = 8^4$ bins) and has $n_{actions} = 2$ outputs to allow representation for each action given the input Fourier modes; the second takes as inputs the state variables and action and has $2n$ outputs which serve as an additional action-dependent multiplier to the connection weights for the hidden layer of the first network. The outputs of the first network and the outputs of the standard DQN are then fully connected to an output layer to learn the updated Q values. We also tested versions of these networks with deep Dyna agents (Peng et al. [34]). See Appendix. E for details. Preliminary results in fig. 4(B) show that the gc-DQN and deep gc-Dyna-Q greatly accelerates learning comparing to the baseline agents, with relatively minor increase in the model complexity and computational costs. The results support our hypothesis that Fourier eigenvectors and the possibility of weighting them by action-specific values aids prediction (in this case, of future value).

The focus of this paper is proposing a theoretical framework for state representation, prediction, planning and path integration via grid-like eigenvectors and action-dependent eigenvalues. The neural network implementation above is only a preliminary attempt at combining this approach with deep learning (see also Mahadevan and Maggioni [30]).

4 A UNIFYING FRAMEWORK FOR MODELS OF GRID CELL FIRING

Our focus so far has been on proposing a flexible and efficient extension of the prediction models of grid cells (Baram et al. [2]; Dordek et al. [16]; Stachenfeld et al. [39]) to arbitrary directed transitions.

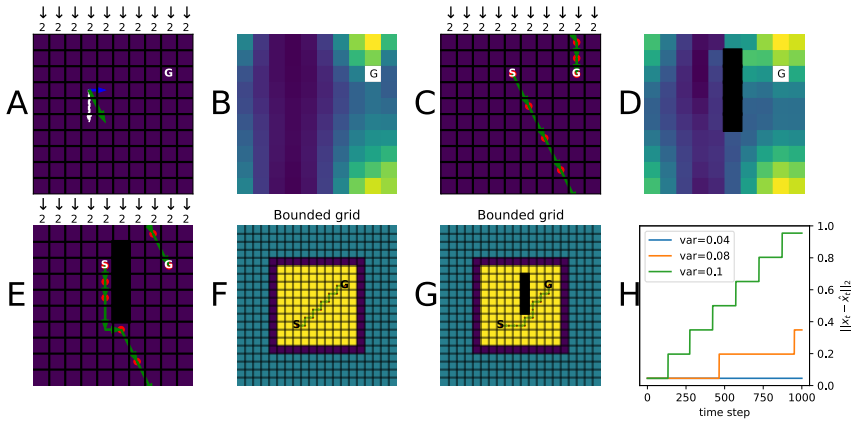

Figure 3: **Application to spatial navigation in grid worlds.** **A**: The 10×10 windy grid world task environment, with toroidal boundary conditions and a constant external force causing two units of displacement southward acting on every state (white arrow: wind; blue arrow: one-step action rightward; green arrow: actual displacement, G: goal state);**B**: Estimated SR (using eq. 2 and eq. 8) under diffusion plus the wind effect (color indicates the strength of future probabilities of occupying the target states); **C**: Path following the diffusion SR plus the wind effect; **D**: Updated diffusion SR plus wind given the insertion of a barrier (dark blocks); **E**: Path following the updated SR; **F**: Navigation in the task space (S_0, yellow) with boundaries (magenta) embedded in a pseudo state space (S_p, blue), utilising the Fourier modes computed from S_p; **G**: Navigation in S_0 with inserted local obstacles (black); **H**: Path integration error over timesteps under different levels of diffusion.

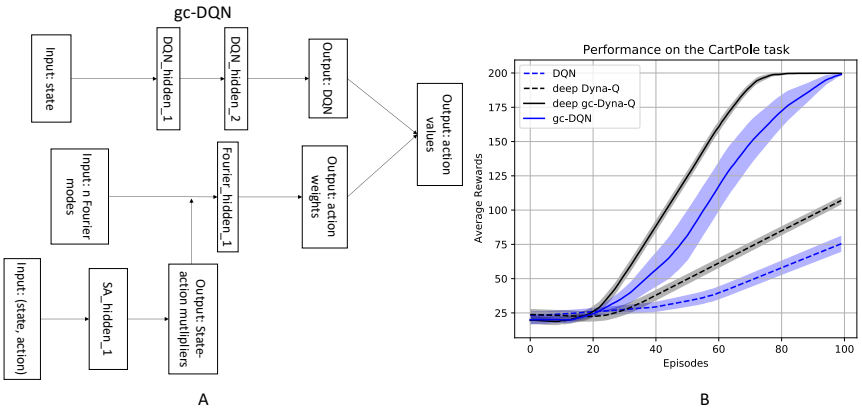

Figure 4: **Neural network implementation of the proposed grid cell model.** **A**: Schematic illustration of the gc-DQN agent. **B**: Comparison between the gc-DQN agents and the baseline standard DQN agents in the CartPole task (Barto et al. [3]). Plots of average reward in each episode of the CartPole task against the episode number. See Appendix. E for details of the simulation.

There exists many other computational models of grid cells that do not rely on inputs from the place cells, such as the continuous attractor network (CAN) models which models the emergence of grid-like patterns in a recurrently connected networks with pre-specified weights via performing path integration (Fuhs and Touretzky [19]; Burak and Fiete [6]; Corneil and Gerstner [14]), and the oscillatory interference (OI) models which models the formation of grid-like patterns via coincidence detection of velocity-dependent inputs during path integration (Welday et al. [42]; Burgess et al. [9]; Burgess [8]). Here we build upon the previous work of unifying the normative and path integration models of grid cells firing by Sorscher et al. [38], extending their discussions and showed the equivalence of our generalised prediction framework with models of path integration by grid cells based on continuous attractor dynamics or oscillatory interference.

4.1 RELATION TO CONTINUOUS ATTRACTOR NETWORK MODELS OF GRID CELLS

Several unified frameworks for grid cell models have been proposed, one of the most prominent by Sorscher et al. [38], where they proved the equivalence between the maximisation of a spatial representation objective function under the prediction models and the pattern formation dynamics of the CANs under the models of path integration. Their analysis, however, does not explicitly address the equivalence between the two classes of models when there is non-zero velocity inputs, corresponding to non-trivial asymmetric velocity-dependent connectivities in the CANs (Fuhs and Touretzky [19]; Burak and Fiete [6]), which is the key element for performing path integration. We explicitly address such equivalence utilising our framework based one Fourier analysis. Assuming that a grid cell is formed via linear combination of selected Fourier modes (e.g., 6 Fourier modes at $\frac{\pi}{3}$ degrees increments with same frequencies):

$$g = \sum_{j=1}^{G} w_j f_j \tag{10}$$

where f_j's are the selected Fourier modes with weights w_j. Inserting the single-cell representation into eqs. (14, 15) from [38] we have the equations governing the dynamics of the grid cells under the normative and mechanistic path integration models respectively as following.

$$\frac{dg}{dt} = \begin{cases} -\alpha g + \gamma(\sum_{j=1}^{G} \lambda_j w_j f_j) + \mu, & g > 0 \\ -\alpha g + \sigma(\gamma(\sum_{j=1}^{G} \lambda_j w_j f_j) + \mu), & g = 0 \end{cases} \tag{11}$$

$$\tau \frac{dg}{dt} = -g + \sigma(\mathbf{W}g + b(v)) \tag{12}$$

where λ_j is the corresponding eigenvalue of f_j with respect to the (symmetric) transition matrix, $\mathbf{T}^0$, $\sigma(\cdot)$ is the Heaviside function, $\mathbf{W}$ is the recurrent connectivity matrix, $b(v)$ is the velocity-dependent feedforward input to the grid cell under continuous attractor model which involves a constant baseline term and a velocity dependent term (Burak and Fiete [6]), α is the multiplicative constant, γ and μ are the multiplicative constant corresponding to the additive penalty terms in the (single-neuron) Lagrangian under the normative models:

$$\mathcal{L} = g^T \mathbf{T}^0 g - \gamma g^T g + \mu \mathbb{1}^T g \tag{13}$$

Now suppose the agent is moving with velocity v, which causes changes in the transition structure, given our previous formulation based on Fourier analysis (eq. 8) and using the first-order Taylor expansion around 0 for approximating the exponential, eq. 11 and eq. 12 becomes equivalent, by identifying the corresponding terms τ, W and $b(v)$, see Appendix B for detailed proof of equivalence. The resulting expression for $b(v)$ is directly related to the Fourier modes inputs, hence it correlates with the underlying grid codes with the baseline μ, hence is equivalent to the asymmetric velocity-dependent parts of the recurrent connectivity matrix in the CAN.

We stress that the proposed model is on the theory-level, instead of the implementation level within which CAN models lies, hence we do not assume any specific neural network structure in the current analysis. What we showed in this section is the equivalence between the differential equation governing the dynamics of updates in the CAN models and the gradients of the Lagrangian of the predictions model under non-zero velocities, based on the aforementioned assumption that the grid cell firing is a linear combination of the Fourier modes (Sec. 3), hence showing that under non-stationary transition dynamics, the CAN models and the prediction models should also yield identical updates to the grid cells (up to scaling factors).

4.2 RELATION TO OSCILLATORY INTERFERENCE MODELS OF GRID CELLS

For the oscillatory interference models of grid cells, path integration is achieved by tracking the phases of the velocity controlled oscillators (VCOs), which encode movement speed and direction by variations in burst firing frequency (Burgess et al. [9]; Burgess [8]). The VCOs generate grid-like firing patterns via coincidence detection by the grid cells (Hasselmo [23]; Welday et al. [42]). The variation of frequency with velocity produces a phase code for displacement, enabling the modelled grid cells to perform path integration. Namely, VCOs change their frequencies relative to the baseline according to the projection of current velocity, $v(t)$, onto the VCO's "preferred direction", $\mathbf{d}$:

$$f_a(t) = f_b(t) + \beta \mathbf{v}(t) \cdot \mathbf{d} \tag{14}$$

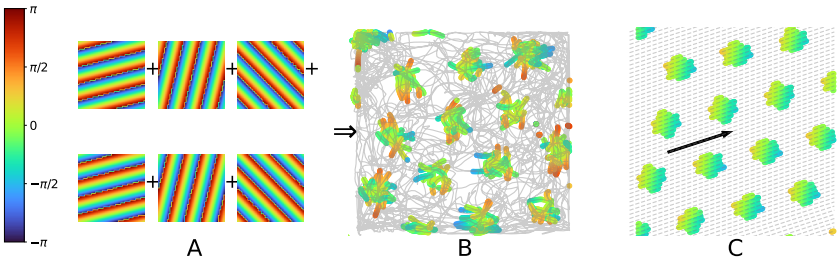

Figure 5: **Generated grid-like patterns given our model**. **A**: 6 input Fourier modes with the same wavelength (complex phases shown); **B**: Simulated grid cell firing given a real rat trajectory (gray line) using coincidence detection and baseline modulation given the 6 Fourier modes inputs in A, each spike is represented by a colored scatter with the color indicates the corresponding "theta" phase at the spiking time; **C**: Simulated grid cell firing fields given multiple runs in the same direction (black arrow) showing theta phase precession.

where β is a positive constant, and $f_b(t)$ is the baseline frequency (the $4-11$Hz EEG theta rhythm). It follows that the VCOs perform linear path integration since the relative phase between VCO's frequency and baseline at time t, $\phi_{ab}(t) = \phi_a(t) - \phi_b(t)$, is proportional to the agent's displacement in the preferred direction:

$$\phi_{ab}(t) - \phi_{ab}(0) = \int_0^t 2\pi[f_a(\tau) - f_b(\tau)]d\tau = 2\pi\beta[\mathbf{x}_t - \mathbf{x}_0] \cdot \mathbf{d} \tag{15}$$

where $\mathbf{x}(t)$ is the agent's location at time t. The interference of VCOs whose preferred directions differ by multiples of $\pi/3$ generate triangular grid-like patterns, provides an explanation of theta phase precession in grid cells in which firing phase relative to the theta rhythm encodes distance travelled through the field (Hafting et al. [22]; Burgess [8]) and complements the attractor dynamics given by symmetrical connections between grid cells (Bush and Burgess [10]). We note that the major challenge to the OI models posed in the neurophysiological community, the main experimental results against the OI models (that bats and humans do not have a reliable theta frequency oscillations) has recently been resolved: the required phase coding (theta phase precession) can occur relative to a variable baseline frequency (Bush and Burgess [11]) and has now been found in both bats and humans (Eliav et al. [18]; Qasim et al. [36]).

We simulated the firing of a grid cell synthesised from 6 Fourier modes inputs (Fig. 5 A-B), each firing a spike at its complex phase in the current state, as a leaky integrate and fire neuron performing coincidence detection using a real trajectory of a rat exploring a 50cm $\times$ 50cm box. At each time step (corresponding to one theta cycle), the phase of each Fourier mode is updated according to eq. 8 given the current velocity, and fires a spike at this phase if it is within the interval $[-\pi/4, \pi/4]$ (modelling modulation by the baseline theta oscillation). The grid cell fires a spike at the current location if the integrated inputs reach a threshold. Note that we could also simulate a set of grid cells, with different offsets (depending on the initial phases of the Fourier modes) and different scales and orientations (depending on the choice of Fourier modes), such that the set of grid cells, like the Fourier modes, comprises a basis for the state space and do so on the basis of path integration (for which environmental inputs are also required to prevent error accumulation [8; 7]).

Grid cells show "precession" in their firing phase relative to theta as the animal moves through the firing field (signaling distance travelled within the field) (Hafting et al. [22]; Jeewajee et al. [24]; Climer et al. [13]), which is captured by our model, similar to the OI model. The Fourier modes whose wavevectors that are aligned with the current direction of translation advance in phase as the movement progresses. Phase precession results from assuming that the Fourier modes aligned with movement direction are dominant influence on grid cell firing (c.f. those aligned to the reverse direction). The baseline "theta frequency" corresponds to the mean rate of change of phase of all Fourier components and so could vary (for noise reduction, see Burgess and Burgess [7]; Burgess [8], without precluding phase coding (Eliav et al. [18]; Bush and Burgess [11])). By simulating random straight runs, we can see clear late-to-early phase precession (Fig. 5 C), analogous to that observed in grid cells (Hafting et al. [22]) and the OI models (Burgess [8]).

We thus conclude that the OI model and our model perform path integration or prediction in the same way: the OI model changes the phase of each VCO relative to baseline corresponding to the component of translation along the VCO's preferred direction, and this is exactly analogous to the complex phases of the Fourier modes being updated to reflect transitions along its wavevectors (multiplication by corresponding eigenvalues, eq. 8), and the phase changes with transitions perform path integration in the same way as the OI models.

5 Discussion

Understanding how different actions affect the agent's state across environments is essential for generalisation. Existing models are capable of such prediction for tasks with an invariant symmetric transition matrix. We propose a computational role for the grid cell firing: generating generalisable "sense of direction" (eq. 9) that captures the underlying transition structure across different environments (ignoring the specific obstacles in each environment), so that they can generalise prediction and path integration under directed transitions in new environments (see also Whittington et al. [43]). Here we generalised the prediction model to provide a mathematical framework for understanding this function, and showed how it corresponds to other computational models of grid cells that are based on path integration, and provide a functional explanation for theta phase precession.

Having obtained the approximate transition structure of the environment, the proposed prediction model can perform the sense-of-direction planning in new environments (i.e., finding combinations of actions that most likely leads to the target state by ignoring local obstacles). We make the prediction that there exists other (e.g., fronto-parietal) areas of the brain that are responsible for detecting and avoiding obstacles following the overall direction provided by the grid cells [17; 29].

The current model applies to translation-invariant transition structures, and use of the Fourier shift theorem to calculate eigenvalues also assumes a Euclidean state space. We demonstrated ways of generalising planning to bounded or locally non-translation invariant transition structures in Section 3. We note that machine learning methods based on a similar premise (creating a single representation to support planning via multiple different actions) might work even when the transitions are not strictly translation invariant (e.g., family trees, see Whittington et al. [43]).

A number of question and predictions can be raised from the proposed model. If there exists a basis of neurons with Fourier-mode-like firing patterns acting as inputs to cells in the entorhinal cortex, then the set of grid firing patterns only comprises a small proportion of the entire set of firing patterns that can be synthesised, how the composition corresponds to the heterogeneity of firing patterns observed in the entorhinal cortex is yet to be understood, but experimental results have indicated the existence of periodic non-grid cells in the entorhinal cortex that also resemble combinations of small number of Fourier modes (Krupic et al. [28]). We predict the use of the same set of grid cells (superpositions of Fourier modes inputs) for indicating direction to goal across different environments with the same global orientation – i.e. different boxes/mazes with the same distal orienting cue, hence supporting generalisability and faster learning. We also predict the existence of topology-dependent grid cells (that exhibits distortions in the firing patterns with respect to changes in the environmental topology) that complements the Euclidean-based grid cells proposed here for spatial navigation in dealing with local changes in the environmental topology. Finally, the proposed model predicts future state occupancy from the transition matrix, future works could look at the reversed direction: vector navigation, i.e., inferring the translation between two locations given the phase codes for each (see Bush et al. [12]; corresponding to discriminative and generative models).

We showed that by giving standard DQN agents the ability to represent the current state as eigenvectors of the state space weighted by action-specific values, the resulting gc-DQN agent accelerates learning in the CartPole experiment (fig. 4). Future work will involve more rigorous study of the neural network implementation of the proposed grid cell model, which we hope will bring insights into the combination of our approach with deep learning. Our approach offers potential generalisation to non-spatial tasks such as transitive inference (Von Fersen et al. [41], see Appendix D) and, in combination with DQN agents, improves performance in the CartPole problem which is both non-spatial and contains actions that are not strictly translation-invariant. A future direction in generalising the current model to non-spatial tasks will be to consider Fourier analysis on groups of operators utilising group-theoretic knowledge (Kondor and Trivedi [26]; Gao et al. [20]).

ACKNOWLEDGEMENTS

The authors thanks support from various funding resources. The authors declare no competing financial interests.

REFERENCES

[1] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. URL `https://www.tensorflow.org/`. Software available from tensorflow.org.

[2] A. B. Baram, T. H. Muller, J. C. Whittington, and T. E. Behrens. Intuitive planning: global navigation through cognitive maps based on grid-like codes. *bioRxiv*, page 421461, 2018.

[3] A. G. Barto, R. S. Sutton, and C. W. Anderson. Neuronlike adaptive elements that can solve difficult learning control problems. *IEEE transactions on systems, man, and cybernetics*, (5): 834–846, 1983.

[4] R. N. Bracewell. *The Fourier transform and its applications*, volume 31999. McGraw-Hill New York, 1986.

[5] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba. Openai gym. *arXiv preprint arXiv:1606.01540*, 2016.

[6] Y. Burak and I. R. Fiete. Accurate path integration in continuous attractor network models of grid cells. *PLoS computational biology*, 5(2), 2009.

[7] C. P. Burgess and N. Burgess. Controlling phase noise in oscillatory interference models of grid cell firing. *Journal of Neuroscience*, 34(18):6224–6232, 2014.

[8] N. Burgess. Grid cells and theta as oscillatory interference: theory and predictions. *Hippocampus*, 18(12):1157–1174, 2008.

[9] N. Burgess, C. Barry, and J. O'keefe. An oscillatory interference model of grid cell firing. *Hippocampus*, 17(9):801–812, 2007.

[10] D. Bush and N. Burgess. A hybrid oscillatory interference/continuous attractor network model of grid cell firing. *Journal of Neuroscience*, 34(14):5065–5079, 2014.

[11] D. Bush and N. Burgess. Advantages and detection of phase coding in the absence of rhythmicity. *Hippocampus*, 2020.

[12] D. Bush, C. Barry, D. Manson, and N. Burgess. Using grid cells for navigation. *Neuron*, 87(3): 507–520, 2015.

[13] J. R. Climer, E. L. Newman, and H. M. E. Phase coding by grid cells in unconstrained environments: two-dimensional phase precession. *European Journal of Neuroscience*, 38(4): 2526–2541, 2013.

[14] D. S. Corneil and W. Gerstner. Attractor network dynamics enable preplay and rapid path planning in maze–like environments. In *Advances in neural information processing systems*, pages 1684–1692, 2015.

[15] P. Dayan. Improving generalization for temporal difference learning: The successor representation. *Neural Computation*, 5(4):613–624, 1993.

[16] Y. Dordek, D. Soudry, R. Meir, and D. Derdikman. Extracting grid cell characteristics from place cell inputs using non-negative principal component analysis. *Elife*, 5:e10094, 2016.

[17] V. Edvardsen, A. Bicanski, and N. Burgess. Navigating with grid and place cells in cluttered environments. *Hippocampus*, 2019.

[18] T. Eliav, M. Geva-Sagiv, M. M. Yartsev, A. Finkelstein, A. Rubin, L. Las, and N. Ulanovsky. Nonoscillatory phase coding and synchronization in the bat hippocampal formation. *Cell*, 175 (4):1119–1130, 2018.

[19] M. C. Fuhs and D. S. Touretzky. A spin glass model of path integration in rat medial entorhinal cortex. *Journal of Neuroscience*, 26(16):4266–4276, 2006.

[20] R. Gao, J. Xie, S.-C. Zhu, and Y. N. Wu. A representational model of grid cells based on matrix lie algebras. *arXiv preprint arXiv:2006.10259*, 2020.

[21] T. Hafting, M. Fyhn, S. Molden, M.-B. Moser, and E. I. Moser. Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436(7052):801–806, 2005.

[22] T. Hafting, M. Fyhn, T. Bonnevie, M.-B. Moser, and E. I. Moser. Hippocampus-independent phase precession in entorhinal grid cells. *Nature*, 453(7199):1248–1252, 2008.

[23] M. E. Hasselmo. Grid cell mechanisms and function: contributions of entorhinal persistent spiking and phase resetting. *Hippocampus*, 18(12):1213–1229, 2008.

[24] A. Jeewajee, C. Barry, V. Douchamps, D. Manson, C. Lever, and B. N. Theta phase precession of grid and place cell firing in open environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1635):20120532, 2014.

[25] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[26] R. Kondor and S. Trivedi. On the generalization of equivariance and convolution in neural networks to the action of compact groups. *arXiv preprint arXiv:1802.03690*, 2018.

[27] E. Kropff and A. Treves. The emergence of grid cells: Intelligent design or just adaptation? *Hippocampus*, 18(12):1256–1269, 2008.

[28] J. Krupic, N. Burgess, and J. O'Keefe. Neural representations of location composed of spatially periodic bands. *Science*, 337(6096):853–857, 2012.

[29] E. A. Maguire, N. Burgess, J. G. Donnett, R. S. Frackowiak, C. D. Frith, and J. O'Keefe. Knowing where and getting there: a human navigation network. *Science*, 280(5365):921–924, 1998.

[30] S. Mahadevan and M. Maggioni. Proto-value functions: A laplacian framework for learning representation and control in markov decision processes. *Journal of Machine Learning Research*, 8:2169–2231, 2007.

[31] B. L. McNaughton, F. P. Battaglia, O. Jensen, E. I. Moser, and M.-B. Moser. Path integration and the neural basis of the'cognitive map'. *Nature Reviews Neuroscience*, 7(8):663–678, 2006.

[32] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*, 2013.

[33] J. O'Keefe and J. Dostrovsky. The hippocampus as a spatial map: preliminary evidence from unit activity in the freely-moving rat. *Brain research*, 1971.

[34] B. Peng, X. Li, J. Gao, J. Liu, K.-F. Wong, and S.-Y. Su. Deep dyna-q: Integrating planning for task-completion dialogue policy learning. *arXiv preprint arXiv:1801.06176*, 2018.

[35] P. Piray and N. D. Daw. A common model explaining flexible decision making, grid fields and cognitive control. *bioRxiv*, page 856849, 2019.

[36] S. E. Qasim, I. Fried, and J. Jacobs. Phase precession in the human hippocampus and entorhinal cortex. *bioRxiv*, 2020.

[37] K. S. Riedel. A sherman–morrison–woodbury identity for rank augmenting matrices with application to centering. *SIAM Journal on Matrix Analysis and Applications*, 13(2):659–662, 1992.

[38] B. Sorscher, G. Mel, S. Ganguli, and S. Ocko. A unified theory for the origin of grid cells through the lens of pattern formation. In *Advances in Neural Information Processing Systems*, pages 10003–10013, 2019.

[39] K. L. Stachenfeld, M. M. Botvinick, and S. J. Gershman. The hippocampus as a predictive map. *Nature neuroscience*, 20(11):1643, 2017.

[40] E. C. Tolman. Cognitive maps in rats and men. *Psychological review*, 55(4):189, 1948.

[41] L. Von Fersen, C. D. Wynne, J. D. Delius, and J. E. Staddon. Transitive inference formation in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 17(3):334, 1991.

[42] A. C. Welday, I. G. Shlifer, M. L. Bloom, K. Zhang, and H. T. Blair. Cosine directional tuning of theta cell burst frequencies: evidence for spatial coding by oscillatory interference. *Journal of Neuroscience*, 31(45):16157–16176, 2011.

[43] J. C. Whittington, T. H. Muller, S. Mark, G. Chen, C. Barry, N. Burgess, and T. E. Behrens. The tolman-eichenbaum machine: Unifying space and relational memory through generalization in the hippocampal formation. *Cell*, 2020.