Rapid approximation of successor representations with STDP and theta phase precession

Summary The successor representation (SR) is a promising candidate principle for hippocampal function. Theory proposes that each place cell encodes the expected state occupancy of its target location in the near future. This framework has desirable consequences on the generalisability and efficiency of reinforcement learning algorithms operating over these representations and is supported by behavioural and electrophysiological evidence. However, it is unclear how the SR might be learnt in the brain. Temporal difference learning, commonly used to learn SRs in artificial agents, is not known to be implemented in hippocampal networks. Instead, we demonstrate that spike-timing dependent plasticity (STDP), a modified form of Hebbian learning, acting on temporally compressed trajectories known as "theta sweeps", is sufficient to rapidly learn a useful approximation to the SR. Our model is biologically plausible: it maps onto validated aspects of hippocampal circuitry, it uses spiking neurons modulated by theta-band oscillations, diffuse and overlapping place cell-like states and experimentally matched parameters. It explains substantial variance in the true successor matrix and gives rise to place cells demonstrating key experimentally observed phenomena associated with the SR including policy-dependent backwards expansion on a 1D track and elongation near walls in 2D. In our model, larger place cells encode longer timescale SRs. We shed insight on the observed topographical ordering of place cell size down the dorsal-ventral axis by showing this is necessary to prevent the detrimental mixing of these timescales.

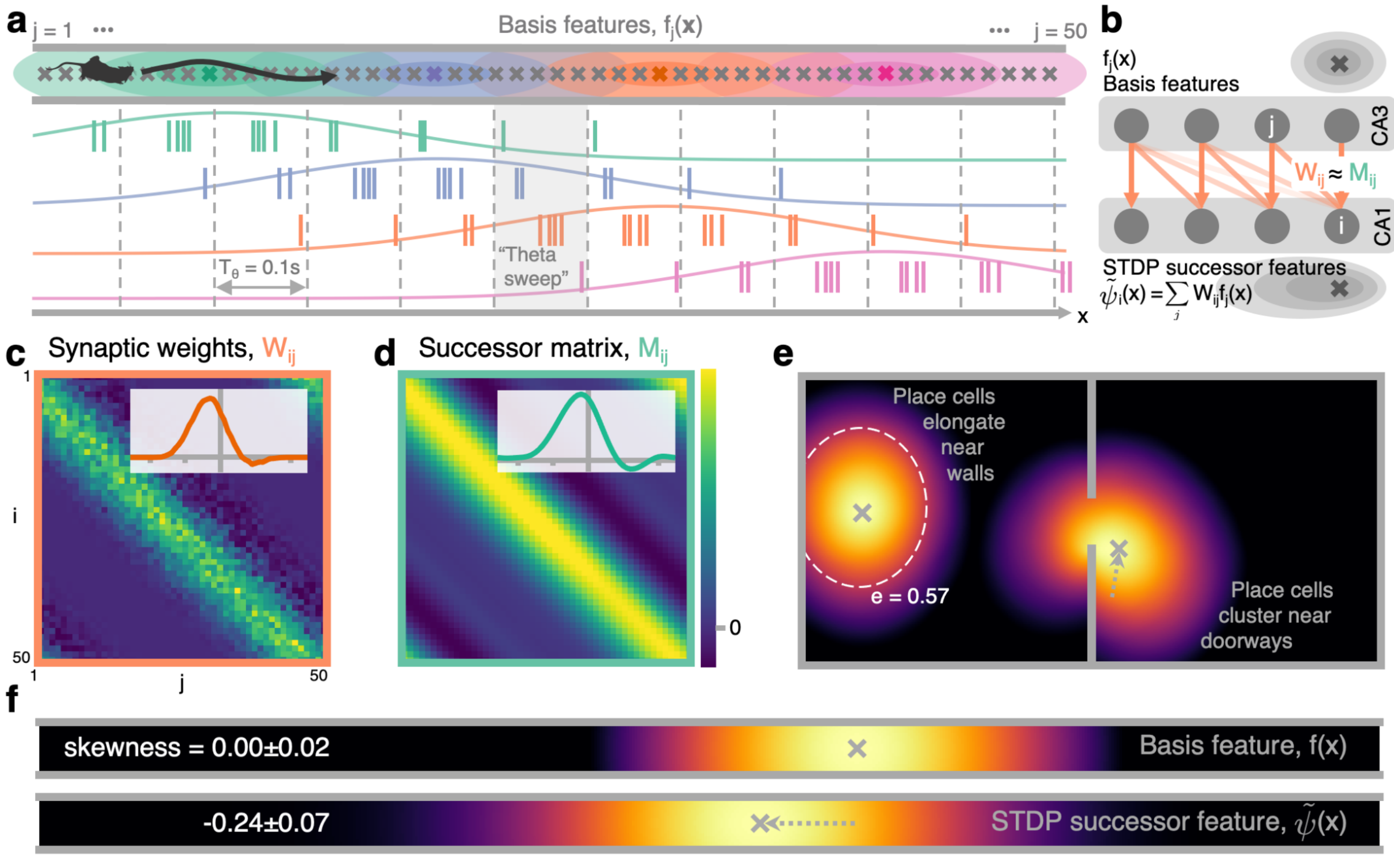

Figure 1: *Learning the SR with STDP and theta phase precession.* ***(a)*** *Basis features phase precess as the agent crosses their receptive field, generating compressed "theta sweeps".* ***(b)*** *Circuit diagram: fixed CA3 basis features drive CA1 cells which learn to approximate the successor features. Synaptic weights between these populations, W_{ij}, are updated according to a Hebbian STDP learning rule.* ***(c)*** *The synaptic weight matrix (STDP learning).* ***(d)*** *The successor matrix (TD learning). Insets show averages over aligned rows.* ***(e)*** *In 2D, CA1 successor features (aka place cells) elongate near walls (e = eccentricity) and cluster near doorways.* ***(f)*** *In 1D, features shift and skew against the motion direction.*

Details In our model, an agent has a set of basis feature cells (Figures 1a & b, upper) analogous to CA3 place cells. A given feature fires at an earlier phase in the theta cycle (10 Hz) as it is traversed by the agent; this corresponds to an observed phenomenon known as phase precession and generates temporally compressed "theta sweeps" of the local trajectory[1] (Figure 1a, lower grey box).

Our CA3 basis features project linearly onto an equivalent set of CA1 cells - which we call 'STDP successor features' - through the synaptic weight matrix W_{ij}. Initially these are identical to the basis features ($W_{ij} = \delta_{ij}$). As the agent explores, W_{ij} is updated according to the Hebbian STDP learning rule[2] shown in Equation 1, where $T_{i/j}(t)$ are the attenuated spike trace and $t_{i/j}$'s are spike times and for cells i and j in CA1 and CA3 respectively (Figure 1b, lower).

$$\frac{d\mathsf{W}_{ij}(t)}{dt} = \eta\left[a_+ \sum_{t_i} \delta(t-t_i)T_j(t) - a_- \sum_{t_j} \delta(t-t_j)T_i(t)\right] \quad \textit{Hebbian STDP, (Equation 1)}$$

For an experiment on a 1D track with circular boundary conditions and a movement policy biased left-to-right (e.g. Figure 1a) the learnt synaptic weight matrix is shown in Figure 1c.

Simultaneously, for comparison, we learn the true 'TD successor features'. These encode the expected discounted future firing of a basis feature**,** in a temporally and spatially continuous setting (Equation 2). τ is the discount horizon, closely related to the discount parameter γ used in the discrete setting. By imposing the constraint that the TD successor features are a linear combination of basis features (Equation 3) we define the successor *matrix*, M_{ij}, which can therefore be directly compared to W_{ij}. We derive a novel TD-based online update rule which optimises M_{ij} directly, simultaneously learning the features.

$$\psi_i(\mathbf{x}) = \mathbb{E}\left[\int_t^\infty e^{-\frac{t'-t}{\tau}} f_i\big(\mathbf{x}(t')\big)dt' \;\middle|\; \mathbf{x}(t) = \mathbf{x}\right] \quad \textit{The TD successor feature, (Equation 2)}$$

$$\psi_i(\mathbf{x}) = \sum_j \mathsf{M}_{ij} f_j(\mathbf{x}) \quad \textit{The successor matrix, (Equation 3)}$$

After learning, the successor and synaptic weight matrices are quantitatively similar ($R^2 = 0.87$). Both are dominated by a band of positive weights left of the diagonal and negative weights on the right (averaging over aligned rows shows this clearly; see Figure 1c & d insets). This reflects the left-to-right behavioural bias: from a TD perspective, basis features nearby and to the right are "predicted" to fire in the immediate future. From an STDP perspective these features will typically fire slightly later in each theta cycle, causing binding. Whereas the overall shape, shift and skew of M_{ij} are well predicted by W_{ij}, long range bindings are not. This is because our STDP model only binds features with fixed overlapping receptive fields. We believe this can be addressed in future work by implementing a recursive model allowing successor features to 'bootstrap' their own predictions, similar to recent models[3,4].

Downstream STDP successor features closely approximate the TD successor features ($R^2 = 0.97$, Figure 1b lower), reflecting the fact that $W_{ij} \approx M_{ij}$. They show behaviourally biased skewing against the direction of travel (Figure 1f). This is a hallmark of classical SR theory[5] which has been experimentally observed in CA1 place cells[6]. Our model also works in 2D, capturing experimental results including the elongation of place cells near walls and boundaries (KS test, $P<0.001$) and the clustering of place cells near doorways ($P<0.001$) as shown in Figure 1e.

We find that when the policy is heavily biased (i.e. non-diffusive) theta sweeps are essential since they force the temporal ordering of spikes from overlapping features to become stereotyped and to reliably reflect the order in which the features are encountered under the policy. Learning is also rapid and has a higher signal-to-noise ratio when theta sweeps are present.

It has been hypothesized that for flexible navigation SRs of multiple timescales must be stored[7]. In our model there is a direct relationship between the place cell-like basis feature size and the timescale of the encoded SR. We show that without strict topographical separation of place cells according to size, short timescale SRs are overwritten due to binding with larger place cells. This provides a functional explanation for ordering of place cell size along the dorsal-ventral axis.

In conclusion, we show modified Hebbian plasticity is sufficient to learn a working approximation to the SR without the need for more complex TD learning. We believe this result has more general applications for learning predictive encodings in other brain regions where compressed behavioural trajectories are known to exist.

References **[1] O'Keefe, J. and Recce, M. L. (1993)**. Phase relationship between hippocampal place units and the EEG theta rhythm. Hippocampus, 3(3):317–330 **[2] Bi, Guo-qiang and Poo, Mu-ming (1998).** Synaptic modifications in cultured hippocampal neurons. 18(24):10464– 10472 **[3] Brea, J., Gaál, A. T., Urbanczik, R., and Senn, W. (2016).** Prospective Coding by Spiking Neurons. PLOS Computational Biology, 12(6):e1005003. **[4] Bono, J., Zannone, S., Pedrosa, V., and Clopath, C. (2021).** Learning predictive cognitive maps with spiking neurons during behaviour and replays **[5] Mehta, M., M.C. Quirk, and Wilson, M. (2000).** Experience-Dependent Asymmetric Shape of Hippocampal Receptive Fields. Neuron, 25:707–715 **[6] Stachenfeld, K. L., Botvinick, M. M., and Gershman, S. J. (2017).** The hippocampus as a predictive map. Nature Neuroscience, 20(11):1643–1653 **[7] Momennejad, I. and Howard, M. W. (2018).** Predicting the Future with Multi-scale Successor Representations, preprint