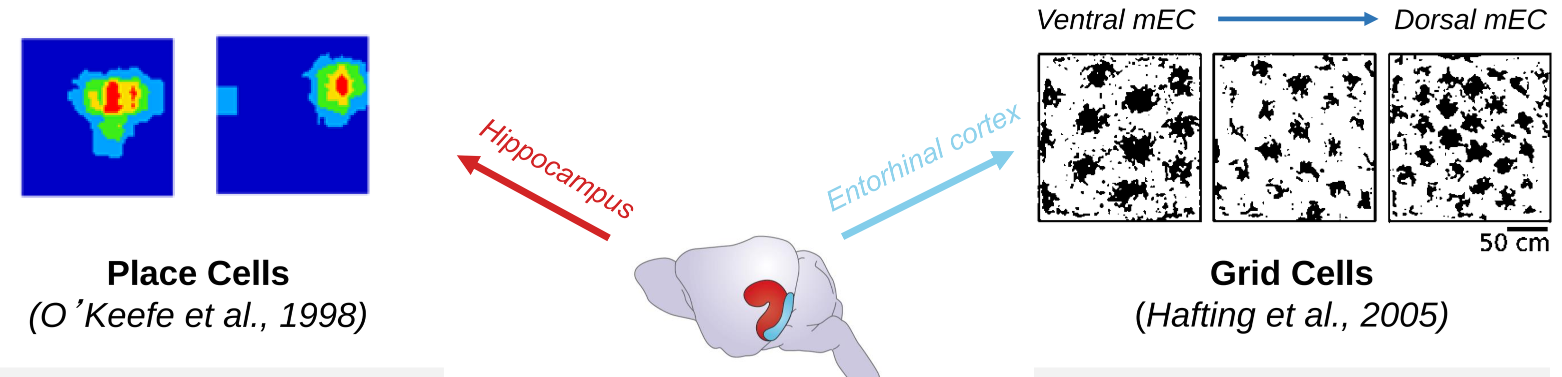
Coordinated hippocampal-entorhinal replay as structural inference

Talfan Evans[1,2,3] and Neil Burgess[1,2]

Institute of Cognitive Neuroscience, UCL[1]; Institute of Neurology, UCL[2]; CoMPLEx, UCL[3]

Overview: Spatial representations in the brain

1. Place cells in the hippocampus (HPC) respond to sensory stimuli
2. Grid cells in entorhinal cortex (mEC) encode location in *metric* space
3. Spatial inference is achieved by message passing in the HPC-mEC system
4. Message scheduling produces coordinated offline 'replay' of place and grid cells



Place Cells
(O'Keefe et al., 1998)

Respond to sensory features

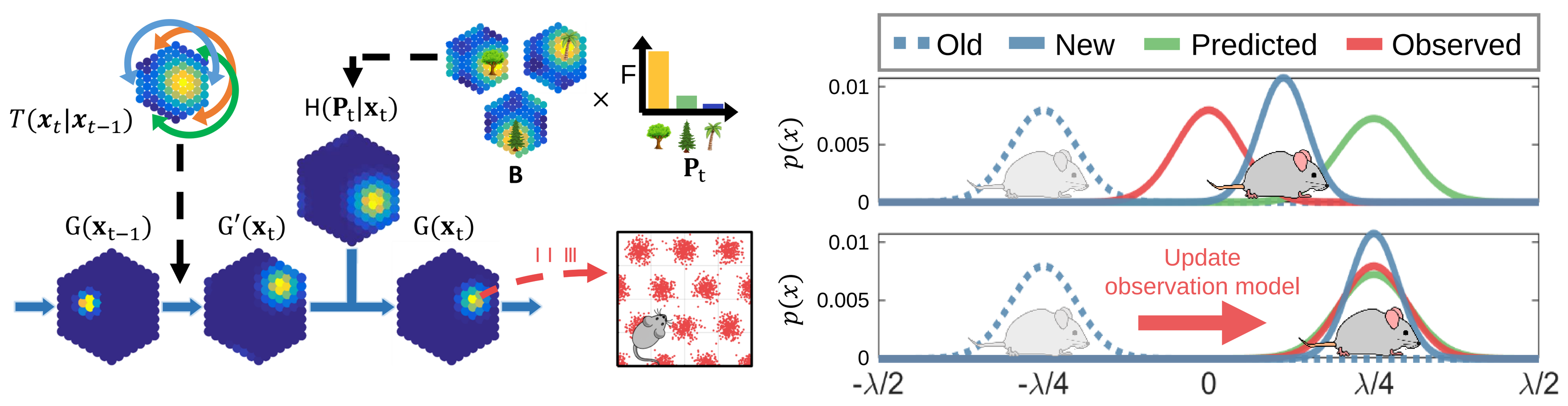
Grid Cells
(Hafting et al., 2005)

Encode transition structure

Online localization given a known map

- Location distribution encoded in GC firing rates
- Landmark beliefs encoded in PC-GC weights

- Online learning forms spatial priors over landmark locations



$$T(\mathbf{x}_t|\mathbf{x}_{t-1},\hat{\mathbf{u}}_t) = \sum_{m,n=-\infty}^{\infty} f(\mathbf{x}_t - \mathbf{x}_{t-1}|\hat{\mathbf{u}}_t + \mathbf{c}_{mn}, \sigma^2_{PI}\hat{\mathbf{u}}_t\mathbf{I})$$

Wrapped Gaussian transition model given self-motion estimate $\hat{\mathbf{u}}_t$ (1)

$$H(\mathbf{P}_t|\mathbf{x}_t) = [\mathbf{P}_t\mathbf{B}]_+$$

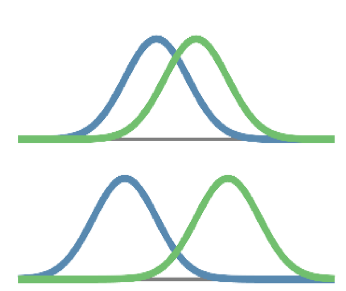
Observation model is weighted sum of PC firing (2)

$$\frac{1}{\alpha}\frac{d\mathbf{B}}{dt} = 2\mathbf{P}_t^\top(\mathbf{G}'_t - \mathbf{P}_t\mathbf{B})$$

Learning corrects prediction errors (3)

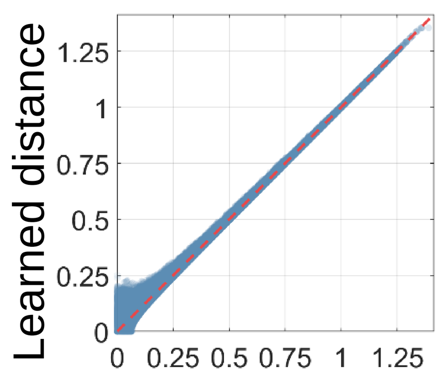
The hippocampus as a probabilistic graph

A

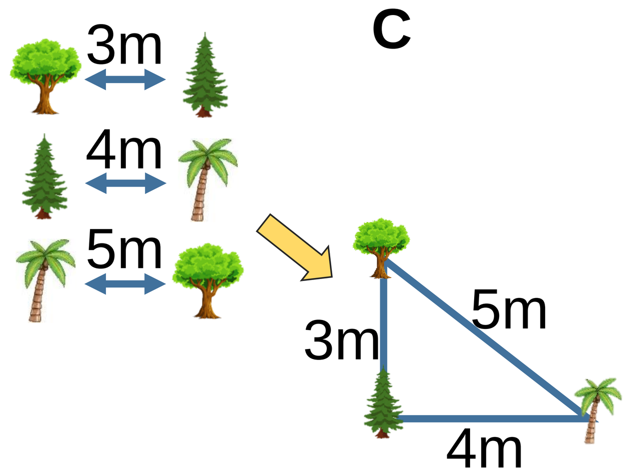


Place fields are approx. Gaussian

B



C



A $A_{ij} \approx \sqrt{< p_i(t), p_j(t) >}$ Hebbian learning is correlative (4)

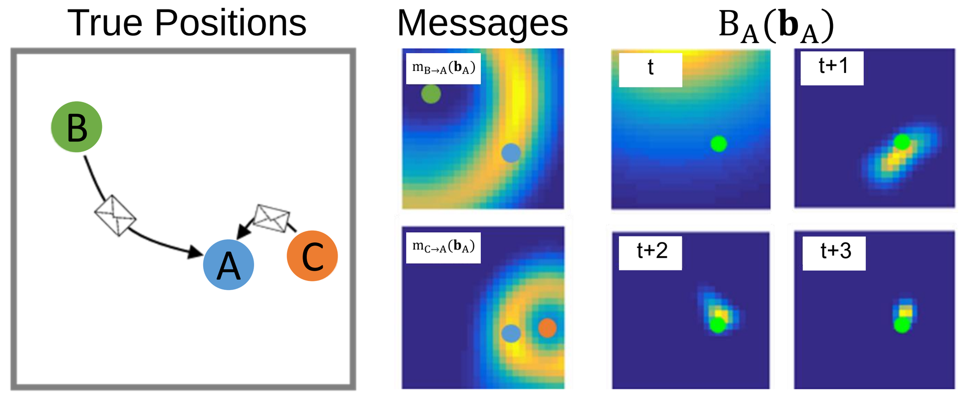
B $d_{ij}^2 = -\log(A_{ij}) = \frac{(\mu_i - \mu_j)^2}{2\sigma^2_{PC}}$ Synaptic weights encode Euclidean distance (5)

C $P(\{\mathbf{b}_p\}_{p=1:N_P}) = \prod_{1\le i\le N_P}\prod_{i\le j\le N_P}\psi_{ij}(\mathbf{b}_i,\mathbf{b}_j)\prod_{1\le i\le N_P}B_i(\mathbf{b}_i)$ Likelihood defines graphical model (6)

$$\psi_{ij}(\mathbf{b}_i,\mathbf{b}_j) = \sum_{m,n=\infty}^{\infty} exp(-\frac{1}{2\sigma^2_{ij}}(d_{ij} - ||\mathbf{b}_i - \mathbf{b}_j + \mathbf{c}_{mn}||_2)^2))$$

Pairwise associations penalize differences between *associative* and *metric* distances (7)

Offline inference through message passing



Belief propagation (Pearl, 1982) performs approximate inference by exchanging messages between landmark nodes (PCs)

Messages from multiple neighbours resolve location

$$B_t^{(n)}(\mathbf{b}_t) \propto B_t^{(0)}(\mathbf{b}_t)\prod_{u\in\Gamma_t} m^{(n)}_{u\to t}(\mathbf{b}_t)$$

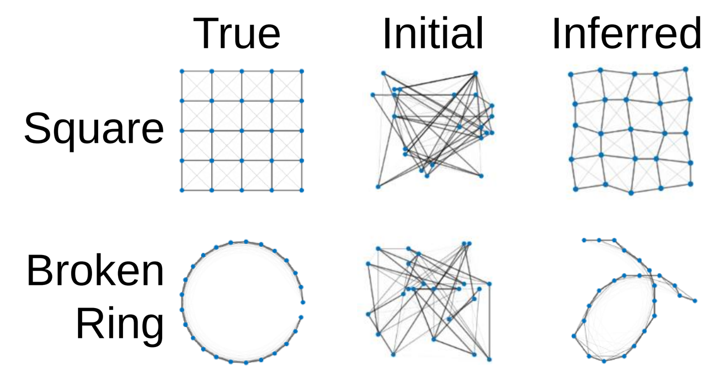
Belief update (8)

$$m^{n+1}_{t\to u}(\mathbf{b}_u) \propto \int \psi_{tu}(\mathbf{b}_t,\mathbf{b}_u)\frac{B_t^{(n)}(\mathbf{b}_t)}{m^{(n)}_{u\to t}(\mathbf{b}_t)}d\mathbf{b}_t$$

Message propagation (9)

Recovered structure is sensitive to topology of pairwise encodings

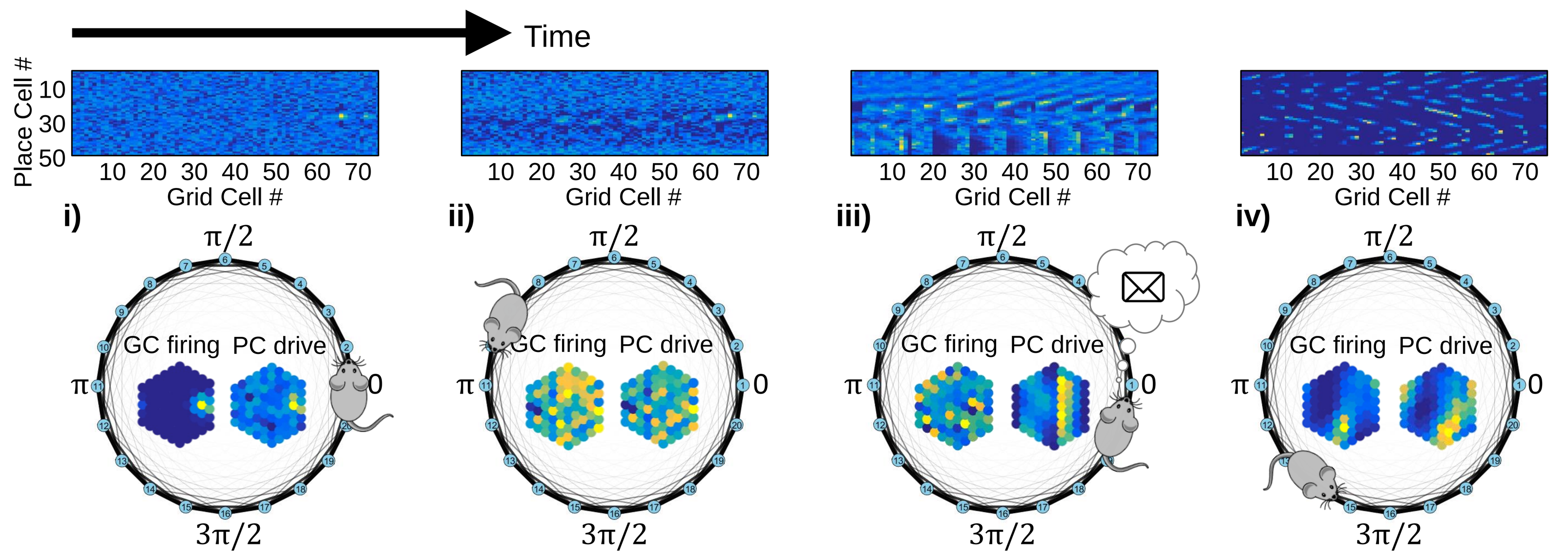
Predicts large structural updates following 'loop closures'



Loop closure experiment

$$\mathcal{E}_t = \mathcal{H}(\mathbf{G}_t) - \mathcal{H}(\mathbf{W}\times\mathbf{P}_t) \quad (10)$$

Offline system triggered by prediction error between the *transition* and *observation* models

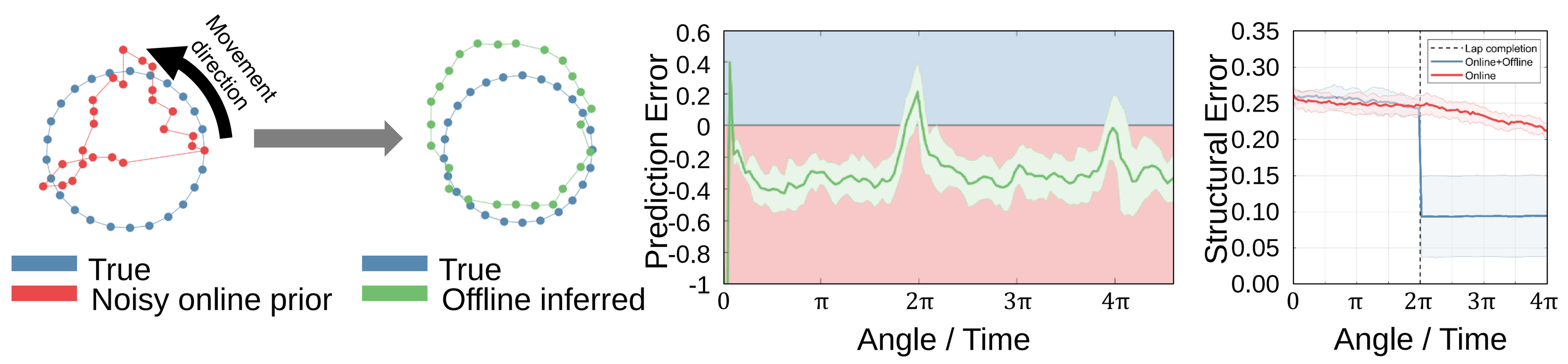


Strong initial sensory associations

Accumulated error from noisy path integration

Strong prediction error on lap completion

Accurate localization with learned structure

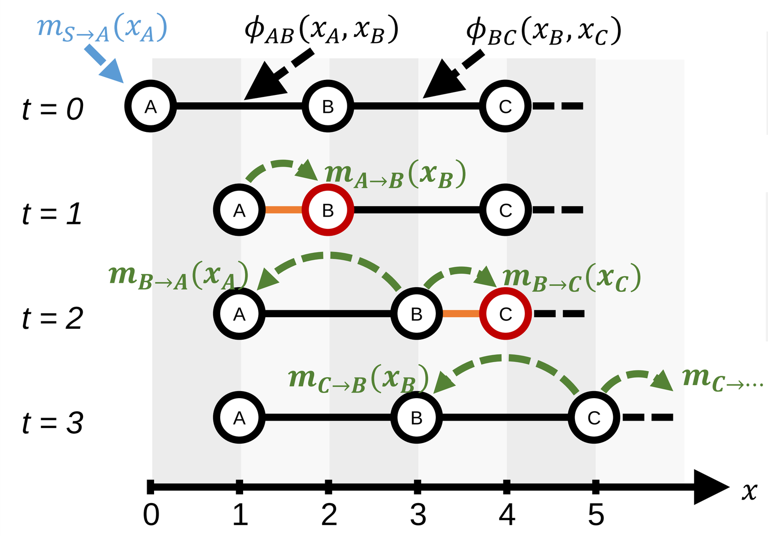


Offline inference corrects noisy spatial priors learned during *online learning*

Prediction errors reduced on subsequent laps

'One-shot' inference vs. Hebbian learning

Coordinated 'replay' as structured information propagation

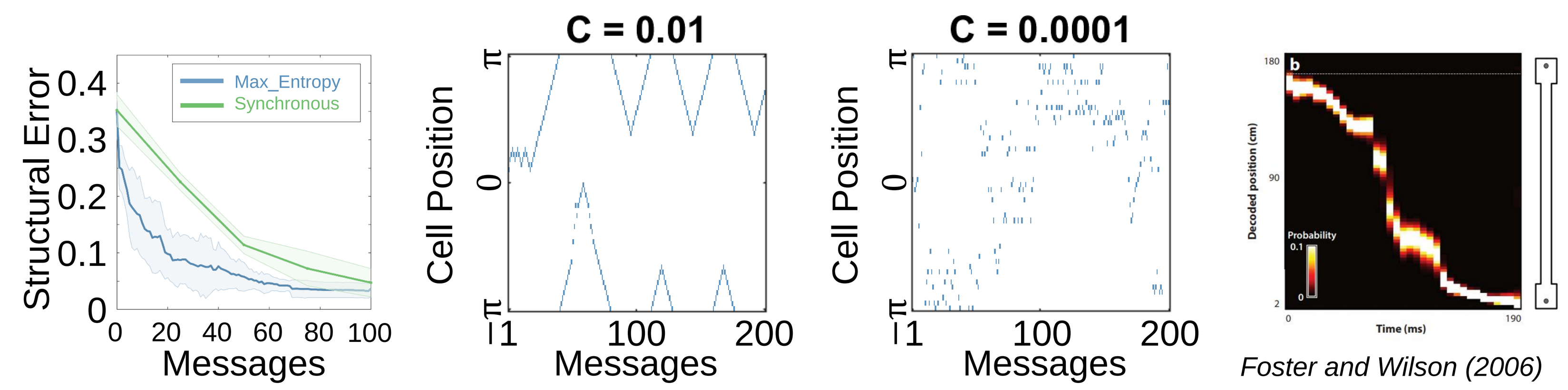


Naïvely scheduled messages contain significant redundant information

Instead, only propagate messages from nodes whose beliefs have changed, defined by message tension (Elidan et al., 2012):

$$T_i^n = T_i^{n-1} + \mathbb{JS}(\mathbf{b}_i^n||\mathbf{b}_i^{n-1}) \quad (11)$$

where JS is the Jensen-Shannon div.



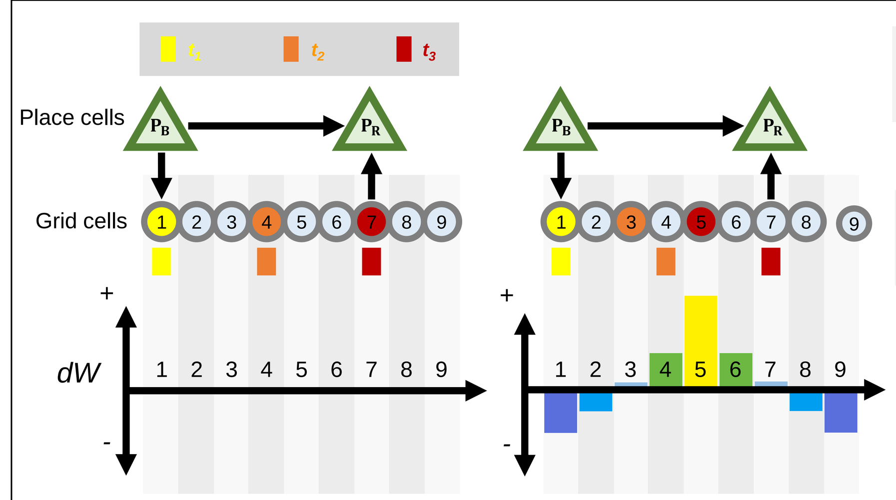
Asynchronous message scheduling is more efficient than naïve schedule

Scheduling produces sequential reactivations that resemble experimentally observed firing of place cells during 'offline' states (Foster and Wilson, 2006)

Sharp-wave **ripples** may correspond to structural prediction errors, which accompany 'replay' events

Roumis an Frank, 2015

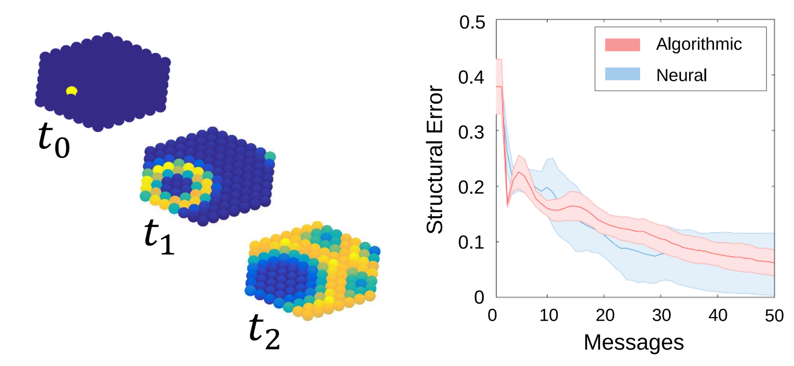
Approximate inference through synchronization



PC fires spike to neighbours with conduction delay proportional to place field separation

Simultaneous activity propagated outwards in *metric* space (GC sheet)

Timing of arrival between PC spike and GC wave causes synaptic changes which approximate belief update



Message broadcasts manifest as travelling fronts of activity in grid cell population

Inference in this system can then be viewed as synchronizing activity propagation through *metric* and *associative* generative models

Human Brain Project
wellcome
EPSRC Engineering and Physical Sciences Research Council
UCL CoMPLEX