Neural recognition and postdiction by temporal distributed distributional code

Li Kevin Wenliang, Maneesh Sahani
Gatsby Computational Neuroscience Unit, University College London kevinli@gatsby.ucl.ac.uk

UCL

Model

internal model: latent $p(\boldsymbol{z}_t|\boldsymbol{z}_{t\text{-}1})$ and observation $p(\boldsymbol{x}_t|\boldsymbol{z}_t)$

DDC: encode $q(\boldsymbol{z}_{1:t}|\boldsymbol{x}_{1:t}) \leftrightarrow \boldsymbol{r}_t := \mathbb{E}_q[\psi(\boldsymbol{z}_{1:t})]$

temporal code: $\psi_t = k(\psi_{t\text{-}1}, \boldsymbol{z}_t)$

sleep phase: train $h(\boldsymbol{r}_{t\text{-}1}, \boldsymbol{x}_t) \to \psi_t$ by MSE (δ-rule)

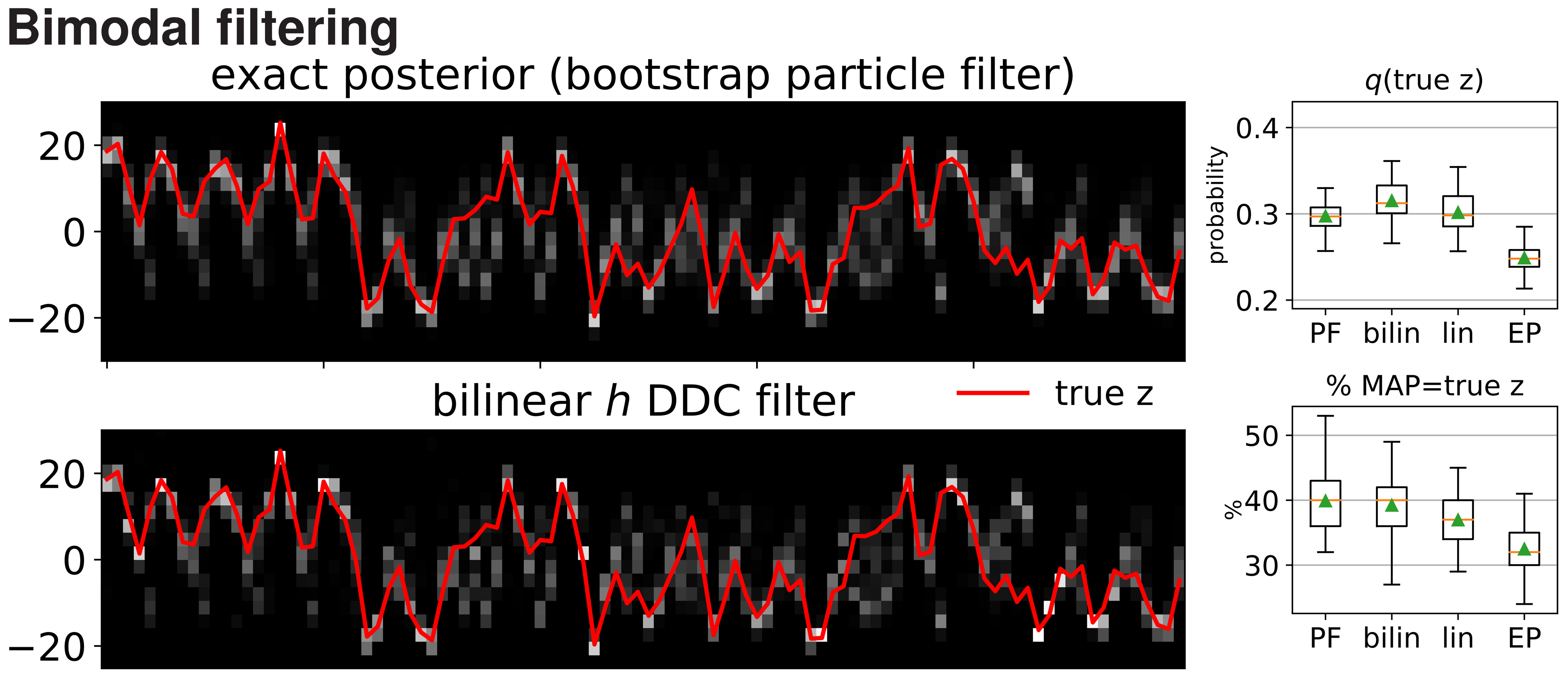
wake phase: predict $\mathbb{E}_q[\psi_t] \approx \boldsymbol{r}_t = h(\boldsymbol{r}_{t\text{-}1}, \boldsymbol{x}_t)$

flexible q (non-Gaussian), neural (δ-rule), postdictive

pre-print

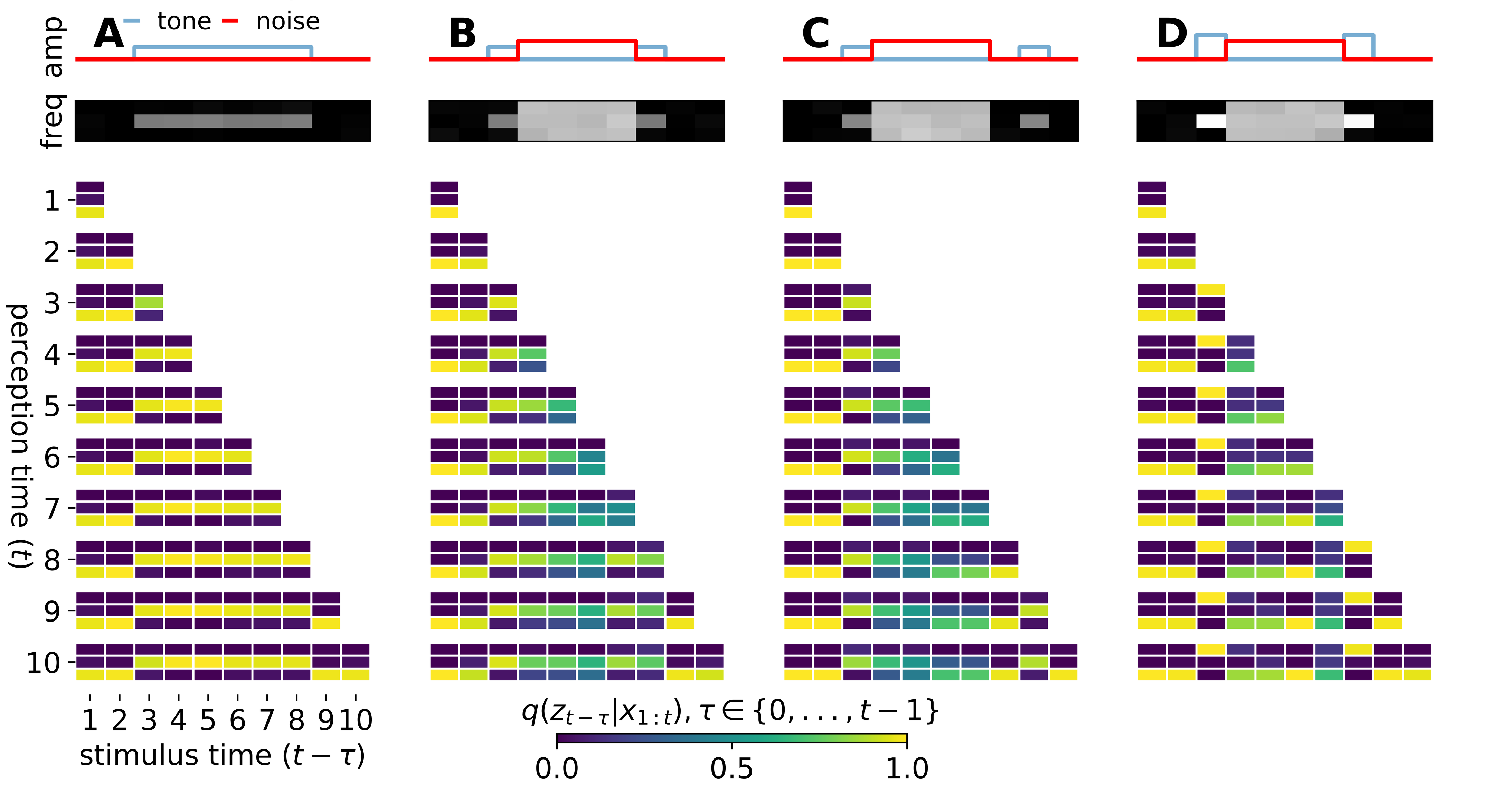
Key results

Bimodal filtering

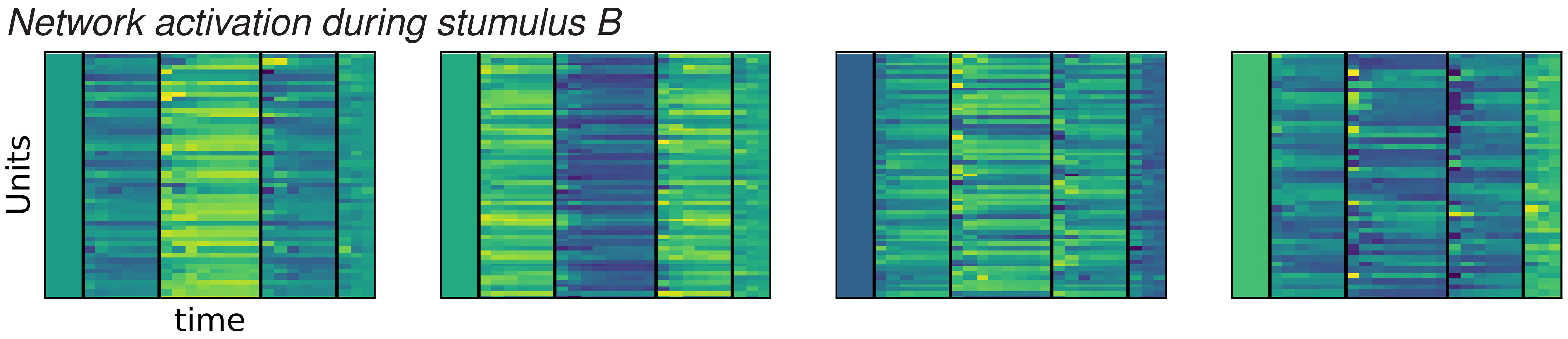


Auditory continuity illusion

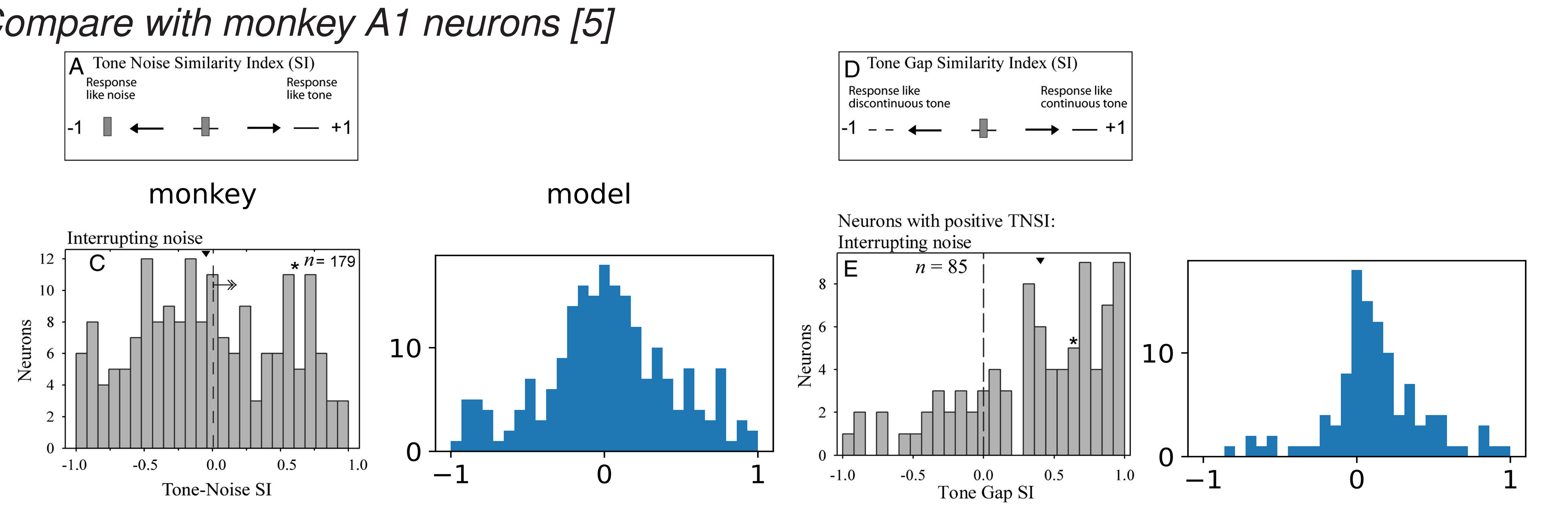
Postdictive filtering



Network activation during stimulus B

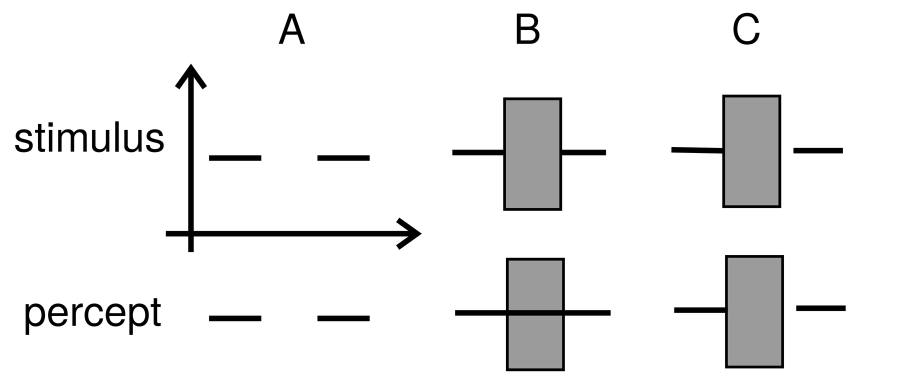


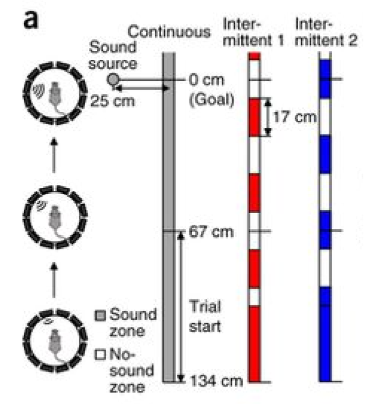
Compare with monkey A1 neurons [5]



Introduction

- Accurate and flexible state representation is crucial
- Experiments suggest optimal "cue combination" [2, 4]
- **Postdiction** is common in dynamic environments [6]
- Examples: continuity illusions [1], localisation [3]

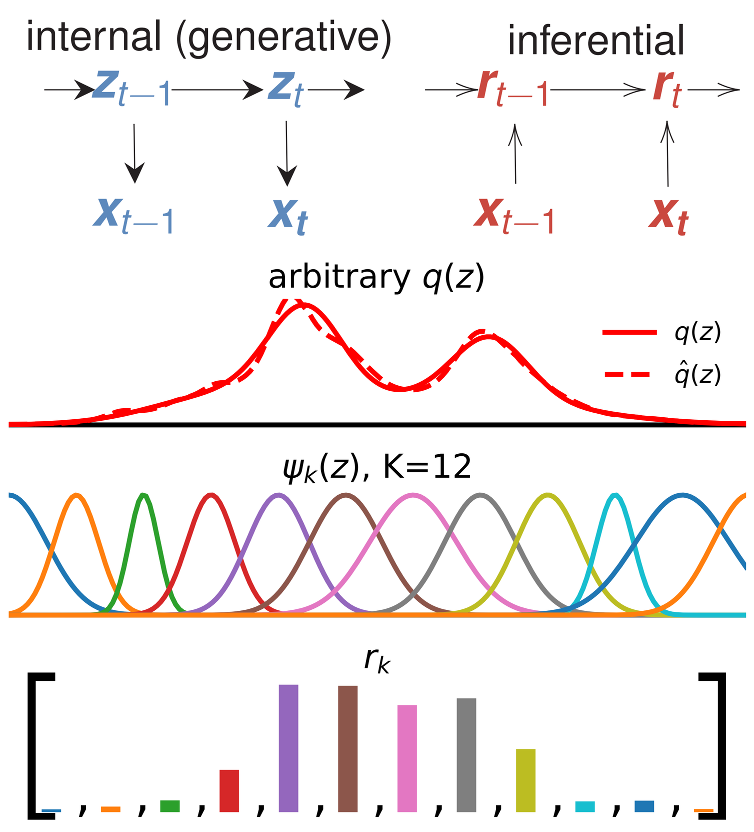




Internal and inferential model

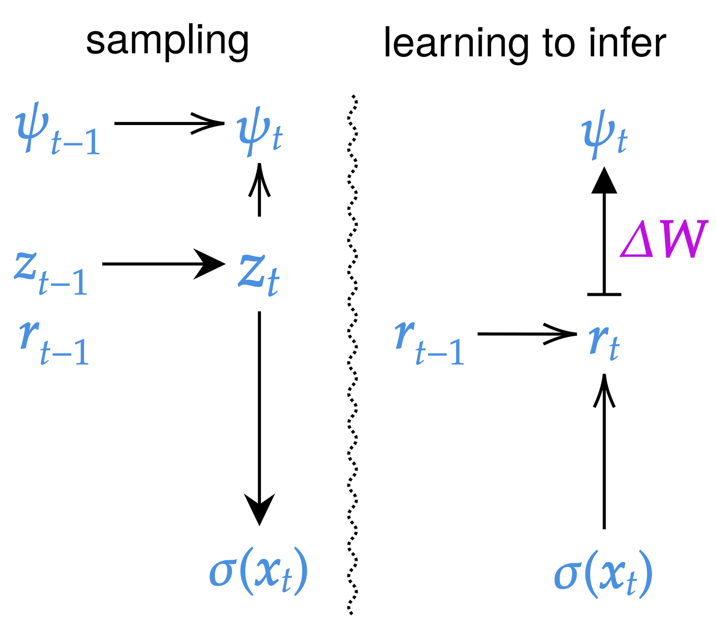
Internal model:
- latent causes
  $\boldsymbol{z}_t = f(\boldsymbol{z}_{t\text{-}1}) + \zeta_z$
- observation
  $\boldsymbol{x}_t = g(\boldsymbol{z}_t) + \zeta_x$
- No assumptions on ζ

Inferential model
- encode $q(\boldsymbol{z}_{1:t}|\boldsymbol{x}_{1:t})$
- by $\boldsymbol{r}_t = \mathbb{E}_q[\psi(\boldsymbol{z}_{1:t})]$ (DDC [7])
- $\psi_t = k(\psi_{t\text{-}1}, \boldsymbol{z}_t)$
- train $\boldsymbol{W}$:
  $h_{\boldsymbol{W}}(\boldsymbol{r}_{t\text{-}1}, \boldsymbol{x}_t) \xrightarrow{MSE} \psi_t$
  outputs $\mathbb{E}_q[\psi_t]$
- assess by max ent



Learning to infer

1. At each time t, have: $\boldsymbol{r}_{t\text{-}1}$, $\boldsymbol{z}_{t\text{-}1}$, and $\psi_{t\text{-}1}$
2. sample $\boldsymbol{z}_t, \boldsymbol{x}_t \sim p$
3. compute $\psi_t = k(\psi_{t\text{-}1}, \boldsymbol{z}_t)$
4. update $\boldsymbol{W}$ to minimise:
   $\|h_{\boldsymbol{W}}(\boldsymbol{r}_{t\text{-}1}, \boldsymbol{x}_t) - \psi_t\|_2^2$
   δ-rule if h is linear/bilinear
5. filter $\boldsymbol{r}_t = k(\boldsymbol{r}_{t\text{-}1}, \boldsymbol{x}_t)$
6. readout: find α
   $V(\boldsymbol{z}_{t\text{-}\tau:t}) \approx \alpha^\intercal \psi_t$
   $\mathbb{E}_q[V(\boldsymbol{z}_{t\text{-}\tau:t})] \approx \alpha^\intercal \boldsymbol{r}_t$



Model details

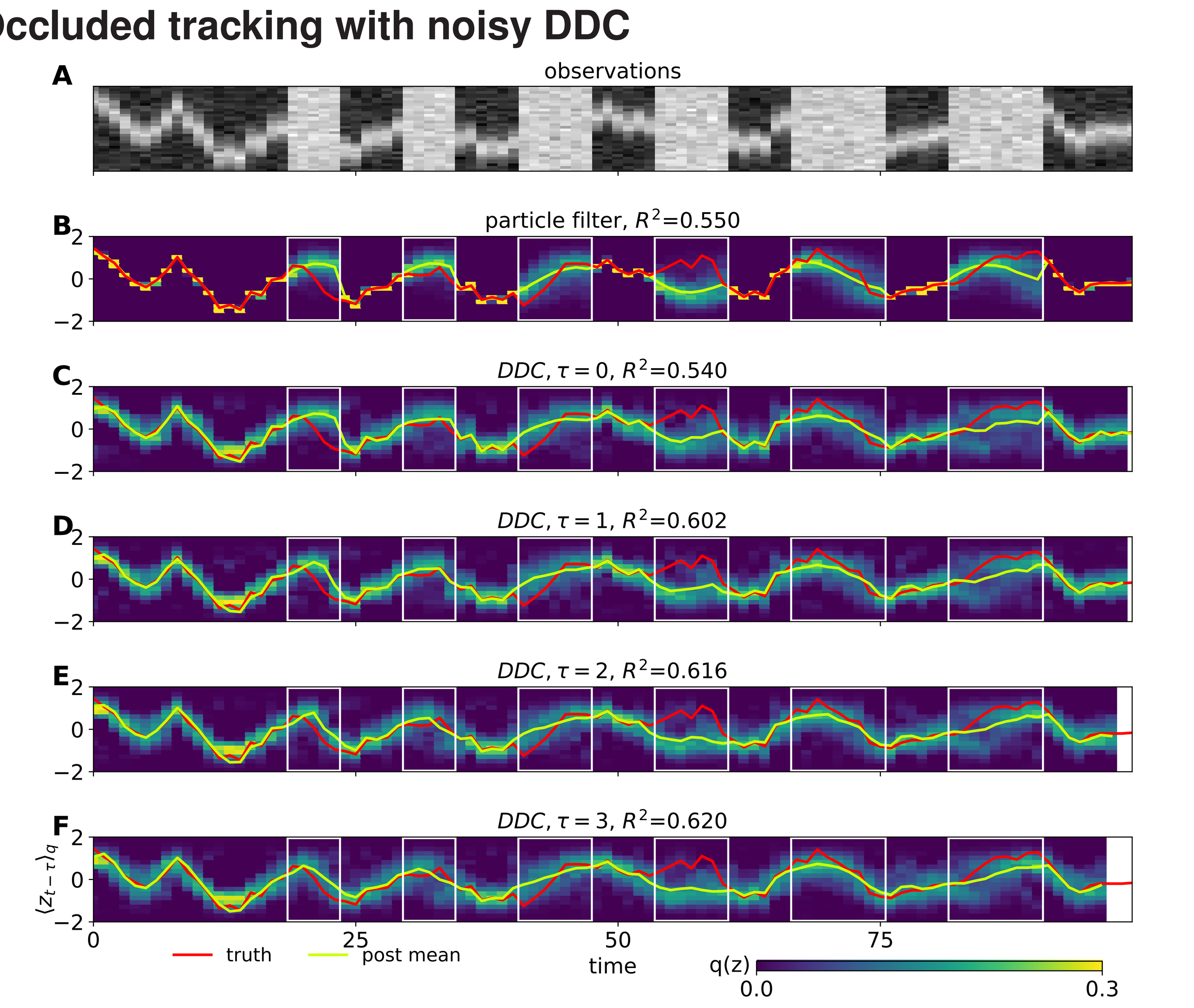
In sleep phase, model solves

$$\min_{\boldsymbol{W}} \mathbb{E}_{p(\boldsymbol{z}_{1:t}, \boldsymbol{x}_{1:t})}\left[\|h_{\boldsymbol{W}}(\boldsymbol{r}_{t\text{-}1}, \boldsymbol{x}_t) - \psi_t\|_2^2\right] \quad (1)$$

- $\boldsymbol{r}_{t\text{-}1}$ is a summary statistics of $\boldsymbol{x}_{1:t\text{-}1}$.
- $\psi_t = \boldsymbol{U}\psi_{t\text{-}1} + \gamma(\boldsymbol{z}_t)$, random but fixed temporal encoding function
- $h^{bil}_{\boldsymbol{W}} = \boldsymbol{W} \cdot (\boldsymbol{r}_{t\text{-}1}\sigma(\boldsymbol{x}_t)^\intercal)$, or $h^{lin}_{\boldsymbol{W}} = \boldsymbol{W}[\boldsymbol{r}_{t\text{-}1}; \sigma(\boldsymbol{x}_t)]$, $\sigma()$ random but fixed
- Possible to assume h is linear only in $\sigma(\boldsymbol{x}_t)$ and derive a formal solution, albeit with complicated neural implementation
- If the state-space model is stationary, $\boldsymbol{W}$ should converge
- Independent noise in ψ_t, σ and $\boldsymbol{r}_t$ averages out for large population
- Adaptation: follow gradient of variational objective $\nabla_\theta \mathcal{F}_\theta(\boldsymbol{z}, \boldsymbol{x})$

Additional results

Occluded tracking with noisy DDC



Reference

Albert S Bregman. MIT press, 1994.
Marc O Ernst and Martin S Banks. In: *Nature* 415.6870 (2002).
Akihiro Funamizu, Bernd Kuhn, and Kenji Doya. In: *Nature neuroscience* 19.12 (2016).
Konrad P. Körding, Shih-pi Ku, and Daniel M. Wolpert. In: *Journal of Neurophysiology* 92.5 (2004).
Christopher I Petkov, Kevin N O'Connor, and Mitchell L Sutter. In: *Neuron* 54.1 (2007).
Shinsuke Shimojo. In: *Frontiers in psychology* 5 (2014).
Eszter Vértes and Maneesh Sahani. In: *AISTATS* (2018).