AARHUS UNIVERSITY

Data Science 7

Chris Mathys

Master's Degree Programme in Cognitive Science

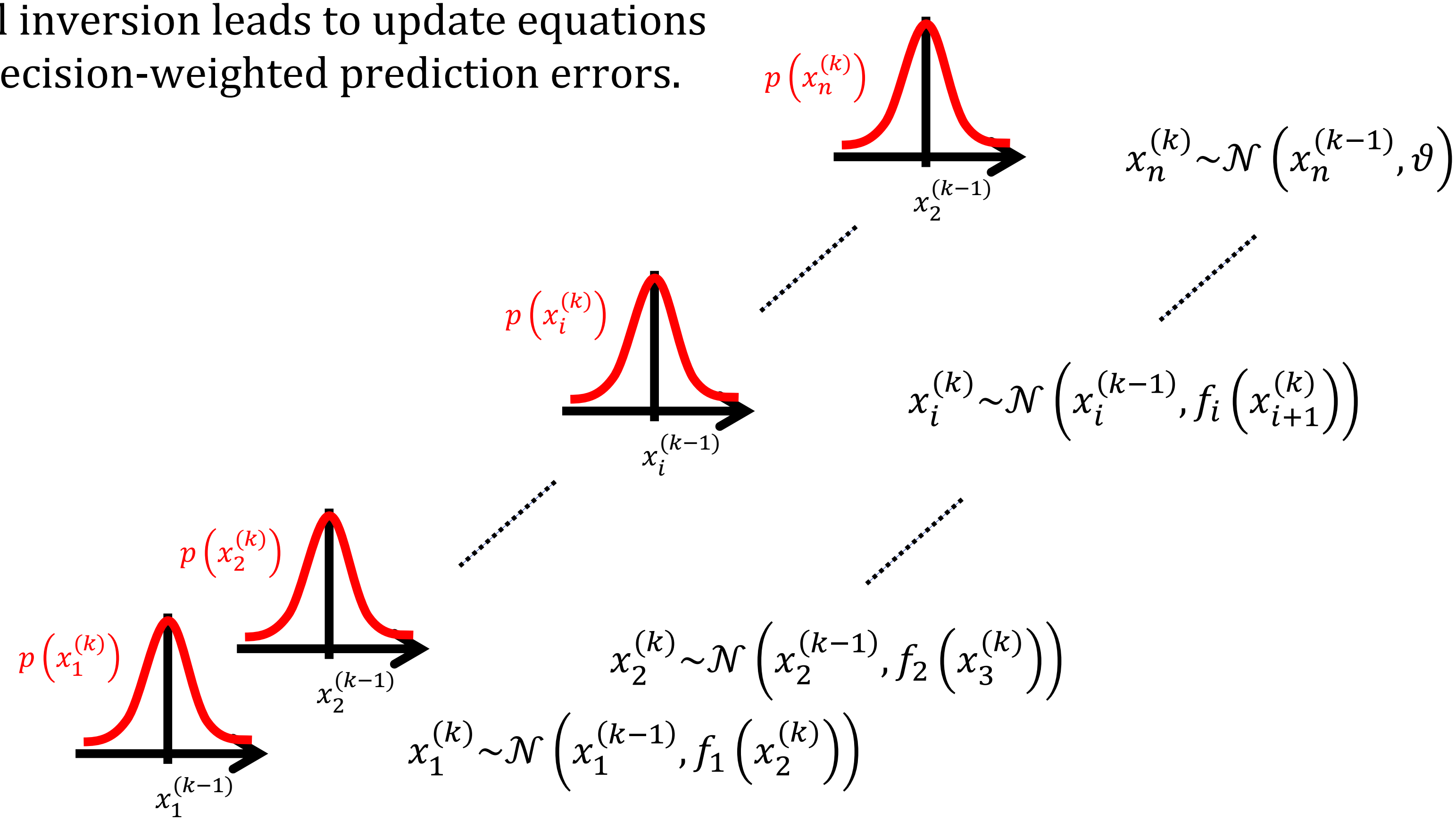
Spring 2023

Hierarchical Gaussian Filtering: introduction

The hierarchical Gaussian filter (HGF, Mathys et al., 2011; 2014)

The HGF provides a generic solution to the problem of adapting one's learning rate in a volatile environment.

Variational inversion leads to update equations that are precision-weighted prediction errors.



$$x_n^{(k)} \sim \mathcal{N}\left(x_n^{(k-1)}, \vartheta\right)$$

$$x_i^{(k)} \sim \mathcal{N}\left(x_i^{(k-1)}, f_i\left(x_{i+1}^{(k)}\right)\right)$$

$$x_2^{(k)} \sim \mathcal{N}\left(x_2^{(k-1)}, f_2\left(x_3^{(k)}\right)\right)$$

$$x_1^{(k)} \sim \mathcal{N}\left(x_1^{(k-1)}, f_1\left(x_2^{(k)}\right)\right)$$

Coupling between levels

Since f has to be everywhere positive, we cannot approximate it by expanding in powers. Instead, we expand its logarithm.

$$f(x) > 0 \; \forall\, x \implies \exists\, g\colon f(x) = \exp\big(g(x)\big) \; \forall\, x$$

$$g(x) = g(a) + g'(a) \cdot (x - a) + O(2) = \log f(x) =$$

$$= \log f(a) + \frac{f'(a)}{f(a)} \cdot (x - a) + O(2) =$$

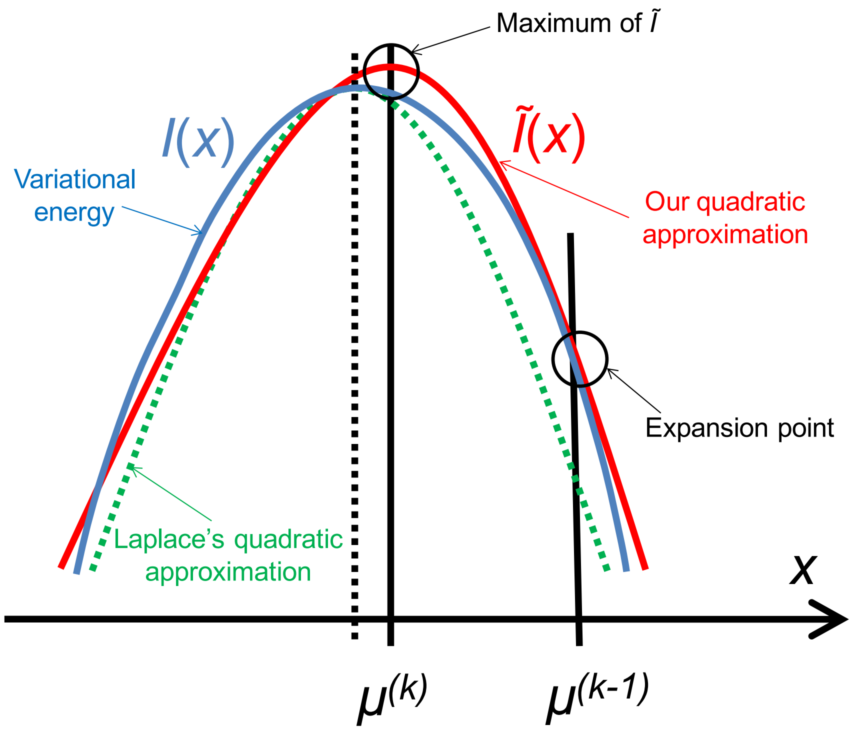
$$= \underbrace{\frac{f'(a)}{f(a)}}_{\stackrel{\text{def}}{\equiv}\kappa} \cdot x + \underbrace{\log f(a) - a \cdot \frac{f'(a)}{f(a)}}_{\stackrel{\text{def}}{\equiv}\omega} + O(2) =$$

$$= \kappa x + \omega + O(2)$$

$$\implies f(x) \approx \exp(\kappa x + \omega)$$

Variational inversion

- A quadratic approximation is found by expanding to second order about the expectation $\mu^{(k-1)}$.
- The update in the sufficient statistics of the approximate posterior is then performed by analytically finding the maximum of the quadratic approximation.



Mathys et al. (2011). *Front. Hum. Neurosci.*, 5:39.

$$\sigma_i^{(k)} = -\frac{1}{\partial^2 I\left(\mu_i^{(k-1)}\right)}$$

$$\mu_i^{(k)} = \mu_i^{(k-1)} - \frac{\partial I\left(\mu_i^{(k-1)}\right)}{\partial^2 I\left(\mu_i^{(k-1)}\right)} = \mu_i^{(k-1)} + \sigma_i^{(k)} \partial I\left(\mu_i^{(k-1)}\right)$$

Variational inversion and update equations

- Inversion proceeds by introducing a mean field approximation and fitting quadratic approximations to the resulting variational energies (Mathys et al., 2011).
- This leads to **simple one-step update equations** for the sufficient statistics (mean and precision) of the approximate Gaussian posteriors of the states x_i.
- The updates of the means have the same structure as value updates in Rescorla-Wagner learning:

Prediction error

$$\Delta\mu_i \propto \frac{\hat{\pi}_{i-1}}{\pi_i} \delta_{i-1}$$

Precisions determine learning rate

- Furthermore, the updates are **precision-weighted prediction errors**.

Precision-weighting of volatility updates

Comparison to the simple non-hierarchical Bayesian update:

HGF:

$$\mu_i^{(k)} = \mu_i^{(k-1)} + \frac{1}{2}\kappa_{i-1}\, v_{i-1}^{(k)} \cdot \frac{\hat{\pi}_{i-1}^{(k)}}{\pi_i^{(k)}} \cdot \delta_{i-1}^{(k)}$$

Precision-weighted prediction error

Simple Gaussian:

$$\mu_{\vartheta|x} = \mu_\vartheta + \frac{\pi_\varepsilon}{\pi_{\vartheta|x}}(x - \mu_\vartheta)$$

Updates at the outcome level

At the outcome level (i.e., at the very bottom of the hierarchy), we have

$$u^{(k)} \sim \mathcal{N}\left(x_1^{(k)}, \hat{\pi}_u^{-1}\right)$$

This gives us the following update for our belief on x_1 (our quantity of interest):

$$\pi_1^{(k)} = \hat{\pi}_1^{(k)} + \hat{\pi}_u$$

$$\mu_1^{(k)} = \mu_1^{(k-1)} + \frac{\hat{\pi}_u}{\pi_1^{(k)}}\left(u^{(k)} - \mu_1^{(k-1)}\right)$$

The familiar structure again – but now with a learning rate that is responsive to all kinds of uncertainty, including environmental (unexpected) uncertainty.

The learning rate in the HGF

Unpacking the learning rate, we see:

outcome uncertainty

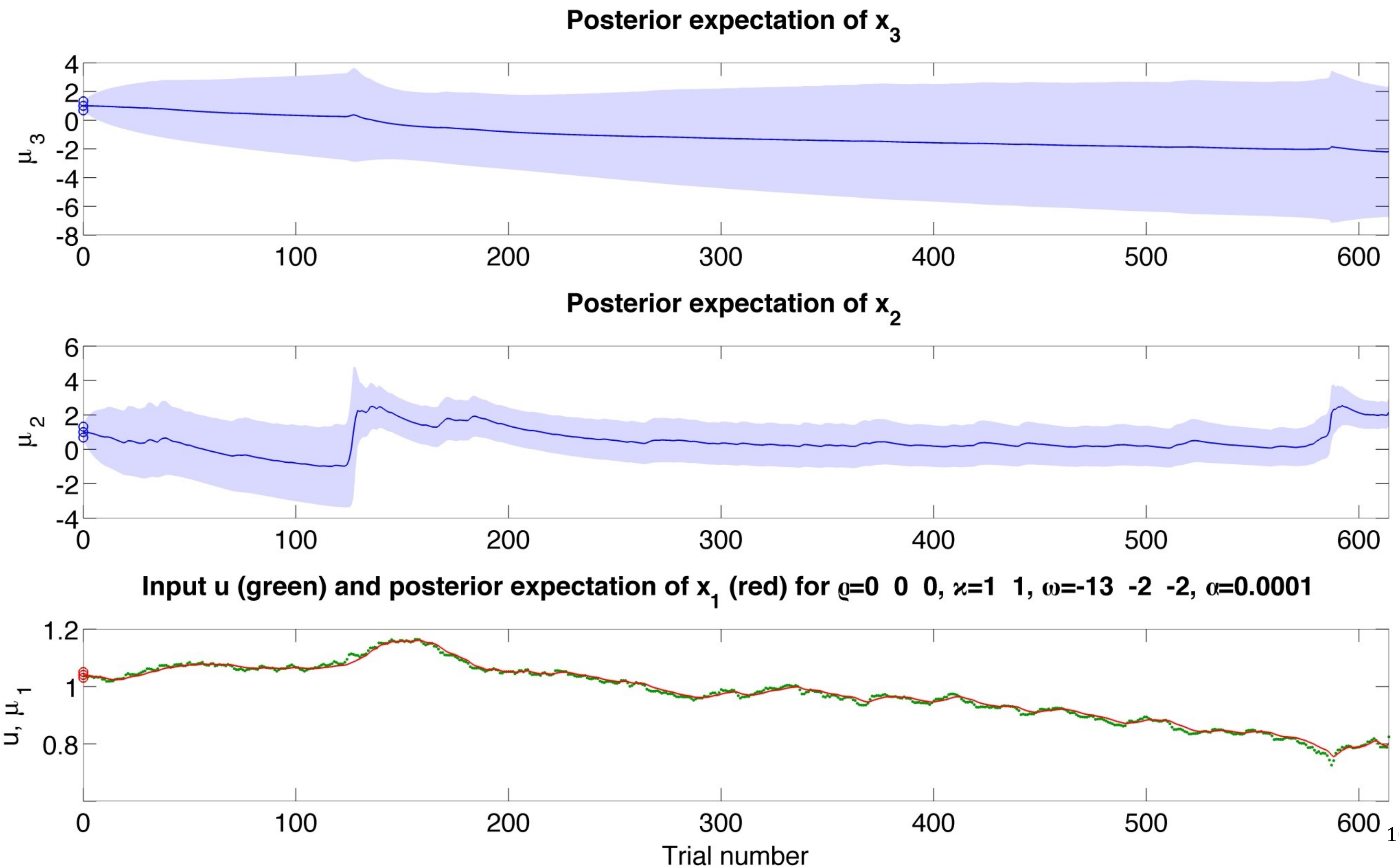
$$\frac{\hat{\pi}_u}{\pi_1^{(k)}} = \frac{\hat{\pi}_u}{\hat{\pi}_1^{(k)} + \hat{\pi}_u} = \frac{\hat{\pi}_u}{\dfrac{1}{\sigma_1^{(k-1)} + \exp\left(\kappa_1 \mu_2^{(k-1)} + \omega_1\right)} + \hat{\pi}_u}$$

informational uncertainty

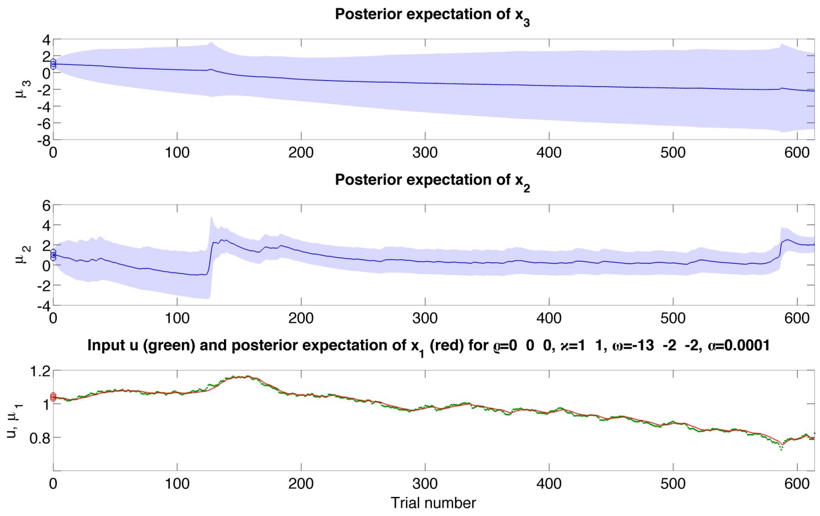
environmental uncertainty

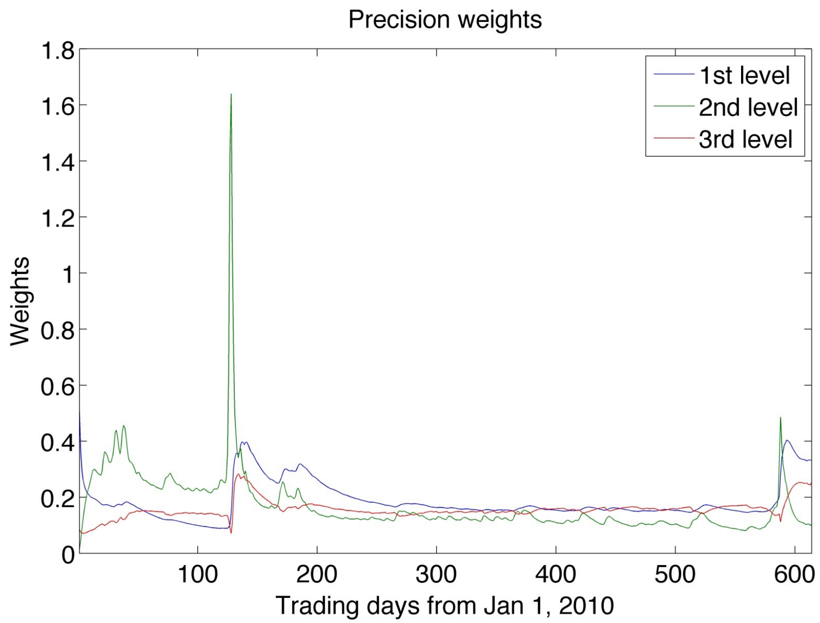
(instead of the constant ϑ in the Kalman filter)

3-level HGF for continuous observations

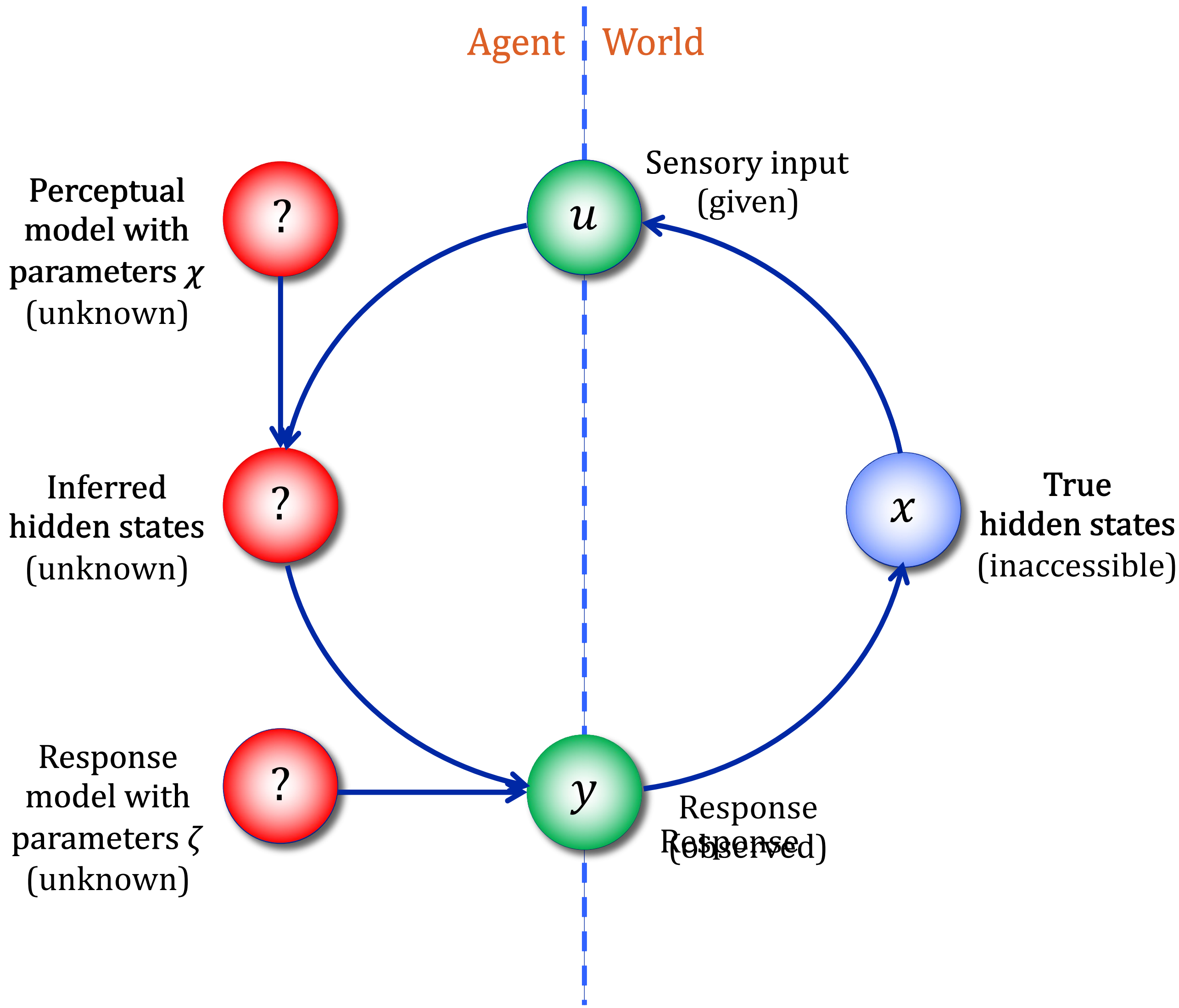


Example of precision weight trajectory

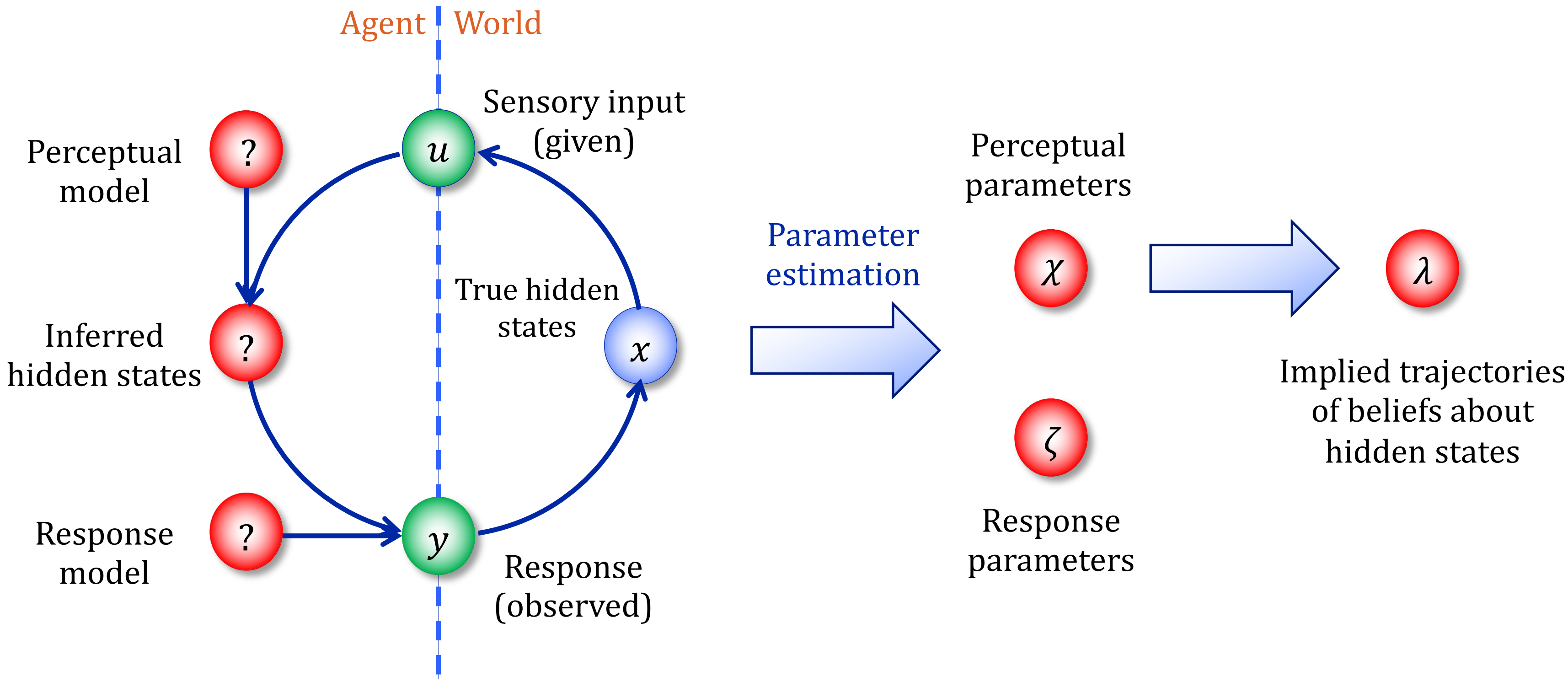




Application to experimental data

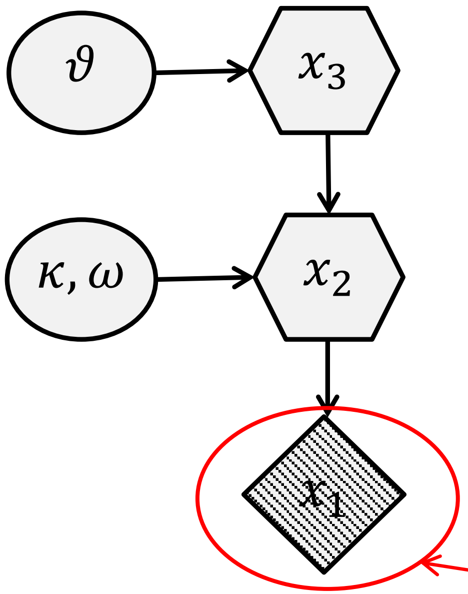


Application to experimental data: parameter estimation



Generative model for trial outcomes in the Iglesias et al. task:

3-level HGF for binary observations



$$x_3^{(k)} \sim \mathcal{N}\left(x_3^{(k-1)}, \vartheta\right)$$

$$x_2^{(k)} \sim \mathcal{N}\left(x_2^{(k-1)}, \exp\left(\kappa x_3^{(k)} + \omega\right)\right)$$

$$x_1^{(k)} \sim \mathrm{Bern}\left(s\left(x_2^{(k)}\right)\right)$$

Outcome: 1 if high/low tone
followed by face/house, 0 otherwise

Mathys et al., 2011; Iglesias et al., 2013; Vossel et al., 2014a; Hauser et al., 2014; Diaconescu et al., 2014; Vossel et al., 2014b; ...

Generative model, perceptual model, and decision model

- The *generative model* model describes (probabilistically) how the states x_i evolve in time: $x_i^{(k-1)} \longrightarrow x_i^{(k)}$
- Variationally inverting the generative gives us the *perceptual model* (also called *inference model* or *recognition model)*
- The perceptual model describes (deterministically, via update equations) how *beliefs* $\{\mu_i, \pi_i\}$ about states evolve in time: $\left\{\mu_i^{(k-1)}, \pi_i^{(k-1)}\right\} \longrightarrow \left\{\mu_i^{(k)}, \pi_i^{(k)}\right\}$
- The decision model (also called observation model or response model) describes (probabilistically) how beliefs are translated into observed actions: $\left\{\mu_i^{(k)}, \pi_i^{(k)}\right\}_{i=1,\dots,l} \longrightarrow y^{(k)}$
- In the process of applying the HGF to data, we only need the perceptual and decision models. The generative model has already done its work: it has supplied the update equations of the perceptual model
- Both the perceptual and decision models have parameters that can be estimated individually for each dataset. Doing so requires defining priors for these parameters.

Parameter estimation with the HGF Toolbox

- Available at

  https://translationalneuromodeling.github.io/tapas

- Start with README, manual, and interactive demo

- Modular, extensible, Matlab-based

```
est2 = tapas_fitModel(sim2.y,...
                      usdchf,...
                      'tapas_hgf_config',...
                      'tapas_gaussian_obs_config',...
                      'tapas_quasinewton_optim_config');
```

```
Parameter estimates for the perceptual model:
    mu_0: [1.0352 1]
    sa_0: [3.7101e-05 0.0996]
     rho: [0 0]
      ka: 1
      om: [-12.8680 -1.8689]
    pi_u: 9.8449e+03
```

```
Parameter estimates for the observation model:
    ze: 2.3970e-05
```

Parameter estimation with the HGF Toolbox

Activations by Precision-Weighted Visual Outcome Prediction Error ε_2

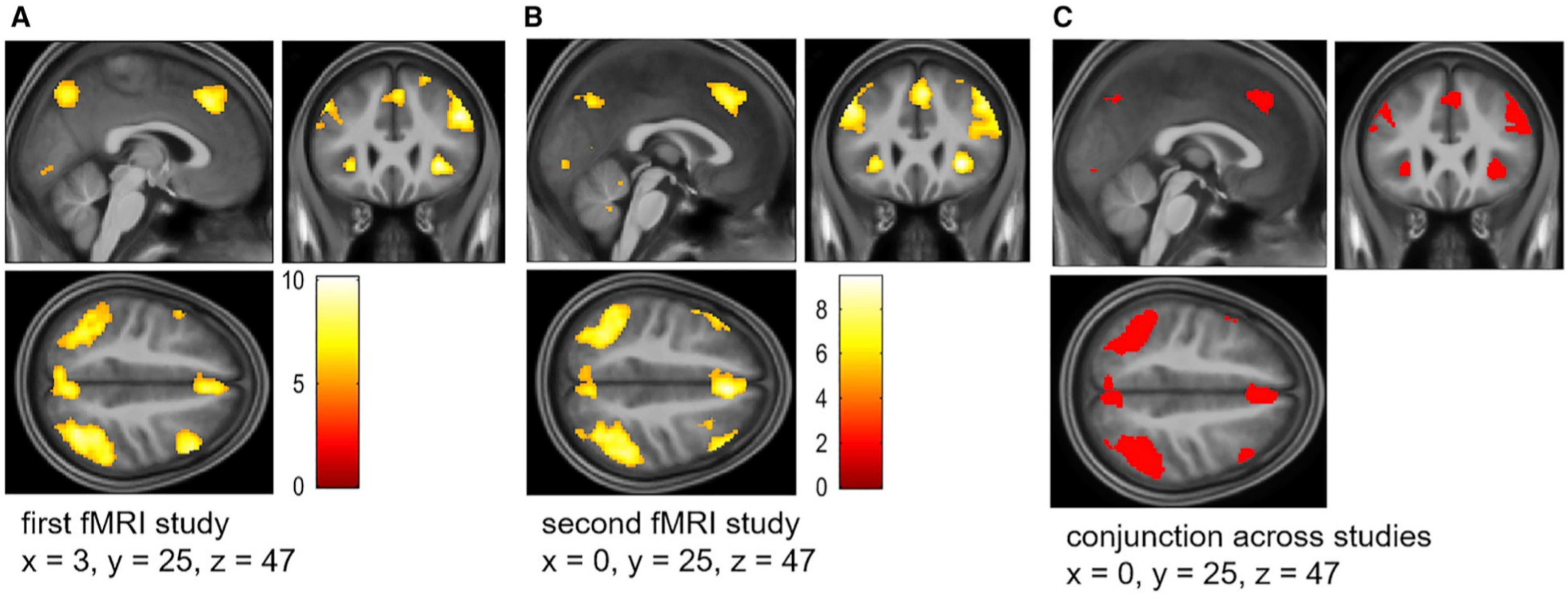


Figure 2. Whole-Brain Activations by ε_2

Activations by precision-weighted prediction errors about visual stimulus outcome, ε_2, in the first fMRI study (A) and the second fMRI study (B). Both activation maps are shown at a threshold of $p < 0.05$, FWE peak-level corrected for multiple comparisons across the whole brain. To highlight replication across studies, (C) shows the results of a "logical AND" conjunction, illustrating voxels that were significantly activated in both studies.

Iglesias et al, 2019 (Correction to Iglesias et al., 2013)

Parameter estimation with the HGF Toolbox

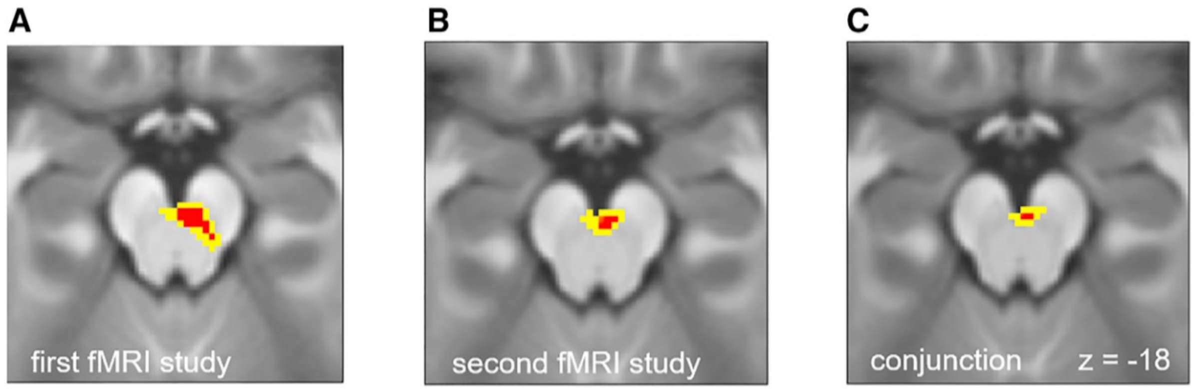


Figure 3. Midbrain Activation by ε_2
Activation of the dopaminergic VTA/SN by precision-weighted prediction errors about visual outcome, ε_2. The activation at $p < 0.05$ FWE peak-level corrected for the volume of our anatomical mask (comprising both dopaminergic and cholinergic brain structures: VTA/SN, PPT/LDT, and basal forebrain) is shown in red. The activation thresholded at $p < 0.001$ uncorrected is shown in yellow.
(A) Results from the first fMRI study. (B) Second fMRI study. (C) Conjunction (logical AND) across both studies.

Iglesias et al, 2019 (Correction to Iglesias et al., 2013)

Parameter estimation with the HGF Toolbox

Activations by Precision-Weighted Prediction Error about Stimulus Probabilities ε_3

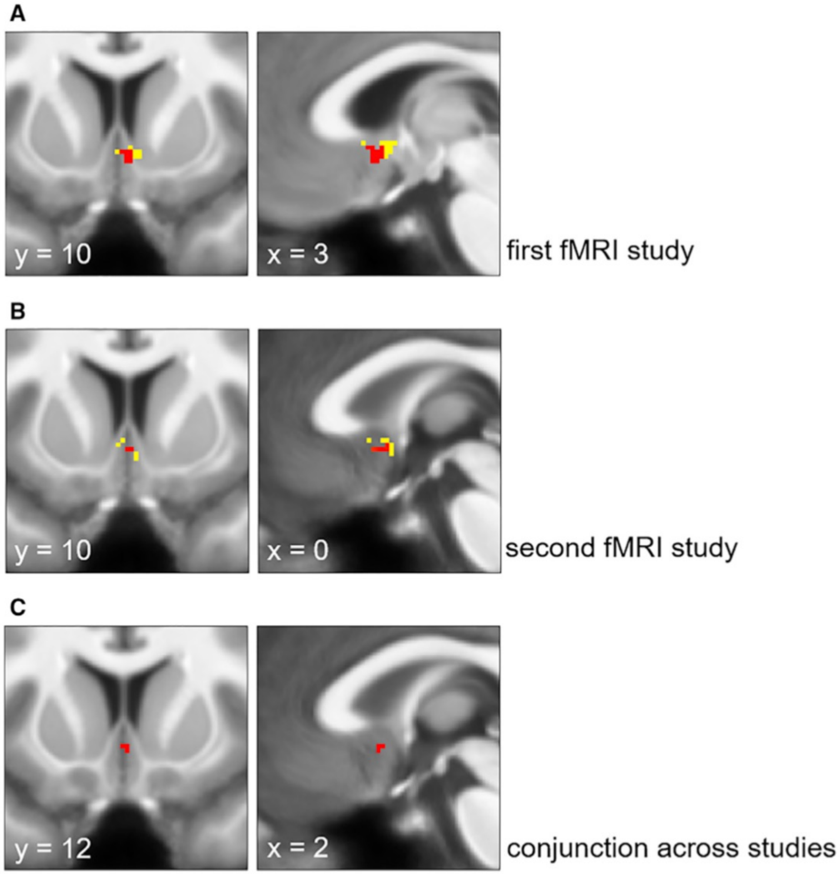


Figure 6. Basal Forebrain Activations by ε_3

Activation of the basal forebrain by precision-weighted prediction error about stimulus probabilities ε_3 within the anatomically defined mask. For visualization of the activation area, we overlay the results thresholded at $p < 0.05$ FWE peak-level corrected for the entire anatomical mask (red) on the results thresholded at $p < 0.001$ (yellow; the yellow cluster also survives $p < 0.05$ FWE cluster-level correction for the entire anatomical mask). The anatomical mask comprised both dopaminergic and cholinergic brain structures: VTA/SN, PPT/LDT, and basal forebrain. (A) and (B) show results from the first (A: local maximum at $x = 4$, $y = 12$, $z = -11$, $t = 4.71$) and the second fMRI study (B: local maximum at $x = 0$, $y = 10$, $z = -8$, $t = 5.09$). (C) shows the conjunction analysis ("logical AND") across both studies. To ease visual comparison with Iglesias et al. (2013), the figure sections (x and y coordinates are indicated on each panel) are not located at the local maxima but correspond closely to those in Iglesias et al. (2013).

Iglesias et al, 2019
(Correction to Iglesias et al., 2013)

Generation of new data: simulation