Inferring the individual nature of Bayesian learning under multiple forms of uncertainty

Christoph Mathys[1,2,†] · Jean Daunizeau[1,4] · Kay H. Brodersen[1,3] · Sandra Iglesias[1] · Karl J. Friston[4] · Klaas E. Stephan[1,4]

[1] Laboratory for Social and Neural Systems Research (SNS Lab), University of Zurich [2] Institute for Biomedical Engineering, ETH Zurich [3] Department of Computer Science, ETH Zurich [4] Wellcome Trust Centre for Neuroimaging, University College London

†chmathys@ethz.ch

ETH Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

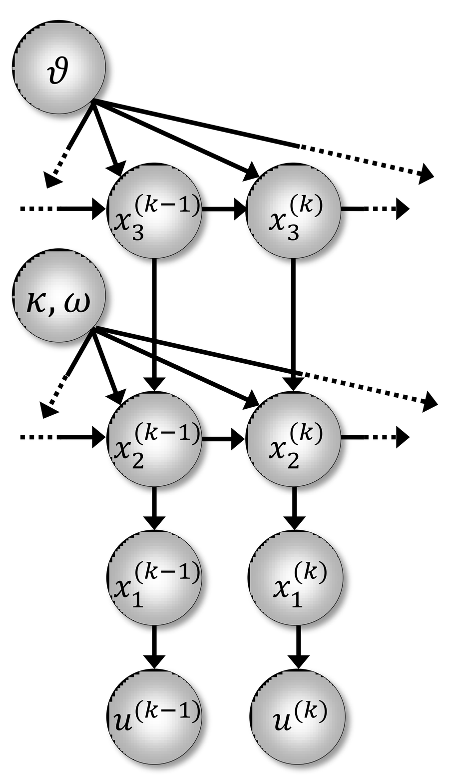
1 Introduction

Probability theory formally prescribes an optimal way for learning about the environment from sensory information: sequential updating of beliefs according to Bayes' theorem. This principle can be extended to hierarchical Bayesian models when dealing with higher-order uncertainty induced by the time-varying structure of the environment. While optimal from the perspective of probability theory, Bayesian belief updating requires evaluating complicated integrals which are not tractable analytically and difficult to evaluate in real time. Recently, however, theoretical advances have enabled computationally efficient approximations to exact Bayesian inference during learning. Here, we focus on a recent derivation of reinforcement learning from Bayesian principles [1] which rests on a generative hierarchical model of the environment and its (in)stability. This model (i) provides analytic and computationally highly efficient update equations, (ii) allows for on-line estimates of the model's states and parameters and (iii) includes a mechanism for time-varying encoding of the precision of beliefs which may correspond to the modulatory effects of dopamine as proposed by [2]. Here, we combine this perceptual model with a model for subjects' responses, such that decisions are informed by current beliefs about the state of the world.

1. Mathys, C., et al. (2011). A Bayesian foundation for individual learning under uncertainty. *Front. Hum. Neurosci.*, **5:** 39.
2. Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences*, **13:** 293-301.

2 The generative model: a hierarchy of Gaussian random walks

An agent is taken to receive a sequence of inputs $u^{(1)}, u^{(2)}, \ldots$ It uses these to make inferences on a hierarchy of hidden states $x_1, x_2, \ldots$ of its environment. While x_1 is binary, all higher states are continuous. Continuous states change by performing Gaussian random walks that are hierarchically coupled with one state's step size determined by the next highest state.



State of the world	Model	
Log-volatility x_3 of tendency	Gaussian random walk with constant step size ϑ	$p(x_3^{(k)}) \sim N(x_3^{(k-1)}, \vartheta)$
Tendency x_2 towards category "1"	Gaussian random walk with step size $\exp(\kappa x_3+\omega)$	$p(x_2^{(k)}) \sim N(x_2^{(k-1)}, \exp(\kappa x_3+\omega))$
Stimulus category x_1 ("0" or "1")	Sigmoid transformation of x_2	$p(x_1=1) = s(x_2)$, $p(x_1=0) = 1-s(x_2)$
Sensory input u	Determined by x_1 and perceptual noise	

Fig. 1 | *The generative model.* *The parameters κ, ω, and ϑ determine the way the agent sees its environment.*

3 A novel variational inversion leads to closed-form update equations based on learning rate and prediction error

We invert the generative model variationally using a mean field approximation and a novel quadratic approximation to the variational energy that defines the posteriors of the continuous states. This permits us to derive closed-form Markovian update equations for the posterior expectations of all states. These update equations are structurally similar to those of reinforcement learning.

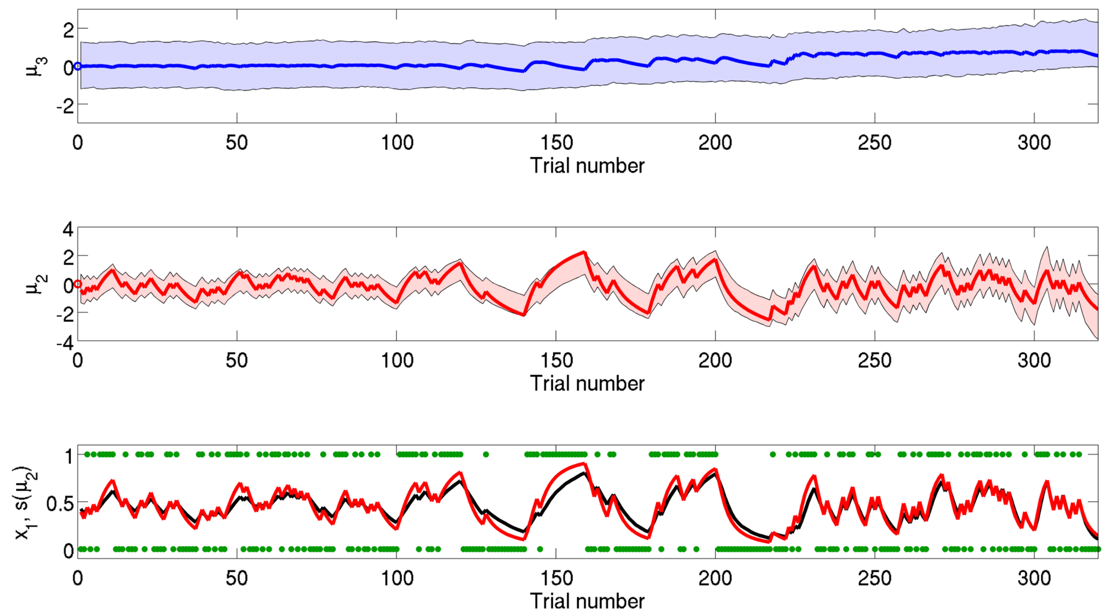


Fig. 2 | *Comparison of new variational inversion to inversion by MCMC sampling.* *Top and middle: Variational approximation (solid colored line) stays within 10-90 percentile range of samples (shaded areas) at both 2nd and 3rd levels. Bottom: Belief that next x_1 (green) will be 1 corresponds closely in approximation (red) and sampling (black)*

4 The response model

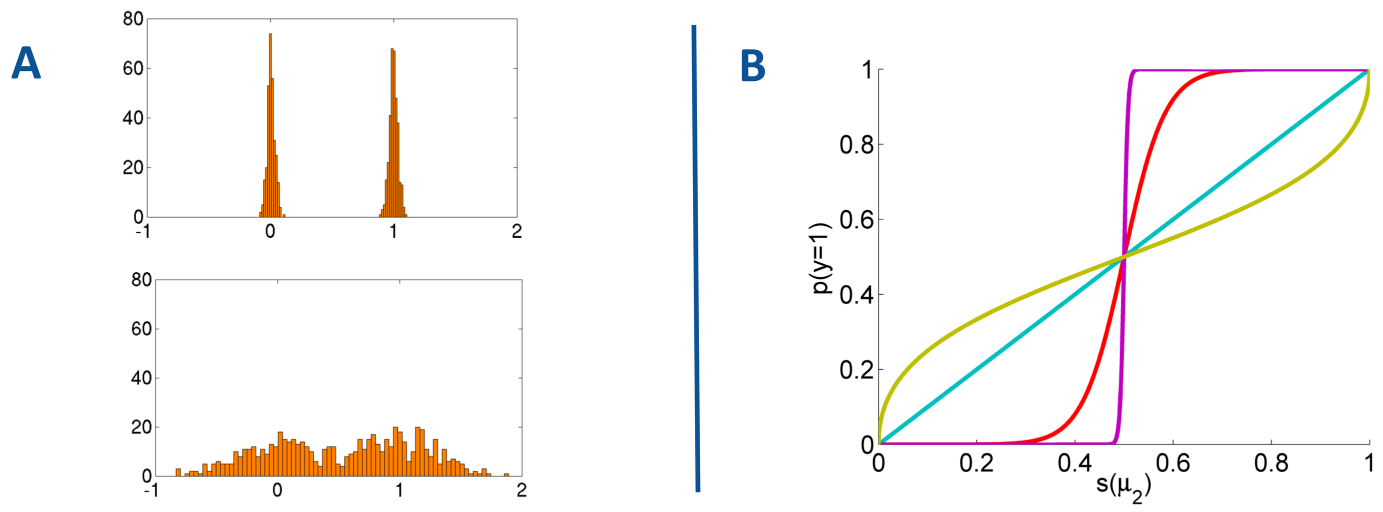


Fig. 3 | *Perceptual input and decision output.* *(A) Distribution of perceptual input in simulations of Section 5. Top: low noise (α=0.001; cf. orange dots in Fig. 4A). Bottom: high noise (α=0.1; cf. orange dots in Fig. 4B). According to the perceptual generative model, input is generated from a mixture of Gaussians with arbitrary means, here 0 and 1. (B) "Softmax" decision rule for different levels of decision noise. The abscissa measures the posterior probability (sigmoid s(μ2) of tendency μ2) that the next input indicates a hidden state of type 1; the ordinate measures the probability of a decision y=1 predicting a hidden state of type 1. The higher the noise, the flatter the curve (i.e., the less the decision reflects the current belief). The curve's shape is determined by the parameter ζ. Purple: ζ=64 (used in Fig. 4A); red: ζ=6; cyan: ζ=1 (used in Fig. 4B); olive: ζ=0.5.*

5 Multiple forms of uncertainty: perceptual and decision noise

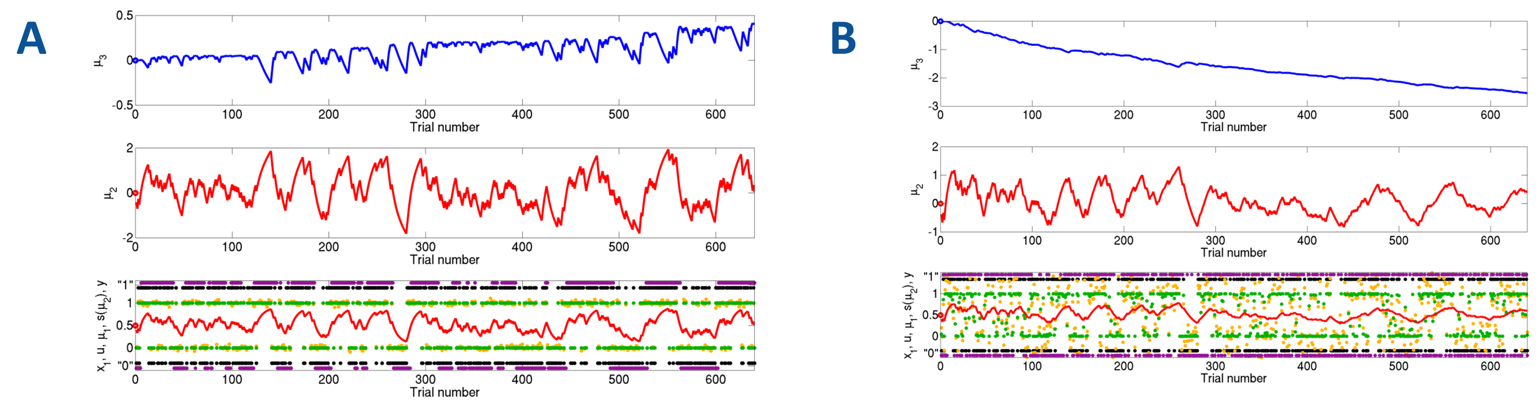


Fig. 4 | *Simulation examples.* *Black: hidden true state, orange: perceptual input; green: belief on probability of true state = 1; red: belief on tendency of true state; blue: belief on volatility of tendency; violet: decision. (A) low noise. (B) high noise.*

6 Parameter estimation and application to behavioral data

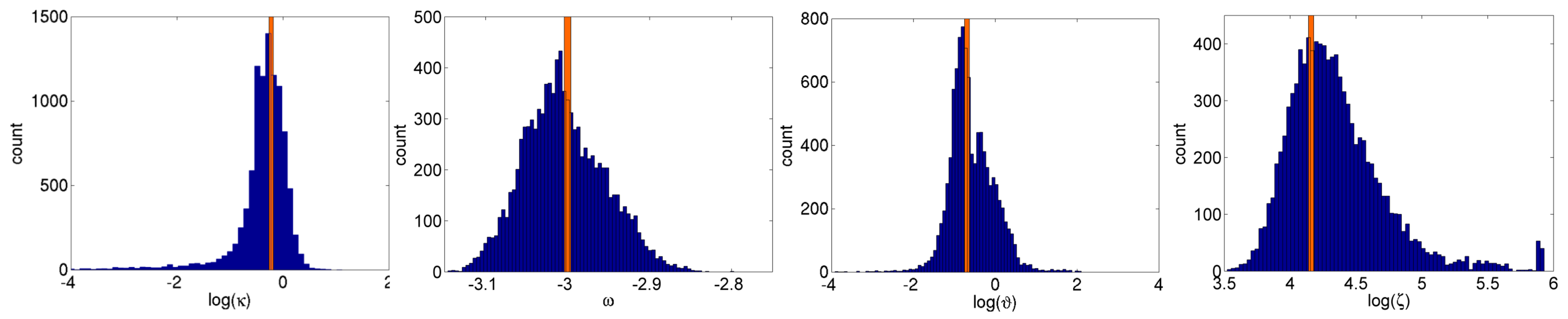


Fig. 5 | *Parameter estimation example.* *10'000 simulation runs generating 640 binary decisions (cf. Fig. 4) were completed under the same parameter setting, with moderate perceptual noise (α=0.01) and weak decision noise (ζ=64). Maximum-a-posteriori parameter estimates were then made for every simulation run. In all panels, the orange line represents the value the decisions were generated with.*

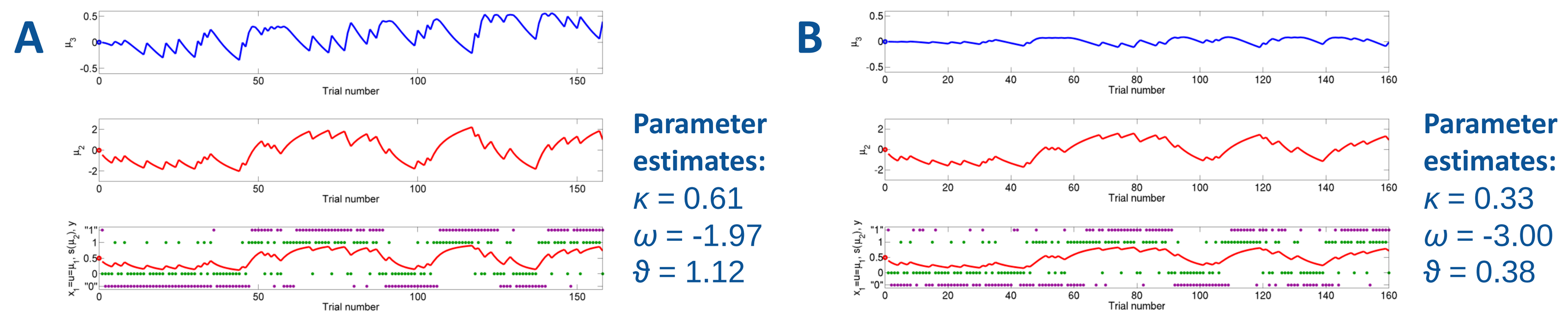


Fig. 6 | *Application to behavioral data.* *Owing to an assumed absence of perceptual noise (unambiguous cues and outcomes), the hidden true state, perceptual input, and belief on probability of true state = 1, are all equal (green); other elements are as in Fig. 4. (A) Healthy control subject. (B) Prodromal schizophrenic patient. Note that while, superficially, the patient shows similar learning and decision-making to the control (red lines and violet dots), more subtle differences appear at a higher level (blue lines) and in the parameter estimates.*

7 Conclusions

We have provided face validity of our Bayesian model for individual learning under uncertainty, showing that parameter estimates can be recovered under simultaneous perceptual and decisional uncertainty. The application to behavioral data has demonstrated the potential of our approach for detecting subtle individual differences in learning that would escape conventional analyses.