Learning complex tasks with probabilistic population codes

Richard E. Turner[1], Pietro Berkes[2] and József Fiser[2,3]

UNIVERSITY OF CAMBRIDGE

1) CBL, Dept Engineering, University of Cambridge, UK. 2) Volen Center for Complex Systems, Brandeis University. 3) Dept of Psychology and Neuroscience Program, Brandeis University

Probabilistic neural representations

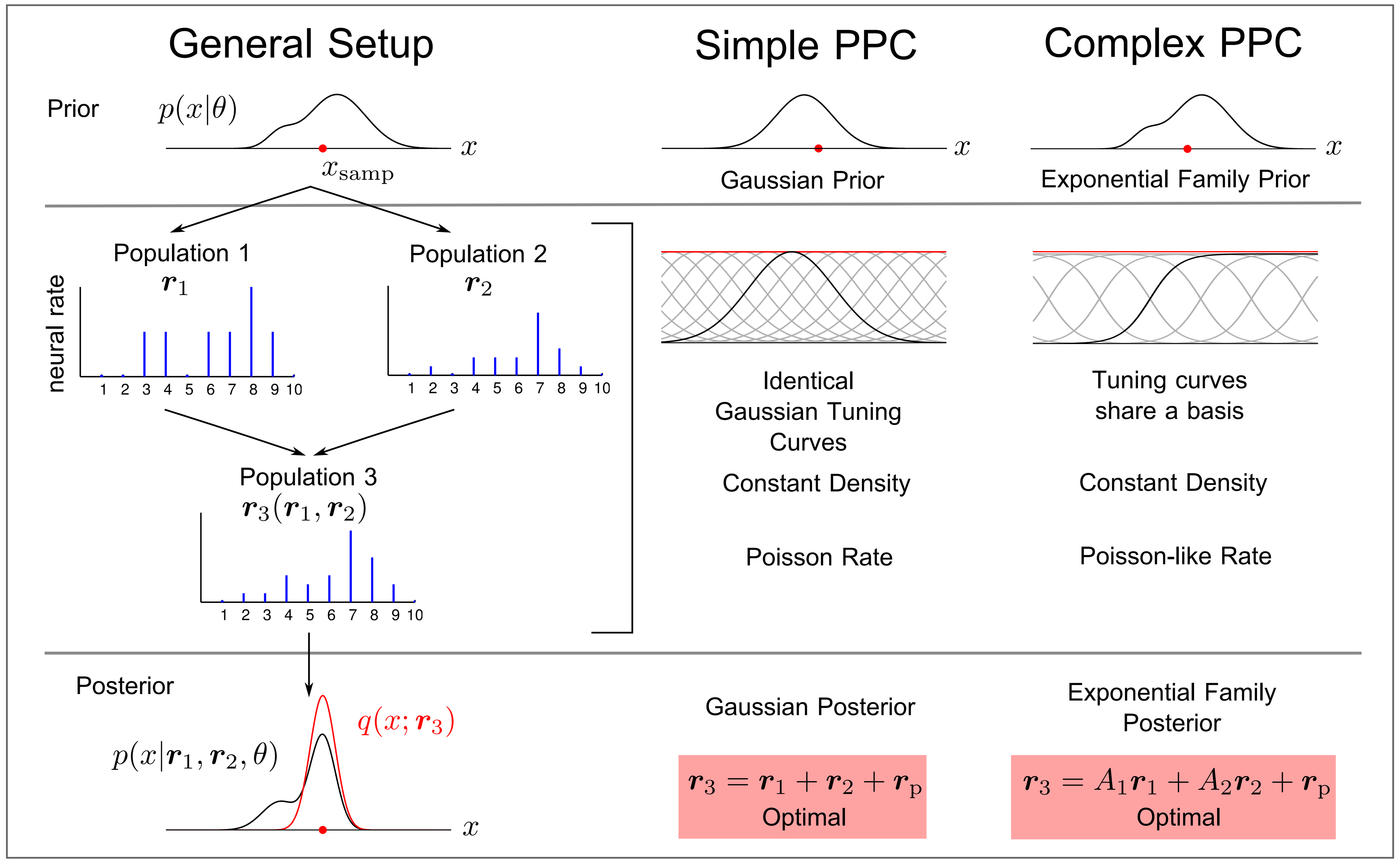
Sensory cues are combined according to:
uncertainty (Ernst and Banks, 2002) and the prior (Knill, 2002)
Priors can be learned (Körding and Wolpert, 2003)
Multiple, dependent variables (Knill, 2002)

How are these computations supported by cortex?
One candidate: probabilistic population codes (PPCs)

Outstanding issues with PPC

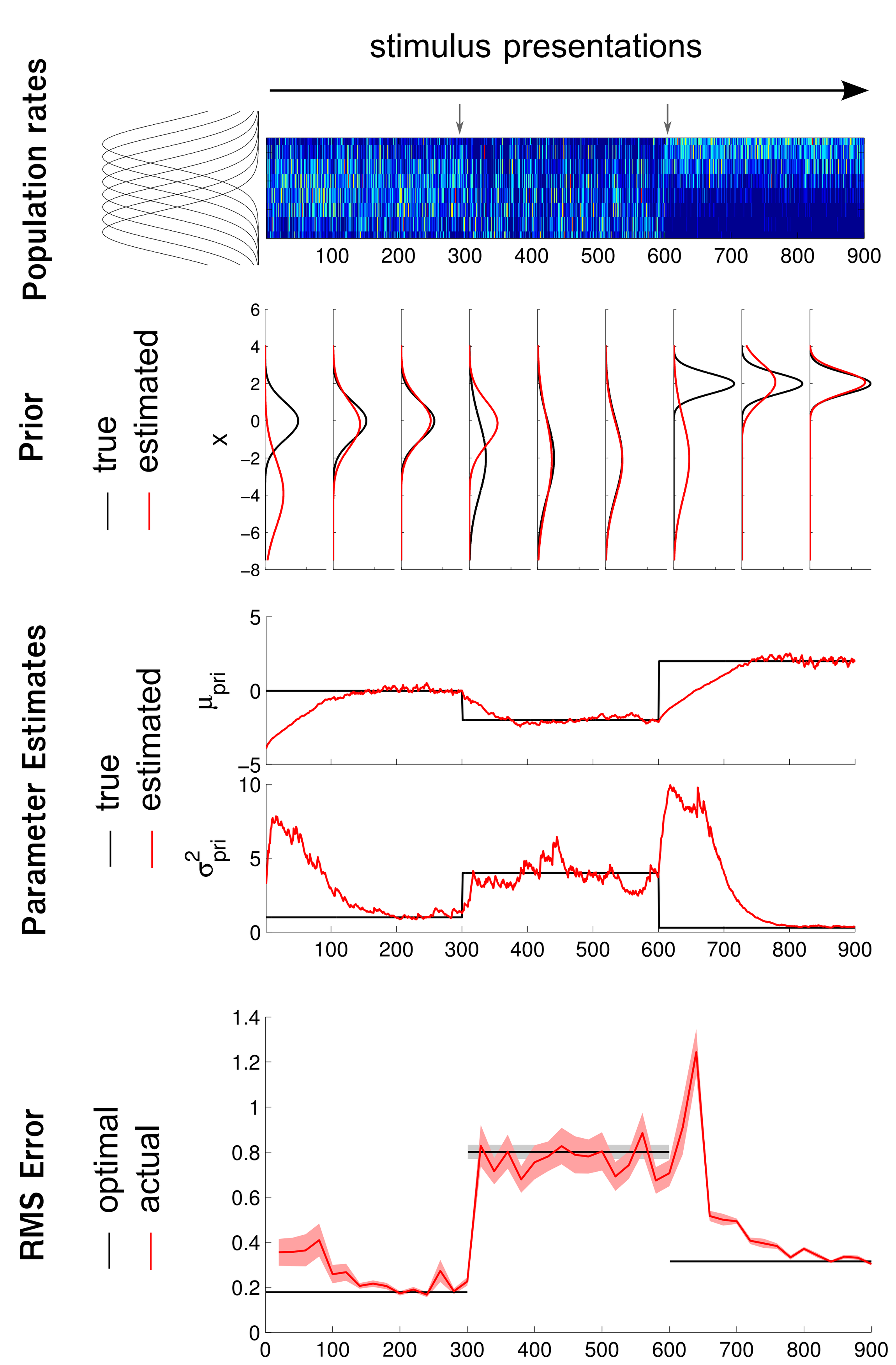
1) **No known learning method**
   Adapting to a changing prior
   Learning the connection strengths
2) Complex models: multimodal/non-linear
3) Encoding multiple dependent variables

Solution: Generalise PPC using vEM



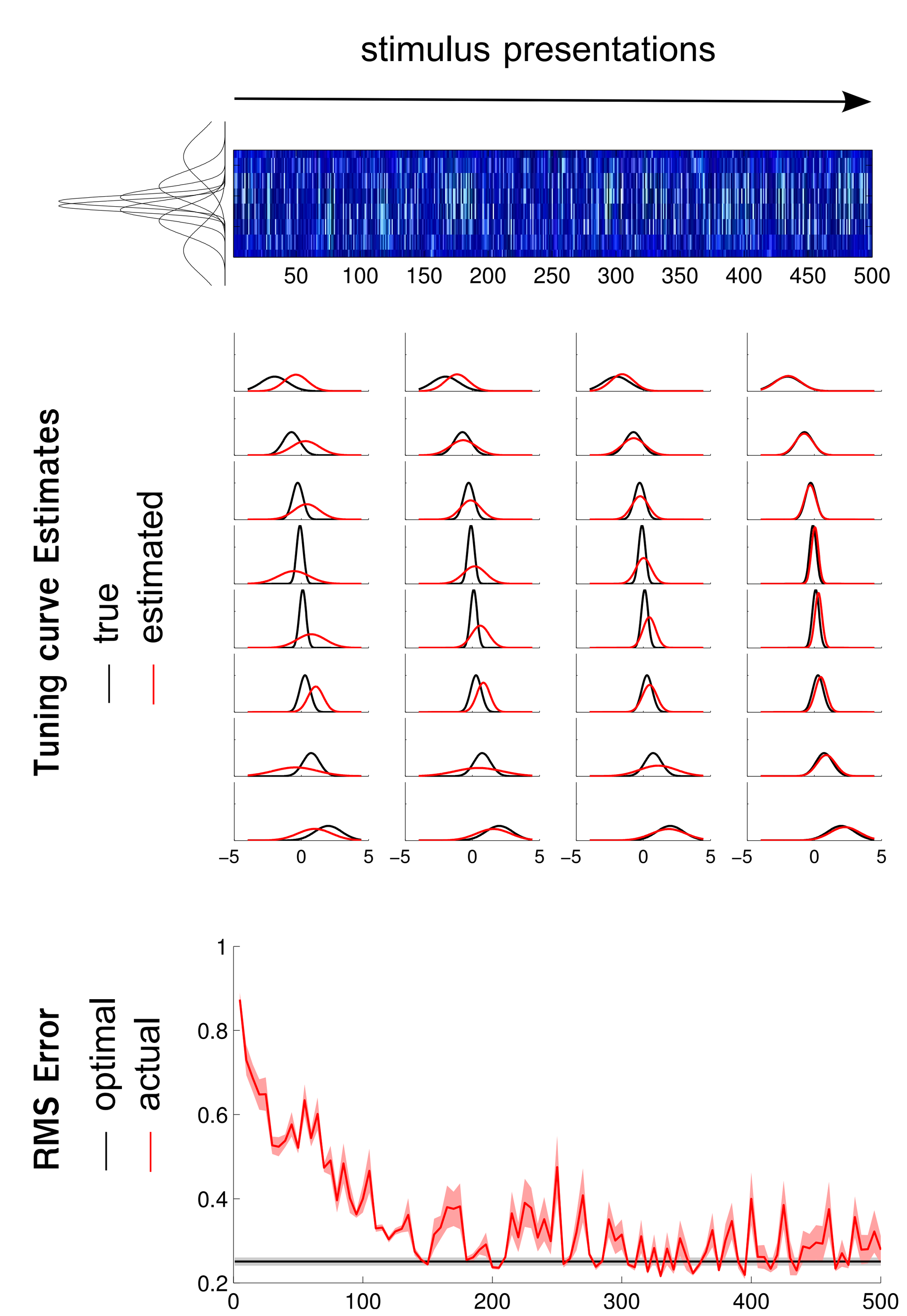
Simulations

Adapting to a changing prior



Parameters of the prior can be learned

Learning the connections



The network can be wired up in an unsupervised manner
Always requires non-linear E-Step

Model Details

Gaussian Prior

$$p(x|\theta) = \mathrm{Norm}(x; \mu_{\mathrm{pri}}, \sigma^2_{\mathrm{pri}})$$

Gaussian tuning curves

$$\lambda_{n,j}(x) = g_n \mathrm{Norm}(x; \mu_{n,j}, \sigma_n^2)$$

Poisson rates

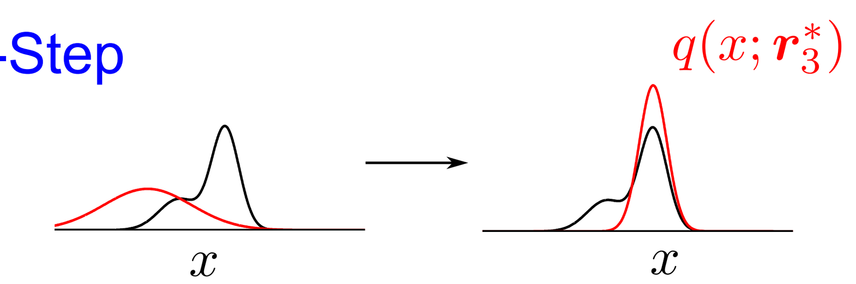
$$p(\mathrm{r}_{n,j}|x, \theta) = \mathrm{Poisson}(\mathrm{r}_{n,j}; \lambda_{n,j}(x))$$

Posterior distribution

$$p(x|\mathrm{R}, \theta) \propto \mathrm{Norm}(x; \mu_*(\mathrm{R}), \sigma^2_*(\mathrm{R})) \times \exp(-\sum_{n,j} \lambda_{n,j}(x))$$

Variational EM

E-Step



$$\mathbf{r}_3^* = \arg\min_{\mathbf{r}_3} \mathrm{KL}(q(x; \mathbf{r}_3) || p(x|\mathbf{r}_1, \mathbf{r}_2, \theta))$$

M-Step

Supervised-like: Fill in x using $q(x; \mathbf{r}_3^*)$

$$\theta_3^* = \arg\max_{\theta} \langle \log p(x, \mathbf{r}_1, \mathbf{r}_2, \theta) \rangle_{q(x;\mathbf{r}_3)}$$

vEM Interpretation of PPC

$$q(x; \mathbf{r}_3) = \mathrm{Norm}(x; \mu_{\mathrm{q}}(\mathbf{r}_3), \sigma^2_{\mathrm{q}}(\mathbf{r}_3))$$

$$\frac{\mu_{\mathrm{q}}(\mathbf{s})}{\sigma^2_{\mathrm{q}}(\mathbf{r}_3)} = \sum_m \mathbf{r}_{3,m} \frac{\mu_m}{\sigma_m^2} \quad \frac{1}{\sigma^2_{\mathrm{q}}(\mathbf{r}_3)} = \sum_m \frac{\mathbf{r}_{3,m}}{\sigma_m^2}$$

E-Step: usual PPC update, modified if sum of tuning curves not constant

$$\tau \frac{d\mathbf{r}_3}{dt} = f(A_1\mathbf{r}_1 + A_2\mathbf{r}_2, B\mathbf{r}_3)$$

M-Step: On-line learning rules

Conclusions

Extended PPC to complex learning tasks
  Adapting to a changing prior and learning the network connection strengths
Viewed PPC as encoding an approximate representation of the posterior
  Connected PPC to vEM which then provides learning rules
Some of the attractive properties of PPC have to be sacrificed
  Updates become non-linear in general and gains must be inferred/learned
Extendedable to more complex models, but neurally plausible implementations will require further approximations

References

Ernst and Banks, Humans integrate visual and haptic information in a statistically optimal fashion, Nature, 2002
Körding and Wolpert, Bayesian integration in sensorimotor learning, Nature, 2004
Knill, Mixture models and the probabilistic structure of depth cues, Vision Research, 2003