Mismatch-Negativity in Somatosensory Cortex: *Theoretical Modelling*

Robert T Lange, Miro Grundei, Sam Gjisem,
Dirk Ostwald, Felix Blankenburg
rtl17@ic.ac.uk
www.rob-lange.com
github.com/RobertTLange/SequentialBayesianLearning

CCNB - FU Berlin

November 16, 2018

Sequence Generating Process

- Design objetives:
  1. Hidden state $s \in \mathcal{S}$ drives dynamics of sampled observation sequence $o_1, o_2, \dots, o_T$
  2. Rowing paradigm: Observation feature does not determine normal/deviant definition
  3. Include "catch" observation/trial to include attention coverage
  4. 2nd order dependencies: Combats problem of fast alternating pairs of observations being **together** perceived as standard.
- Markov Dependencies:
  1. 1st order: $p^{(s_t)}_{o_t|o_{t-1}} := p(o_t|o_{t-1}, s_t)$
  2. 2nd order: $p^{(s_t)}_{o_t|o_{t-1},o_{t-2}} := p(o_t|o_{t-1}, o_{t-2}, s_t)$
- p^{catch}: Probability of catch trial
- $p^{reg-change}$: Probability of regime change

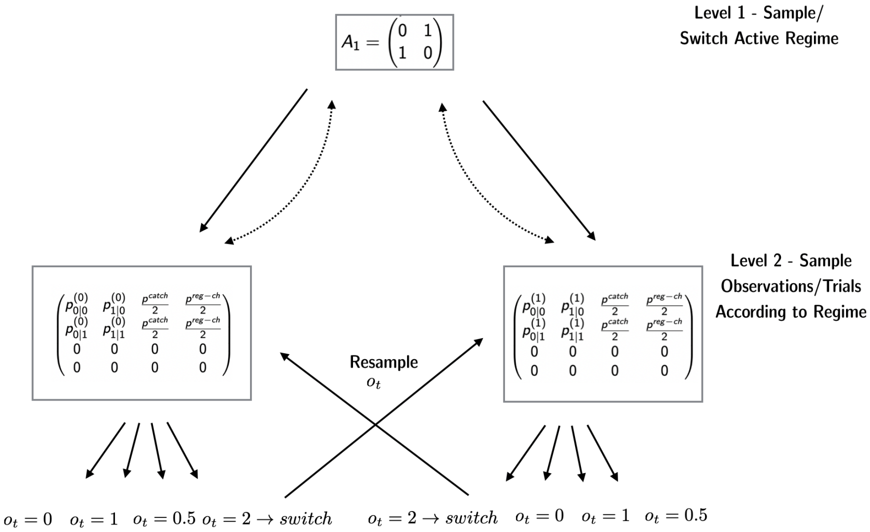
1st order Markov Dependencies in HHMM

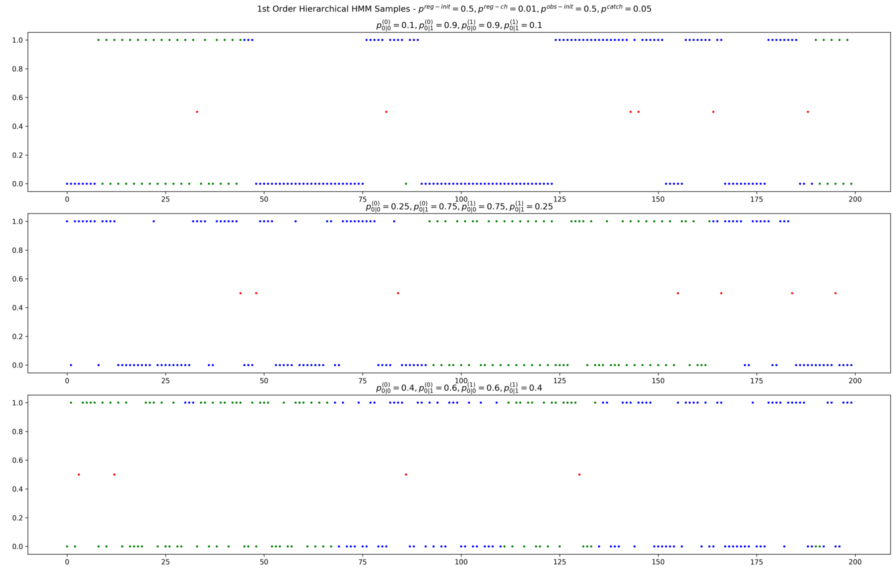
- 1st Level (Regime Sampling/Switching):
  $\pi_1 := (p^{reg-init}, 1-p^{reg-init})$; $A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
- 2nd Level (Observation Sampling):

$$\pi_1 := (p^{obs-init}, 1-p^{obs-init}, 0, 0)$$

$$A_{2.0} = \begin{pmatrix} p_{0|0}^{(0)} & p_{1|0}^{(0)} & \frac{p^{catch}}{2} & \frac{p^{reg-ch}}{2} \\ p_{0|1}^{(0)} & p_{1|1}^{(0)} & \frac{p^{catch}}{2} & \frac{p^{reg-ch}}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; A_{2.1} = \begin{pmatrix} p_{0|0}^{(1)} & p_{1|0}^{(1)} & \frac{p^{catch}}{2} & \frac{p^{reg-ch}}{2} \\ p_{0|1}^{(1)} & p_{1|1}^{(1)} & \frac{p^{catch}}{2} & \frac{p^{reg-ch}}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

- 8 P.: $p^{reg-init}, p^{obs-init}, p^{reg-ch}, p^{catch}, p_{0|0}^{(i)}, p_{0|1}^{(i)}, i \in \{0, 1\}$
- So far - symmetric (6 P.): $p_{0|0}^{(i)} = p_{1|1}^{(i)}$ and $p_{j|k}^{(0)} = p_{k|j}^{(1)}$


$A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
Level 1 - Sample/ Switch Active Regime
Level 2 - Sample Observations/Trials According to Regime
$\begin{pmatrix} p_{0|0}^{(0)} & p_{1|0}^{(0)} & \frac{p^{catch}}{2} & \frac{p^{reg-ch}}{2} \\ p_{0|1}^{(0)} & p_{1|1}^{(0)} & \frac{p^{catch}}{2} & \frac{p^{reg-ch}}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$
$\begin{pmatrix} p_{0|0}^{(1)} & p_{1|0}^{(1)} & \frac{p^{catch}}{2} & \frac{p^{reg-ch}}{2} \\ p_{0|1}^{(1)} & p_{1|1}^{(1)} & \frac{p^{catch}}{2} & \frac{p^{reg-ch}}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$
Resample o_t
$o_t = 0$ $o_t = 1$ $o_t = 0.5$ $o_t = 2 \rightarrow switch$
$o_t = 2 \rightarrow switch$ $o_t = 0$ $o_t = 1$ $o_t = 0.5$


1st Order Hierarchical HMM Samples - $p^{reg-init} = 0.5, p^{reg-ch} = 0.01, p^{obs-init} = 0.5, p^{catch} = 0.05$
$p_{0|0}^{(0)} = 0.1, p_{0|1}^{(0)} = 0.9, p_{0|0}^{(1)} = 0.9, p_{0|1}^{(1)} = 0.1$
$p_{0|0}^{(0)} = 0.25, p_{0|1}^{(0)} = 0.75, p_{0|0}^{(1)} = 0.75, p_{0|1}^{(1)} = 0.25$
$p_{0|0}^{(0)} = 0.4, p_{0|1}^{(0)} = 0.6, p_{0|0}^{(1)} = 0.6, p_{0|1}^{(1)} = 0.4$
0.0 0.2 0.4 0.6 0.8 1.0
0 25 50 75 100 125 150 175 200

2nd order Markov Dependencies in HHMM I

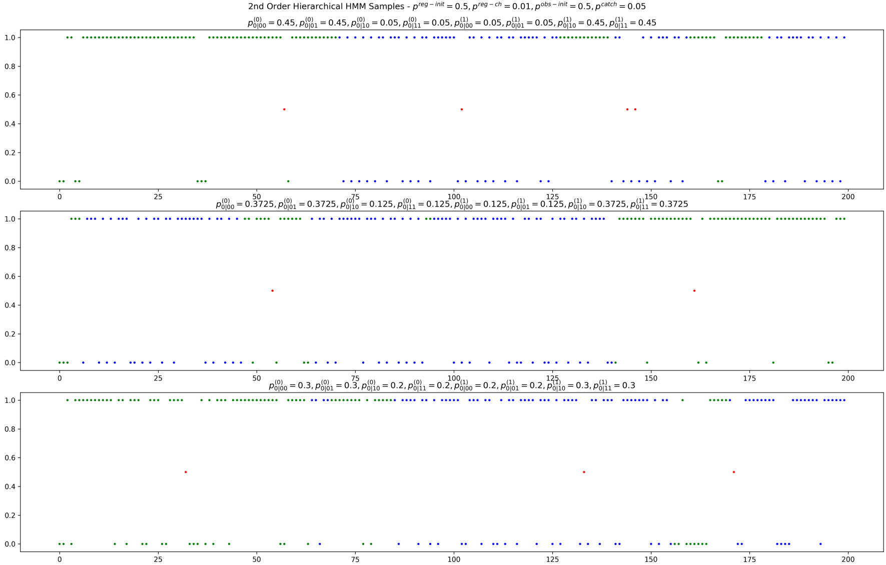
- First level same $\rightarrow$ deterministic regime transition matrix

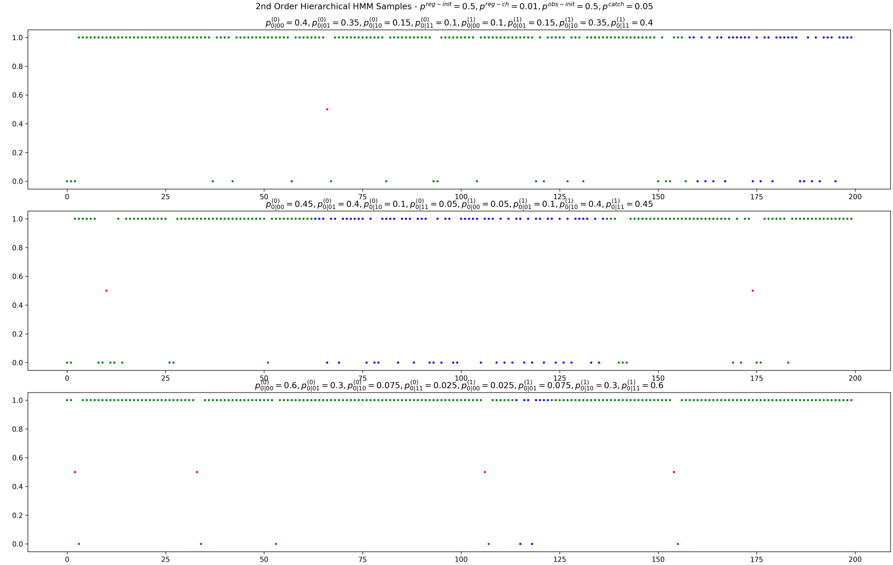
$$A_{2.i} = \begin{array}{c} \\ o_{t-1}{=}0,\ o_{t-2}{=}0 \\ o_{t-1}{=}0,\ o_{t-2}{=}1 \\ o_{t-1}{=}1,\ o_{t-2}{=}0 \\ o_{t-1}{=}1,\ o_{t-2}{=}1 \\ o_{t-1}{=}2,\ o_{t-2}{=}1 \\ \ddots \\ o_{t-1}{=}3,\ o_{t-2}{=}3 \end{array} \begin{array}{c} \begin{array}{cccc} o_t{=}0 & o_t{=}1 & \text{Catch} & \text{Reg change} \end{array} \\ \begin{pmatrix} p^{(i)}_{0|00} & p^{(i)}_{1|00} & p^{catch}/4 & p^{reg-ch}/4 \\ p^{(i)}_{0|01} & p^{(i)}_{1|01} & p^{catch}/4 & p^{reg-ch}/4 \\ p^{(i)}_{0|10} & p^{(i)}_{1|01} & p^{catch}/4 & p^{reg-ch}/4 \\ p^{(i)}_{0|11} & p^{(i)}_{1|11} & p^{catch}/4 & p^{reg-ch}/4 \\ 0 & 0 & 0 & 0 \\ \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \in \mathbb{R}^{16\times 4}$$

- 12 P.: $p^{reg-init}, p^{obs-init}, p^{reg-ch}, p^{catch}, p^{(i)}_{0|00}, p^{(i)}_{0|01}, p^{(i)}_{0|10}, p^{(i)}_{0|11}$

2nd order Markov Dependencies in HHMM II

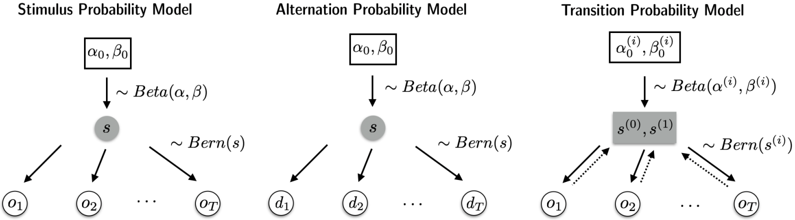
- 0th: $p(o_t|s_t = i) = \sum_{j,k} p(o_t|o_{t-1} = j, o_{t-2} = k) = \sum_{j,k} p^{(i)}_{o_t|jk}$
- 1st: $p(o_t|o_{t-1}, s_t = i) = \begin{cases} p^{(i)}_{o_t|10} + p^{(i)}_{o_t|11}, & \text{for } o_{t-1} = 1 \\ p^{(i)}_{o_t|00} + p^{(i)}_{o_t|01}, & \text{for } o_{t-1} = 0 \end{cases}$
- On which dependency level do we define fast and slow?
  - 1st: $p^{(s_t)}_{o_t|o_{t-1}} := p(o_t|o_{t-1}, s_t) \rightarrow$ (0): slow, (1): fast
    - $p^{(0)}_{0|00} + p^{(0)}_{0|01} > 0.5 > p^{(1)}_{0|00} + p^{(1)}_{0|01}$
    - $p^{(0)}_{1|10} + p^{(0)}_{1|11} > 0.5 > p^{(1)}_{1|10} + p^{(1)}_{1|11}$
  - 2nd: $p^{(s_t)}_{o_t|o_{t-1},o_{t-2}} := p(o_t|o_{t-1}, o_{t-2}, s_t)$
    - $p^{(0)}_{0|00} > p^{(1)}_{0|00}$ and $p^{(0)}_{0|01} > p^{(1)}_{0|01}$
    - $p^{(0)}_{1|10} > p^{(1)}_{1|10}$ and $p^{(0)}_{1|11} > p^{(1)}_{1|11}$


2nd Order Hierarchical HMM Samples - $p^{reg-init} = 0.5, p^{reg-ch} = 0.01, p^{obs-init} = 0.5, p^{catch} = 0.05$
$p^{(0)}_{0|00} = 0.45, p^{(0)}_{0|01} = 0.45, p^{(0)}_{0|10} = 0.05, p^{(0)}_{0|11} = 0.05, p^{(1)}_{0|00} = 0.05, p^{(1)}_{0|01} = 0.05, p^{(1)}_{0|10} = 0.45, p^{(1)}_{0|11} = 0.45$
$p^{(0)}_{0|00} = 0.3725, p^{(0)}_{0|01} = 0.3725, p^{(0)}_{0|10} = 0.125, p^{(0)}_{0|11} = 0.125, p^{(1)}_{0|00} = 0.125, p^{(1)}_{0|01} = 0.125, p^{(1)}_{0|10} = 0.3725, p^{(1)}_{0|11} = 0.3725$
$p^{(0)}_{0|00} = 0.3, p^{(0)}_{0|01} = 0.3, p^{(0)}_{0|10} = 0.2, p^{(0)}_{0|11} = 0.2, p^{(1)}_{0|00} = 0.2, p^{(1)}_{0|01} = 0.2, p^{(1)}_{0|10} = 0.3, p^{(1)}_{0|11} = 0.3$


2nd Order Hierarchical HMM Samples - $p^{reg-init} = 0.5, p^{reg-ch} = 0.01, p^{obs-init} = 0.5, p^{catch} = 0.05$
$p^{(0)}_{0|00} = 0.4, p^{(0)}_{0|01} = 0.35, p^{(0)}_{0|10} = 0.15, p^{(0)}_{0|11} = 0.1, p^{(1)}_{0|00} = 0.1, p^{(1)}_{0|01} = 0.15, p^{(1)}_{0|10} = 0.35, p^{(1)}_{0|11} = 0.4$
$p^{(0)}_{0|00} = 0.45, p^{(0)}_{0|01} = 0.4, p^{(0)}_{0|10} = 0.1, p^{(0)}_{0|11} = 0.05, p^{(1)}_{0|00} = 0.05, p^{(1)}_{0|01} = 0.1, p^{(1)}_{0|10} = 0.4, p^{(1)}_{0|11} = 0.45$
$p^{(0)}_{0|00} = 0.6, p^{(0)}_{0|01} = 0.3, p^{(0)}_{0|10} = 0.075, p^{(0)}_{0|11} = 0.025, p^{(1)}_{0|00} = 0.025, p^{(1)}_{0|01} = 0.075, p^{(1)}_{0|10} = 0.3, p^{(1)}_{0|11} = 0.6$

SBL: Beta-Bernoulli Agent

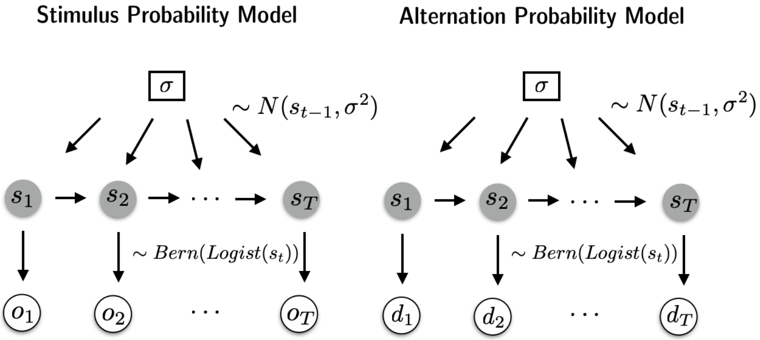
- $\mathcal{O} := \{0, 1\}, \mathcal{S} := [0, 1]$
- $p(s_t|s_{t-1}) := \delta_{s_{t-1}}(s_t)$
- $d_t := \mathbf{1}_{o_t \neq o_{t-1}}$
- SP: $p(o_t|s_t)$
- AP: $p(d_t|s_t)$
- TP: $p(o_t|o_{t-1}, s_t)$



- Limited memory via exponential decay in parameter updates
- Closed-form posterior/suprise via Beta-Bernoulli conjugacy

SBL: Gaussian Random Walk Agent

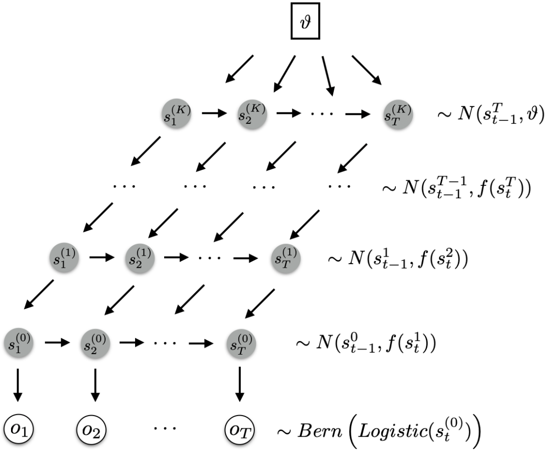
- $\mathcal{O} = \{0, 1\}, \mathcal{S} = \mathcal{R}$
- $p(s_t|s_{t-1}) = \mathcal{N}(s_{t-1}, \sigma^2)$
- SP: $p(o_t|s_t) = Bern(Logistic(s_t))$
- AP: $p(d_t|s_t) = Bern(Logistic(s_t))$



- Numerical inversion: Discrete resolution and restricted support

SBL: Gaussian Hierarchical Filter Agent

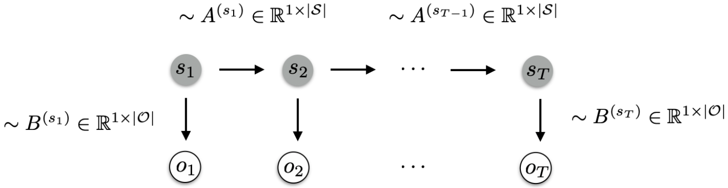
- $\mathcal{O} = \{0, 1\}, \mathcal{S} = \mathcal{R}^K$
- Hidden states: $p(s_t^{(k)} | s_{t-1}^k, s_t^{(k+1)}) = \mathcal{N}(s_{t-1}^k, f(s_t^{(k+1)}))$
- Observations: $p(o_t | s_t^{(0)}) = Bern(Logistic(s_t^{(0)}))$



SBL: Hidden Markov Model Agent

- $\mathcal{O} = \{0,1\}, \mathcal{S} = \{1,2,\ldots,K\}$
- Hidden state transitions: $p(s_t|s_{t-1}) = A \in \mathbb{R}^{|\mathcal{S}|\times|\mathcal{S}|}$
- Observations: $p(o_t|s_t) = B \in \mathbb{R}^{|\mathcal{S}|\times|\mathcal{O}|}$

$$p(o_t|o_1,\ldots o_{t-1}) = \sum_{s_t=1}^{K} \underbrace{p(o_t|s_t)}_{\text{Emissions}} \underbrace{p(s_t|s_{t-1})}_{\text{Transitions}} \underbrace{p(s_{t-1}|o_1,\ldots,o_{t-1})}_{\text{Filtering}}$$



- Fitting via Expectation-Maximization
- Hidden state inference via decoding (Viterbi algorithm)

Decoding (in the latent variable sense)

- Decoding (e.g. Viterbi algorithm) in the latent variable literature refers to hidden state inference!

$$arg \max_{s_1,\ldots,s_t} p(s_1,\ldots,s_t|o_1,\ldots,o_t)$$

- Form of dynamic programming solution based on forward-backwards solution to EM