EM algorithm derivations and supplementary data

1 Algorithm derivations
1.1 Derivation of E-step filter updates
1.1.1 Probability mass function of \( m_k \)

\[
P(m_k | x_k) = \frac{m_k}{p_k} (1 - p_k)^{1-m_k}
\]

\[
= \exp \left\{ \log \left[ \frac{m_k}{p_k} (1 - p_k)^{1-m_k} \right] \right\}
\]

\[
= \exp \left[ m_k \log(p_k) + (1 - m_k) \log(1 - p_k) \right]
\]

\[
= \exp \left[ m_k \log \left( \frac{p_k}{1 - p_k} \right) + \log(1 - p_k) \right]
\]

\[
= \exp \left[ m_k (\beta_0 + \beta_1 x_k) + \log \left( \frac{1}{1 + e^{\beta_0 + \beta_1 x_k}} \right) \right] \tag{1}
\]

1.1.2 Mean and variance of the posterior density function \( p(x_k | y^k) \)

We follow an approach similar to [1] in deriving the filter updates. Recall that linear models are assumed to relate sympathetic arousal to the phasic-derived and tonic components.

\[
r_k = \gamma_0 + \gamma_1 x_k + v_k \tag{2}
\]

\[
s_k = \delta_0 + \delta_1 x_k + w_k \tag{3}
\]

We take the two noise terms \( v_k \) and \( w_k \) to be independent of each other. Consequently, the density functions \( p(r_k | x_k) \) and \( p(s_k | x_k) \) conditioned on already having observed \( x_k \) are independent of each other in the following derivation.

\[
p(x_k | y^k) = \frac{p(x_k | y^{k-1})p(y_k | x_k)}{p(y_k | y^{k-1})}
\]

\[
= \frac{p(x_k | y^{k-1})P(m_k | x_k)p(r_k | x_k)p(s_k | x_k)P(n_{k,1:t} | x_k)}{p(y_k | y^{k-1})}
\]

\[
\propto \exp \left[ \frac{-(x_k - x_k) \|k-1 \|^2}{2\sigma_{v}\|k-1 \|^2} + m_k \log \left( \frac{p_k}{1 - p_k} \right) + \log(1 - p_k) \right.
\]

\[
- \left. \frac{(r_k - \gamma_0 - \gamma_1 x_k)^2}{2\sigma_{v}^2} - \frac{(s_k - \delta_0 - \delta_1 x_k)^2}{2\sigma_{w}^2} + \sum_{j=1}^{J} \log(\lambda_{k,j} \Delta) n_{k,j} - \lambda_{k,j} \Delta \right] \tag{4}
\]

\[
\log \left[ p(x_k | y^k) \right] = \frac{-(x_k - x_k) \|k-1 \|^2}{2\sigma_{v}^2} + m_k \log \left( \frac{p_k}{1 - p_k} \right) + \log(1 - p_k) \]

\[
- \frac{(r_k - \gamma_0 - \gamma_1 x_k)^2}{2\sigma_{v}^2} - \frac{(s_k - \delta_0 - \delta_1 x_k)^2}{2\sigma_{w}^2} + \sum_{j=1}^{J} \log(\lambda_{k,j} \Delta) n_{k,j} - \lambda_{k,j} \Delta \right] \quad \text{+ const} \tag{5}
\]
We take the partial derivative of the logarithm term above and set it to 0 to solve for the mean.

\[
\frac{\partial}{\partial x_k} \log \left[ p(x_k|y^k) \right] = \frac{(x_k - x_{k|k-1})}{\sigma_{k|k-1}^2} + \beta_1 (m_k - p_k) + \frac{\gamma_1 (r_k - \gamma_0 - \gamma_1 x_k)}{\sigma_v^2} + \delta_1 (s_k - \delta_0 - \delta_1 x_k) + \sum_{j=1}^J \frac{1}{\lambda_{k,j|k}} \frac{\partial \lambda_{k,j|k}}{\partial x_k} (n_{k,j} - \lambda_{k,j|k} \Delta) = 0. \tag{6}
\]

Solving for \( x_k \) in the equation above provides the filter update for \( x_{k|k} \). We have taken,

\[
\frac{\partial p_k}{\partial x_k} = \beta_1 p_k (1 - p_k) \tag{7}
\]

when calculating the partial derivative. Similarly, the second partial derivative is,

\[
\frac{\partial^2}{\partial x_k^2} \log \left[ p(x_k|y^k) \right] = -\frac{1}{\sigma_{k|k-1}^2} - \beta_1^2 p_k (1 - p_k) - \frac{\gamma_1^2}{\sigma_v^2} - \frac{\delta_1^2}{\sigma_w^2} + \sum_{j=1}^J \frac{1}{\lambda_{k,j|k}} \frac{\partial \lambda_{k,j|k}}{\partial x_k} (n_{k,j} - \lambda_{k,j|k} \Delta) + \sum_{j=1}^J \left[ \frac{1}{\lambda_{k,j|k}} \frac{\partial^2 \lambda_{k,j|k}}{\partial x_k^2} (n_{k,j} - \lambda_{k,j|k} \Delta) - \frac{n_{k,j}}{\lambda_{k,j|k}^2} \left( \frac{\partial \lambda_{k,j|k}}{\partial x_k} \right)^2 \right]. \tag{8}
\]

The filter update for \( \sigma_{k|k}^2 \) is given by [1],

\[
\sigma_{k|k}^2 = \left\{ - \frac{\partial^2}{\partial x_k^2} \log \left[ p(x_k|y^k) \right] \right\}^{-1}. \tag{9}
\]

### 1.2 Derivation of the M-step updates

#### 1.2.1 Complete data log-likelihood

Taking \( \mathcal{X}^K = \{x_1, x_2, \ldots, x_K\} \), the complete data likelihood conditioned on the model parameters \( \Theta \) is given by,

\[
p(Y^K, \mathcal{X}^K | \Theta) = \prod_{k=1}^K p_k^{m_k} (1 - p_k)^{1-m_k} \times \prod_{k=1}^K \frac{1}{\sqrt{2\pi \sigma_v^2}} e^{-\frac{(x_k - \gamma_0 - \gamma_1 x_k)^2}{2\sigma_v^2}} \
\times \prod_{k=1}^K \frac{1}{\sqrt{2\pi \sigma_w^2}} e^{-\frac{(s_k - \delta_0 - \delta_1 x_k)^2}{2\sigma_w^2}} \times \prod_{k=1}^K e^{\sum_{j=1}^J \log(\lambda_{k,j}) n_{k,j} - \lambda_{k,j} \Delta} \
\times \prod_{k=1}^K \frac{1}{\sqrt{2\pi \sigma_\varepsilon^2}} e^{-\frac{(x_k - \rho x_{k-1} - \alpha_1 x_k)^2}{2\sigma_\varepsilon^2}} . \tag{10}
\]

The expected log-likelihood is,

\[
Q = \sum_{k=1}^K \mathbb{E} \left[ m_k (\beta_0 + \beta_1 x_k) - \log \left( 1 + e^{\beta_0 + \beta_1 x_k} \right) \right] + \frac{(-K)}{2} \log \left( 2\pi \sigma_v^2 \right)
\]
\[-\sum_{k=1}^{K} \mathbb{E}\left[\frac{(r_k - \gamma_0 - \gamma_1 x_k)^2}{2\sigma_v^2}\right] + \frac{(-K)}{2} \log (2\pi\sigma_v^2) - \sum_{k=1}^{K} \mathbb{E}\left[\frac{(s_k - \delta_0 - \delta_1 x_k)^2}{2\sigma_w^2}\right] + \sum_{k=1}^{K} \sum_{j=1}^{J} \mathbb{E}\left[\log(\lambda_{k,j} \Delta) n_{k,j} - \lambda_{k,j} \Delta\right] + \frac{(-K)}{2} \log (2\pi\sigma_w^2) - \sum_{k=1}^{K} \mathbb{E}\left[\frac{(x_k - \rho x_{k-1} - \alpha I_k)^2}{2\sigma^2}\right]. \tag{11}\]

Following [2], we take
\[
x_k|K = \mathbb{E}\left[ x_k | Y^K, \Theta \right] \tag{12} \]
\[
U_k = \mathbb{E}\left[ x_k^2 | Y^K, \Theta \right] \tag{13} \]
\[
U_{k,k+1} = \mathbb{E}\left[ x_k x_{k+1} | Y^K, \Theta \right]. \tag{14} \]

1.2.2 M-step updates for $\alpha$ and $\rho$

Let $Q_1$ denote the term in $Q$ that contains $\alpha$ and $\rho$.

\[
Q_1 = \frac{1}{2\sigma^2} \sum_{k=1}^{K} \mathbb{E}\left[ (x_k - \rho x_{k-1} - \alpha I_k)^2 \right] \tag{15} \]

While it is possible to determine the starting state $x_0$ as a separate parameter, we follow one of the options in [2, 3] and set $x_0 = x_1$. This permits some bias at the beginning. Therefore,

\[
Q_1 = \frac{1}{2\sigma^2} \left\{ \sum_{k=2}^{K} \mathbb{E}\left[ (x_k - \rho x_{k-1} - \alpha I_k)^2 \right] + \mathbb{E}\left[ (\alpha I_1)^2 \right] \right\}. \tag{16} \]

We take the partial derivatives of $Q_1$ with respect to $\alpha$ and $\rho$ and set them to 0 to obtain the M-step updates.

\[
\frac{\partial Q_1}{\partial \alpha} = \sum_{k=2}^{K} \mathbb{E}\left[ -2I_k (x_k - \rho x_{k-1} - \alpha I_k) \right] + 2\alpha I_1^2
\]

\[
0 = - \sum_{k=2}^{K} I_k \mathbb{E}\left[ x_k \right] + \rho \sum_{k=2}^{K} I_k \mathbb{E}\left[ x_{k-1} \right] + \alpha \sum_{k=1}^{K} I_k^2
\]

\[
= - \sum_{k=2}^{K} I_k x_k|K + \rho \sum_{k=2}^{K} I_k x_{k-1}|K + \alpha \sum_{k=1}^{K} I_k^2 \tag{17} \]

\[
\frac{\partial Q_1}{\partial \rho} = \sum_{k=2}^{K} \mathbb{E}\left[ -2x_{k-1} (x_k - \rho x_{k-1} - \alpha I_k) \right]
\]

\[
0 = - \sum_{k=2}^{K} \mathbb{E}\left[ x_k x_{k-1} \right] + \rho \sum_{k=2}^{K} \mathbb{E}\left[ x_k^2 \right] + \alpha \sum_{k=2}^{K} I_k \mathbb{E}\left[ x_{k-1} \right]
\]

\[
= - \sum_{k=1}^{K-1} U_{k,k+1} + \rho \sum_{k=1}^{K-1} U_k + \alpha \sum_{k=2}^{K-1} I_k x_{k-1}|K \tag{18} \]

The solutions to these simultaneous equations provide $\alpha$ and $\rho$. 

3
1.2.3 M-step updates for $\gamma_0$, $\gamma_1$, $\delta_0$ and $\delta_1$

Let $Q_2$ denote the term in $Q$ containing $\gamma_0$ and $\gamma_1$.

$$Q_2 = \sum_{k=1}^{K} \frac{\mathbb{E}[(r_k - \gamma_0 - \gamma_1 x_k)^2]}{2\sigma^2_v}$$

(20)

Taking the partial derivatives with respect to $\gamma_0$ and $\gamma_1$ yields,

$$\frac{\partial Q_2}{\partial \gamma_0} = \sum_{k=1}^{K} -2\mathbb{E}[r_k - \gamma_0 - \gamma_1 x_k]$$

$$0 = -\sum_{k=1}^{K} r_k + \gamma_0 K + \gamma_1 \sum_{k=1}^{K} \mathbb{E}[x_k]$$

$$= -\sum_{k=1}^{K} r_k + \gamma_0 K + \gamma_1 \sum_{k=1}^{K} x_k \pi_1$$

(21)

$$\frac{\partial Q_2}{\partial \gamma_1} = \sum_{k=1}^{K} -2\mathbb{E}[x_k (r_k - \gamma_0 - \gamma_1 x_k)]$$

$$0 = -\sum_{k=1}^{K} r_k \mathbb{E}[x_k] + \gamma_0 \sum_{k=1}^{K} \mathbb{E}[x_k] + \gamma_1 \sum_{k=1}^{K} \mathbb{E}[x_k^2]$$

$$= -\sum_{k=1}^{K} r_k x_k \pi_1 + \gamma_0 \sum_{k=1}^{K} x_k \pi_1 + \gamma_1 \sum_{k=1}^{K} U_k.$$  

(22)

The solutions to these simultaneous equations provide $\gamma_0$ and $\gamma_1$. $\delta_0$ and $\delta_1$ may be obtained similarly from the term in $Q$ containing $s_k$.

1.2.4 M-step updates for $\sigma^2_v$ and $\sigma^2_w$

Let $Q_3$ denote the term in $Q$ containing $\sigma^2_v$.

$$Q_3 = -\frac{K}{2} \log (2\pi \sigma^2_v) - \sum_{k=1}^{K} \frac{\mathbb{E}[(r_k - \gamma_0 - \gamma_1 x_k)^2]}{2\sigma^2_v}$$

(23)

We take the partial derivative with respect to $\sigma^2_v$ and set it to 0.

$$\frac{\partial Q_3}{\partial \sigma^2_v} = -\frac{K}{2\sigma^2_v} + \frac{1}{2\sigma^2_v} \sum_{k=1}^{K} \mathbb{E}[(r_k - \gamma_0 - \gamma_1 x_k)^2] = 0$$

(24)

$$\sigma^2_v = \frac{1}{K} \sum_{k=1}^{K} \mathbb{E}[(r_k - \gamma_0 - \gamma_1 x_k)^2]$$

$$= \frac{1}{K} \left\{ \sum_{k=1}^{K} r_k^2 + \gamma_0^2 + \gamma_1^2 \sum_{k=1}^{K} \mathbb{E}[x_k^2] - 2\gamma_0 \sum_{k=1}^{K} r_k \mathbb{E}[x_k] + 2\gamma_0 \gamma_1 \sum_{k=1}^{K} \mathbb{E}[x_k] \right\}$$

$$= \frac{1}{K} \left\{ \sum_{k=1}^{K} r_k^2 + \gamma_0^2 + \gamma_1^2 \sum_{k=1}^{K} U_k - 2\gamma_0 \sum_{k=1}^{K} r_k \mathbb{E}[x_k] + 2\gamma_0 \gamma_1 \sum_{k=1}^{K} x_k \pi_1 \right\}.$$  

(25)

The update for $\sigma^2_w$ may be obtained likewise.
1.2.5 M-step update for $\sigma^2_z$

Let $Q_4$ denote the term in $Q$ containing $\sigma^2_z$.

$$Q_4 = -\frac{K}{2} \log (2 \pi \sigma^2_z) - \sum_{k=1}^{K} \mathbb{E} \left[ \left( x_k - \rho x_{k-1} - \alpha I_k \right)^2 \right] \frac{1}{2 \sigma^2_z}$$

$$= -\frac{K}{2} \log (2 \pi \sigma^2_z) - \sum_{k=2}^{K} \mathbb{E} \left[ \left( x_k - \rho x_{k-1} - \alpha I_k \right)^2 \right] \frac{1}{2 \sigma^2_z} - \mathbb{E} \left[ (\alpha I_k)^2 \right] \frac{1}{2 \sigma^2_z} \quad (26)$$

We take the partial derivative with respect to $\sigma^2_z$ and set it to 0.

$$\frac{\partial Q_4}{\partial \sigma^2_z} = -\frac{K}{2 \sigma^2_z} + \frac{1}{2 \sigma^2_z} \sum_{k=2}^{K} \mathbb{E} \left[ (x_k - \rho x_{k-1} - \alpha I_k)^2 \right] + \frac{(\alpha I_k)^2}{2 \sigma^2_z} = 0 \quad (27)$$

$$\sigma^2_z = \frac{1}{K} \sum_{k=2}^{K} \mathbb{E} \left[ x_k^2 \right] - 2 \rho \mathbb{E} \left[ x_k x_{k-1} \right] + \rho^2 \mathbb{E} \left[ x_{k-1}^2 \right] - 2 \alpha I_k \mathbb{E} \left[ x_k \right] + 2 \alpha I_k \mathbb{E} \left[ x_{k-1} \right] + \frac{\alpha^2}{K} \sum_{k=1}^{K} I_k^2 \quad (28)$$

1.2.6 M-step updates for $\beta_0$ and $\beta_1$

Let $Q_5$ denote the expectation term containing $\beta_0$ and $\beta_1$.

$$Q_5 = \sum_{k=1}^{K} \mathbb{E} \left[ m_k (\beta_0 + \beta_1 x_k) - \log \left( 1 + e^{\beta_0 + \beta_1 x_k} \right) \right] \quad (29)$$

We perform a Taylor expansion of the logarithm term around $x_{k|K}[1]$.

$$\log \left( 1 + e^{\beta_0 + \beta_1 x_k} \right) \approx \log \left( 1 + e^{\beta_0 + \beta_1 x_{k|K}} \right) + \beta_1 p_{k|K} (x_k - x_{k|K}) + \frac{\beta_1^2}{2} p_{k|K} (1 - p_{k|K}) (x_k - x_{k|K})^2 \quad (30)$$

Taking the expected value on both sides,

$$\mathbb{E} \left[ \log \left( 1 + e^{\beta_0 + \beta_1 x_k} \right) \right] \approx \log \left( 1 + e^{\beta_0 + \beta_1 x_{k|K}} \right) + \beta_1 p_{k|K} \mathbb{E} \left[ x_k - x_{k|K} \right] + \frac{\beta_1^2}{2} p_{k|K} (1 - p_{k|K}) \mathbb{E} \left[ (x_k - x_{k|K})^2 \right]$$

$$= \log \left( 1 + e^{\beta_0 + \beta_1 x_{k|K}} \right) + 0 + \frac{\beta_1^2}{2} p_{k|K} (1 - p_{k|K}) \sigma^2_{k|K} \quad (31)$$

Therefore,

$$Q_5 \approx \sum_{k=1}^{K} \left[ m_k (\beta_0 + \beta_1 x_{k|K}) - \log \left( 1 + e^{\beta_0 + \beta_1 x_{k|K}} \right) - \frac{\beta_1^2}{2} p_{k|K} (1 - p_{k|K}) \sigma^2_{k|K} \right] \quad (32)$$

Now,

$$\frac{\partial p_{k|K}}{\partial \beta_0} = \frac{\partial}{\partial \beta_0} \left[ \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_{k|K})}} \right]$$
\[
\begin{align*}
\frac{\partial Q}{\partial \beta} &= \left[1 + e^{-(\beta_0 + \beta_1 x)}\right]^2 \times \left[ -e^{-(\beta_0 + \beta_1 x)} \right] \\
&= p_k | K (1 - p_k | K).
\end{align*}
\]

And similarly for \( \beta_1 \),

\[
\frac{\partial p_k | K}{\partial \beta_1} = p_k | K (1 - p_k | K) x_k | K.
\]

Taking the partial derivative w.r.t. \( \beta_0 \),

\[
\frac{\partial Q_5}{\partial \beta_0} = \sum_{k=1}^{K} \left\{ m_k - p_k | K - \frac{\beta_1^2 \sigma^2}{2} \frac{\partial}{\partial \beta_0} \left[ p_k | K (1 - p_k | K) \right] \right\}
\]

\[
0 = \sum_{k=1}^{K} \left[ m_k - p_k | K - \frac{\beta_1^2 \sigma^2}{2} (1 - p_k | K) (1 - 2 p_k | K) p_k | K \right].
\]

And similarly for \( \beta_1 \), we arrive at,

\[
\frac{\partial Q_5}{\partial \beta_1} = \sum_{k=1}^{K} \left[ m_k x_k | K - x_k | K p_k | K \right.
\]

\[
- \frac{\beta_1^2 \sigma^2}{2} p_k | K (1 - p_k | K) \left[ 2 + \beta_1 x_k | K (1 - 2 p_k | K) \right]
\]

\[
= 0
\]

\subsection{1.2.7 Approximation for the expectation term containing \( \lambda_{k,j} \)}

Let \( Q_6 \) denote the expectation term containing \( \lambda_{k,j} \).

\[
Q_6 = \sum_{k=1}^{K} \sum_{j=1}^{J} \mathbb{E} \left[ \log(\lambda_{k,j} \Delta) n_{k,j} - \lambda_{k,j} \Delta \right]
\]

We perform a Taylor expansion of the summed term around \( x_k | K \) [1].

\[
\log(\lambda_{k,j} \Delta) n_{k,j} - \lambda_{k,j} \Delta \approx \log(\lambda_{k,j} | K \Delta) n_{k,j} - \lambda_{k,j} | K \Delta + \frac{1}{\lambda_{k,j} | K} \frac{\partial \lambda_{k,j} | K}{\partial x_k} (n_{k,j} - \lambda_{k,j} | K \Delta) (x_k - x_k | K)
\]

\[
+ \frac{1}{2} \left[ \frac{1}{\lambda_{k,j} | K} \frac{\partial^2 \lambda_{k,j} | K}{\partial x_k^2} (n_{k,j} - \lambda_{k,j} | K \Delta) - \frac{n_{k,j}}{\lambda^2_{k,j} | K} \left( \frac{\partial \lambda_{k,j} | K}{\partial x_k} \right)^2 \right] (x_k - x_k | K)^2
\]

Taking the expected value on both sides,

\[
\mathbb{E} \left[ \log(\lambda_{k,j} \Delta) n_{k,j} - \lambda_{k,j} | K \Delta \right] \approx \log(\lambda_{k,j} | K \Delta) n_{k,j} - \lambda_{k,j} | K \Delta + \frac{1}{\lambda_{k,j} | K} \frac{\partial \lambda_{k,j} | K}{\partial x_k} (n_{k,j} - \lambda_{k,j} | K \Delta) \mathbb{E} [x_k - x_k | K]
\]

\[
+ \frac{1}{2} \left[ \frac{1}{\lambda_{k,j} | K} \frac{\partial^2 \lambda_{k,j} | K}{\partial x_k^2} (n_{k,j} - \lambda_{k,j} | K \Delta) - \frac{n_{k,j}}{\lambda^2_{k,j} | K} \left( \frac{\partial \lambda_{k,j} | K}{\partial x_k} \right)^2 \right] \mathbb{E} [x_k - x_k | K]^2
\]

\[
\approx \log(\lambda_{k,j} | K \Delta) n_{k,j} - \lambda_{k,j} | K \Delta + 0
\]
\[
+ \frac{1}{2} \left[ \frac{1}{\lambda_{k,j|K}} \frac{\partial^2 \lambda_{k,j|K}}{\partial x_k^2} (n_{k,j} - \lambda_{k,j|K} \Delta) - \frac{n_{k,j}}{\lambda_{k,j|K}^2} \left( \frac{\partial \lambda_{k,j|K}}{\partial x_k} \right)^2 \right] \sigma_{k|K}^2.
\]

Therefore,

\[
Q_6 \approx \sum_{k=1}^{K} \sum_{j=1}^{J} \log(\lambda_{k,j|K} \Delta) n_{k,j} - \lambda_{k,j|K} \Delta + \frac{1}{2} \left[ \frac{1}{\lambda_{k,j|K}} \frac{\partial^2 \lambda_{k,j|K}}{\partial x_k^2} (n_{k,j} - \lambda_{k,j|K} \Delta) - \frac{n_{k,j}}{\lambda_{k,j|K}^2} \left( \frac{\partial \lambda_{k,j|K}}{\partial x_k} \right)^2 \right] \sigma_{k|K}^2.
\]

2 Experimental data – model parameter estimates

The experimental model parameters estimated for each participant are shown in Table 1. Recall that we estimate \( x_k \) at the E-step and calculate the model parameters at the M-step. Recall also that due to computational complexity we split the estimation of the model parameters related to heart rate (i.e., the \( \theta_i \)'s and the \( \eta \) coefficient) into two parts and calculate them separately. We calculate the \( \theta_i \)'s offline based on maximum likelihood estimation (MLE) and select \( \eta \) based on which value maximized a log-likelihood term. As pointed out in the “Discussion” section of the main text, this separated-out calculation is a limitation of our model (e.g. it can result in numerical issues). The separate estimation of the heart rate parameters may be the reason why \( \eta \) values are small in the final estimates and why larger \( \eta \) values cause convergence issues in the Newton-Raphson solution for the state update \( x_{k|k} \) since the MLE estimation of the \( \theta_i \)'s may account for much of the heart rate variability. The value of \( \beta_1 \) also turned out to be negative for two participants (the M-step updates turned out to be negative even after the first iteration). Our algorithm provides two options for calculating \( \beta_0 \) and \( \beta_1 \) and it is possible to select the alternate option which sets \( \beta_0 = 1 \) and calculates \( \beta_1 \) empirically as well if this is to be avoided beforehand. As noted in the main text, a model with a less complex form of the conditional intensity function may enable all the parameters to be recovered at once at the M-step. Lower computational load is also likely to have the benefit of easier deployment onto a wearable platform.
Table 1: Estimated model parameters on experimental data

| Participant | α    | ρ    | δ_0 | δ_1 | σ_w^2 | γ_0 | γ_1 | σ_v^2 | β_0 | β_1 | σ_ε^2 | η  |
|-------------|------|------|-----|-----|--------|-----|-----|--------|-----|-----|--------|----|
| 1           | 0.1339 | 0.9964 | 6.0453 | 0.8501 | 0.4475 | -2.1516 | 0.8182 | 0.4881 | -4.5533 | 0.1642 | 0.0147 | -10^{-4} |
| 2           | 0.1752 | 0.9984 | 10.4601 | 0.8507 | 0.4731 | -6.3457 | 0.8019 | 0.5319 | -4.9209 | 0.1145 | 0.0149 | -10^{-6} |
| 3           | 0.2050 | 0.9944 | 15.4847 | 0.9157 | 0.5967 | -2.9270 | 0.8805 | 0.6270 | -3.9055 | 0.2620 | 0.0196 | -10^{-6} |
| 4           | 0.2122 | 0.9889 | 14.8302 | 0.9337 | 0.5191 | -3.3621 | 0.9018 | 0.5515 | -3.1620 | 0.0282 | 0.0195 | -10^{-6} |
| 5           | 0.1779 | 0.9963 | 13.4671 | 0.8726 | 0.4388 | -3.6769 | 0.7916 | 0.5381 | -3.7924 | -0.1577 | 0.0150 | -10^{-6} |
| 6           | 0.1532 | 0.9888 | 14.6993 | 0.8869 | 0.4876 | -0.5512 | 0.8542 | 0.5247 | -3.8582 | 0.1047 | 0.0170 | -10^{-5} |
| 7           | 0.0934 | 0.9986 | 4.0806  | 0.8944 | 0.2311 | -2.6035 | 0.8653 | 0.2804 | -4.1830 | 0.5256 | 0.0109 | -10^{-3} |
| 8           | 0.1231 | 0.9933 | 13.5384 | 0.8755 | 0.4095 | -0.9875 | 0.8193 | 0.4829 | -3.6242 | 0.3115 | 0.0148 | -10^{-4} |
| 9           | 0.1587 | 0.9966 | 6.0052  | 0.9050 | 0.3256 | -4.0313 | 0.8408 | 0.4178 | -3.6640 | -0.1002 | 0.0141 | -10^{-6} |
| 10          | 0.1365 | 0.9911 | 5.6286  | 0.8804 | 0.3722 | -2.1238 | 0.8515 | 0.4127 | -4.3647 | 0.1101 | 0.0142 | -10^{-4} |
| 11          | 0.1533 | 0.9947 | 17.6648 | 0.9119 | 0.4269 | -0.7794 | 0.8601 | 0.4901 | -3.4974 | 0.1716 | 0.0171 | -10^{-3} |
| 12          | 0.1536 | 0.9931 | 11.5030 | 0.9069 | 0.3077 | -1.6278 | 0.8558 | 0.3835 | -4.0583 | 0.2806 | 0.0140 | -10^{-6} |

| Participant | θ_i’s             |
|-------------|-------------------|
| 1           | 0.0744, 1.1497, -0.1054, -0.1686, 1289.5 |
| 2           | 0.0733, 1.0793, -0.2674, -0.0028, 0.4085, -0.4041, -0.0461, 0.1289, 0.0149, 1404.5 |
| 3           | 0.2484, 0.7572, 231.5820 |
| 4           | 0.2079, 1.3273, -0.4968, -0.0615, -0.0765, -0.0068, 0.0194, 341.5199 |
| 5           | 0.1319, 1.3161, -0.4226, -0.1202, 0.0481, -0.0603, 0.0376, 561.4294 |
| 6           | 0.0677, 0.9083, 625.1079 |
| 7           | 0.1478, 1.0039, -0.2359, 0.0760, 585.6449 |
| 8           | 0.1289, 1.1844, -0.3928, 699.4618 |
| 9           | 0.0229, 0.7706, -0.0228, -0.1910, 0.2809, 0.1220, 1253.4 |
| 10          | 0.1575, 1.1033, -0.3104, 714.4876 |
| 11          | 0.0990, 1.0544, -0.4583, 0.0098, 0.1616, 0.0254, -0.0710, 0.0540, 0.1015, 107.0561 |
| 12          | 0.0818, 0.7980, -0.2500, 0.0853, 0.2650, 693.1574 |
References

[1] Coleman TP, Yanike M, Suzuki WA, Brown EN. A mixed-filter algorithm for dynamically tracking learning from multiple behavioral and neurophysiological measures. In: The dynamic brain: An exploration of neuronal variability and its functional significance. Oxford Univ. Press; 2011. p. 1–16.

[2] Smith AC, Frank LM, Wirth S, Yanike M, Hu D, Kubota Y, et al. Dynamic analysis of learning in behavioral experiments. Journal of Neuroscience. 2004;24(2):447–461.

[3] Prerau MJ, Smith AC, Eden UT, Kubota Y, Yanike M, Suzuki W, et al. Characterizing learning by simultaneous analysis of continuous and binary measures of performance. Journal of Neurophysiology. 2009;102(5):3060–3072.