Quantifying the uncertainty of activation periods in fMRI data via changepoint analysis

Christopher F. H. Nam
Department of Statistics, University of Warwick
Joint work with John A. D. Aston and Adam M. Johansen

CRiSM Changepoint Workshop,
Tuesday 27th March 2012

fMRI data from an Anxiety Induced Experiment

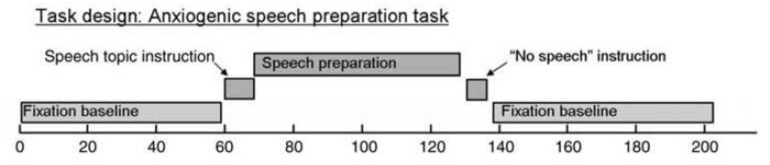


Figure: Design of Anxiety Induced Experiment

- 215 images/time points
- 24 subjects

Image from Lindquist et al. [2007]

fMRI Data

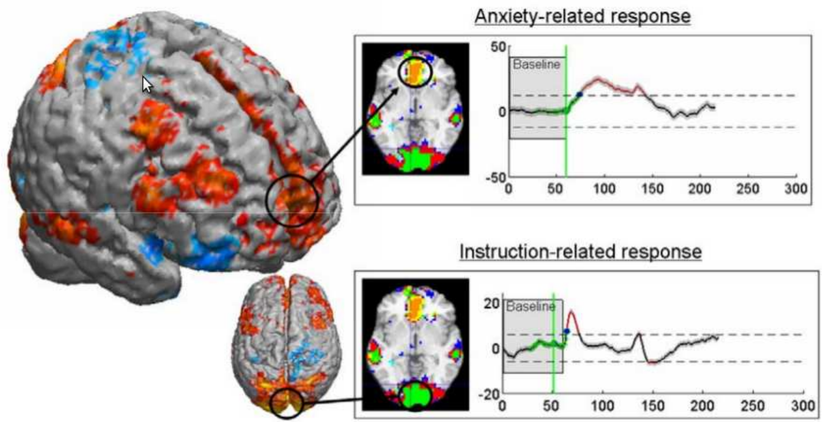
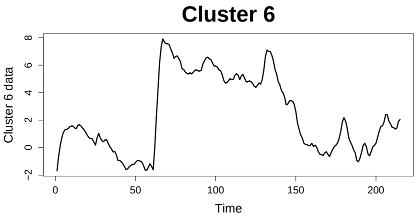
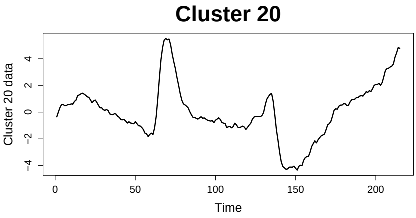


Image from Lindquist et al. [2007]

Brain Regions of interest



(a) RMPFC time series



(b) VC time series

Questions of interest

- The exact design of the experiment.
- When might activations have occurred?
- How many activations may have occurred?
- How long are these activation periods?

Aim

Propose a methodology to fully capture the uncertainty of changepoint characteristics given a sequence of data $y = y_{1:n} = (y_1, \ldots, y_n)$.

- Changepoint Probability (CPP) to a regime occurring at time t.
- Probability of m changepoints occurring in the data.

Focus our attention to these quantities although others are easily accessible.

Hidden Markov Models

- Assume y can be generated by a finite number of states.
- **Hidden Markov Models (HMMs)** to model time series y with $\{X_t\}$ as our underlying Markov Chain, with finite state space Ω_X.

General Finite State Hidden Markov Model

$$\text{Emission:} \quad y_t | y_{1:t-1}, x_{1:t}, \theta \sim f(y_t | x_{t-r:t}, y_{1:t-1}, \theta)$$
$$\text{Transition:} \quad p(x_t | y_{1:t-1}, x_{-r+1:t-1}, \theta) = p(x_t | x_{t-1}, \theta)$$

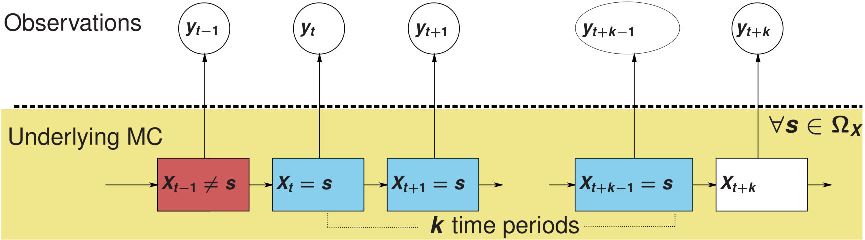
- Changepoints in this HMM framework.

Changepoint Definition

Changepoint Definition

Changepoint to a regime is said to have occurred at time t when a state persists for at least k time periods in $\{X_t\}$.

$$x_{t-1} \neq x_t = x_{t+1} = \ldots = x_{t+j} \quad \text{where } j \geq k - 1$$



Define $\tau^{(k)}$ and $M^{(k)}$ to be the time and number of changepoints.

Model Parameters

- θ indicates our model parameter vector which needs to be estimated.
- Components will vary; dependent on the particular type of HMM used.
  - **P**, the $|\Omega_X| \times |\Omega_X|$ probability transition matrix for $\{X_t\}$
  - Parameters for the emission distribution which can be dependent on the underlying state
    Means μ_{X_t}, variances $\sigma^2_{X_t}$, Poisson intensity rates λ_{X_t}, ...

Suppose we are interested in the changepoint probability $P(\tau^{(k)} = t|y)$.

$$P(\tau^{(k)} = t|y) = \int_{\theta} P(\tau^{(k)} = t, \theta|y)d\theta = \int_{\theta} P(\tau^{(k)} = t|\theta, y)p(\theta|y)d\theta$$

Suppose we are interested in the changepoint probability $P(\tau^{(k)} = t|y)$.

$$P(\tau^{(k)} = t|y) = \int_\theta P(\tau^{(k)} = t, \theta|y)d\theta = \int_\theta P(\tau^{(k)} = t|\theta, y)p(\theta|y)d\theta$$

$$\approx \widehat{P^N}(\tau^{(k)} = t|y) = \sum_{i=1}^{N} W^i P(\tau^{(k)} = t|\theta^i, y)$$

by standard Monte Carlo results.

- $\{W^i, \theta^i\}_{i=1}^N$ approximates $p(\theta|y)$
- $P(\tau^{(k)} = t|\theta^i, y)$

Proposed Methodology

1. Approximate $p(\theta|y)$ via SMC samplers to obtain a normalised weighted cloud of particles $\{W^i, \theta^i\}_{i=1}^N$.

Proposed Methodology

1. Approximate $p(\theta|y)$ via SMC samplers to obtain a normalised weighted cloud of particles $\{W^i, \theta^i\}_{i=1}^N$.
2. For $i = 1, \ldots, N$, compute the associated exact changepoint distribution conditioned on θ^i, $P(\tau^{(k)} = t|\theta^i, y)$, via FMCI in a HMM context.

Proposed Methodology

1. Approximate $p(\theta|y)$ via SMC samplers to obtain a normalised weighted cloud of particles $\{W^i, \theta^i\}_{i=1}^N$.
2. For $i = 1, \ldots, N$, compute the associated exact changepoint distribution conditioned on θ^i, $P(\tau^{(k)} = t|\theta^i, y)$, via FMCI in a HMM context.
3. **Weighted average** of these exact changepoint distributions to obtain the changepoint distributions in light of parameter uncertainty, $P(\tau^{(k)} = t|y)$.

Proposed Methodology

1. Approximate $p(\theta|y)$ via SMC samplers to obtain a normalised weighted cloud of particles $\{W^i, \theta^i\}_{i=1}^N$.
2. For $i = 1, \ldots, N$, compute the associated exact changepoint distribution conditioned on θ^i, $P(\tau^{(k)} = t|\theta^i, y)$, via FMCI in a HMM context.
3. **Weighted average** of these exact changepoint distributions to obtain the changepoint distributions in light of parameter uncertainty, $P(\tau^{(k)} = t|y)$.

Proposed Methodology

1. Approximate $p(\theta|y)$ via SMC samplers to obtain a normalised weighted cloud of particles $\{W^i, \theta^i\}_{i=1}^N$.
2. For $i = 1, \dots, N$, compute the associated exact changepoint distribution conditioned on θ^i, $P(\tau^{(k)} = t|\theta^i, y)$, via FMCI in a HMM context.
3. **Weighted average** of these exact changepoint distributions to obtain the changepoint distributions in light of parameter uncertainty, $P(\tau^{(k)} = t|y)$.

Combines recent work of Del Moral et al. [2006] and Aston et al. [2009].
Specific details can be found in Nam et al. [in press].

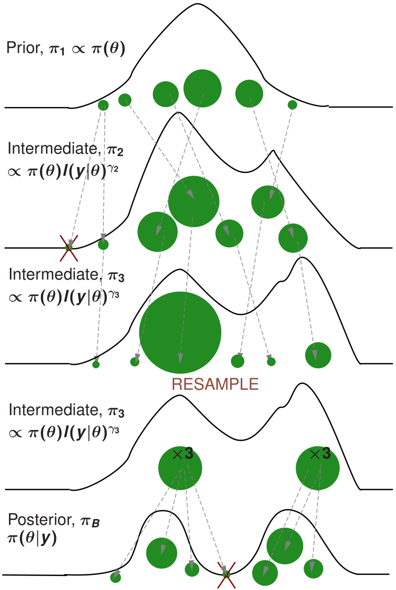
Sequential Monte Carlo (SMC) Samplers

SMC Samplers

$$\pi_b \propto p(\theta) l(y|\theta)^{\gamma_b}$$

where

- $p(\theta)$ = prior of the model parameters
- $l(y|\theta)$ = likelihood
- $0 = \gamma_1 \leq \gamma_2 \leq \ldots \leq \gamma_B = 1$, a tempering schedule



Exact distributions

- Exact computation of $P(\tau^{(k)} = t|\theta^i, y)$.

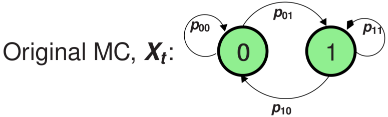
Exact distributions

- Exact computation of $P(\tau^{(k)} = t|\theta^i, y)$.
- $\tau_u^{(k)}$ denote the uth changepoint under our definition.
- $P(\tau^{(k)} = t|\theta^i, y) = \sum_{u=1,2,\dots} P(\tau_u^{(k)} = t|\theta^i, y)$.
- $P(\tau_u^{(k)} = t|\theta^i, y) \equiv P(W(k, u) = t + k - 1|\theta^i, y)$ re-express as a waiting time for runs.

Exact distributions

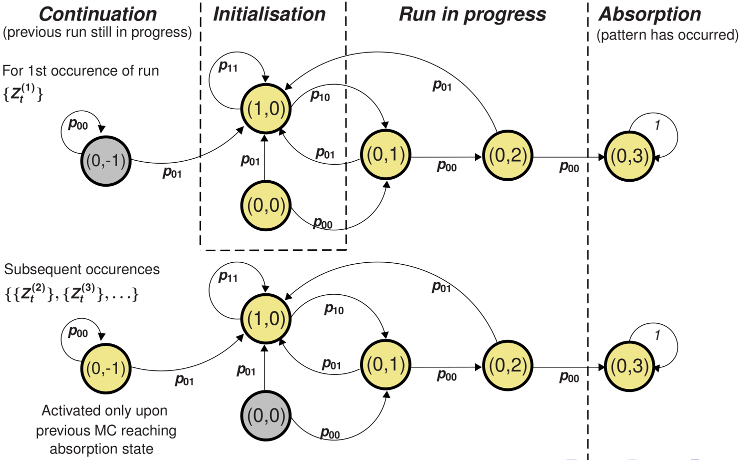
- Exact computation of $P(\tau^{(k)} = t|\theta^i, y)$.
- $\tau_u^{(k)}$ denote the uth changepoint under our definition.
- $P(\tau^{(k)} = t|\theta^i, y) = \sum_{u=1,2,\ldots} P(\tau_u^{(k)} = t|\theta^i, y)$.
- $P(\tau_u^{(k)} = t|\theta^i, y) \equiv P(W(k,u) = t + k - 1|\theta^i, y)$ re-express as a waiting time for runs.
- Waiting time distribution for runs can be computed exactly via Finite Markov Chain Imbedding (FMCI).

Finite Markov Chain Imbedding (FMCI)



Run of interest: **000** $(s = 0, k = 3)$

Via FMCI: Auxiliary MCs $\{\{Z_t^{(1)}\}, \{Z_t^{(2)}\}, \ldots\}$



Waiting Time Distributions via FMCI

- $P(W(k,u) \leq t+k-1|\theta^i) = P(Z^{(u)}_{t+k-1} \in A|\theta^i)$ where A denotes the set of absorption states in Ω_Z.
- Computed by standard Markov Chain results.

FMCI in a HMM context

- Inference on the underlying state sequence.

FMCI in a HMM context

- Inference on the underlying state sequence.
- Compute posterior transition probabilities $P(X_t|X_{t-1}, y)$ to consider all possible state sequences.
- Sequence of time dependent posterior transition probability matrices $\{\tilde{\mathbf{P}}_1, \tilde{\mathbf{P}}_2, \ldots, \tilde{\mathbf{P}}_n\}$.

FMCI in a HMM context

- Inference on the underlying state sequence.
- Compute posterior transition probabilities $P(X_t|X_{t-1}, y)$ to consider all possible state sequences.
- Sequence of time dependent posterior transition probability matrices $\{\tilde{\mathbf{P}}_1, \tilde{\mathbf{P}}_2, \ldots, \tilde{\mathbf{P}}_n\}$.
- Results in a sequence of time dependent posterior transition probability matrices $\{\tilde{\mathbf{M}}_1, \tilde{\mathbf{M}}_2, \ldots, \tilde{\mathbf{M}}_n\}$ for the auxiliary MCs $\{Z_t\}$.

Exact Distributions for Changepoint characteristics

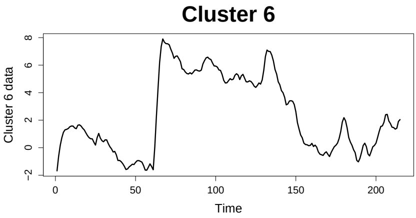
Probability of uth changepoint at specific time

$$P(\tau_u^{(k)} = t|\theta^i, y) = P(W(k,u) = t+k-1|\theta^i, y)$$
$$= P(W(k,u) \leq t+k-1|\theta^i, y) - P(W(k,u) \leq t+k-2|\theta^i, y)$$

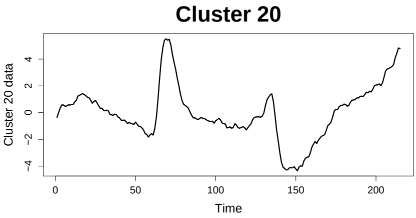
Distribution of the number of changepoints,

$$P(M^{(k)} = u|\theta^i) = P(W(k,u) \leq n|\theta^i, y) - P(W(k,u+1) \leq n|\theta^i, y)$$

fMRI data



(c) RMPFC time series



(d) VC time series

Preliminary I

- AR error process leads to a HMS-AR model.
- Detrending within model.

HMS-AR(r) with detrending

$$\begin{aligned} y_t - \mu_{x_t} - \mathbf{m}_t^{'}\beta &= a_t \\ a_t &= \phi_1 a_{t-1} + \ldots + \phi_r a_{t-r} + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma^2) \end{aligned}$$

- Detrending parameters within model $\rightarrow$ Estimated within SMC samplers.

Preliminary II

- Detrending types:
  No detrending, Polynomial of order 3, Discrete Cosine Basis

Preliminary II

- Detrending types:
  No detrending, Polynomial of order 3, Discrete Cosine Basis
- HMS-AR of order 0 and 1

Preliminary II

- Detrending types:
  No detrending, Polynomial of order 3, Discrete Cosine Basis
- HMS-AR of order 0 and 1
- $\Omega_X = \{\text{"resting"}, \text{"active"}\}$

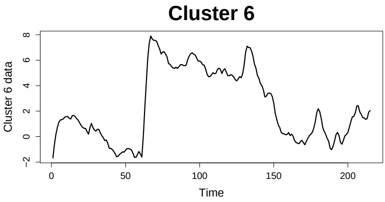
Preliminary II

- Detrending types:
  No detrending, Polynomial of order 3, Discrete Cosine Basis
- HMS-AR of order 0 and 1
- $\Omega_X = \{\text{``resting''}, \text{``active''}\}$
- 5 consecutive active states for activated region $(k = 5)$

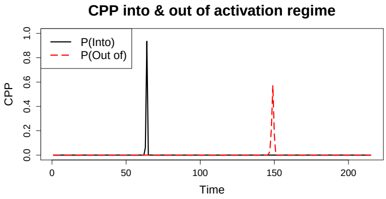
Preliminary II

- Detrending types:
  No detrending, Polynomial of order 3, Discrete Cosine Basis
- HMS-AR of order 0 and 1
- $\Omega_X = \{\text{"resting"}, \text{"active"}\}$
- 5 consecutive active states for activated region $(k = 5)$
- SMC samplers: $500 = N$ particles, $100 = B$ iterations

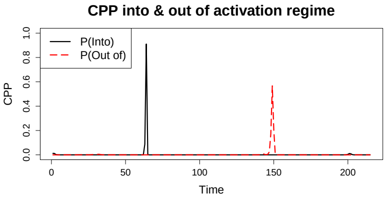
Cluster 6, CPP HMS-AR(0)



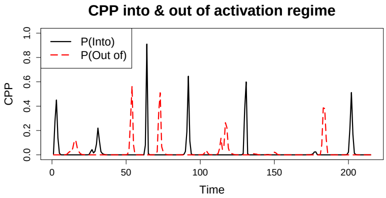
(e) RMPFC time series



(f) AR(0), No detrending



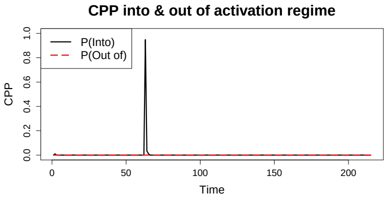
(g) AR(0), Poly detrending



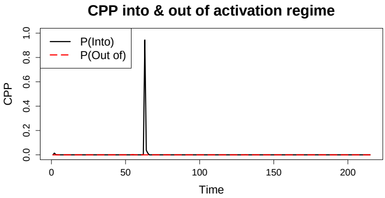
(h) AR(0), DCB detrending

Cluster 6, CPP, HMS-AR(1)

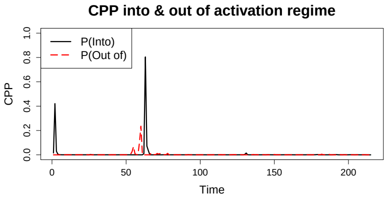
(i) RMPFC time series



(j) AR(1), No detrending

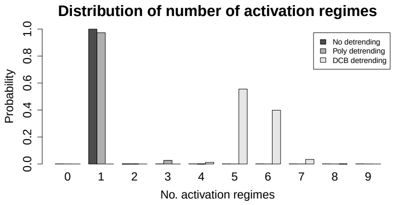


(k) AR(1), Poly detrending

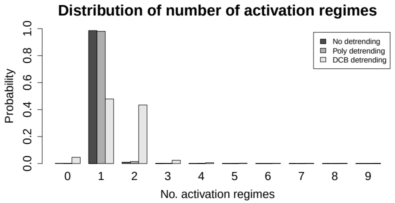


(l) AR(1), DCB detrending

Cluster 6, Distribution of Number Activation Regimes

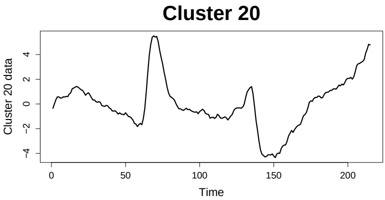


(m) AR(0)

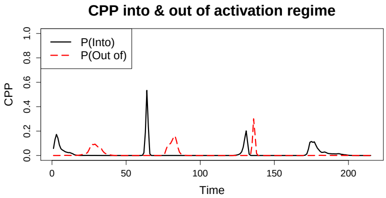


(n) AR(1)

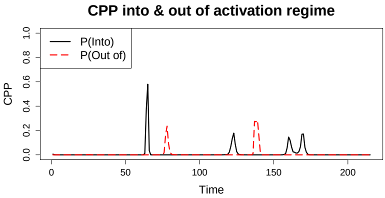
Cluster 20, CPP HMS-AR(0)



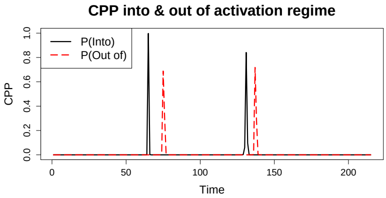
(o) VC time series



(p) AR(0), No detrending

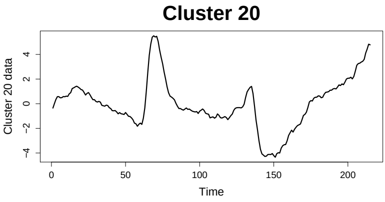


(q) AR(0), Poly detrending

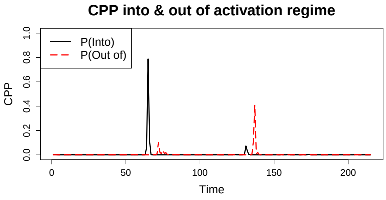


(r) AR(0), DCB detrending

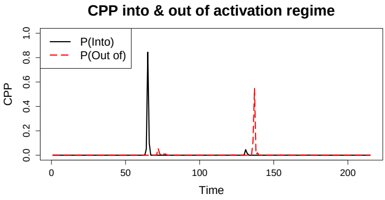
Cluster 20, CPP, HMS-AR(1)



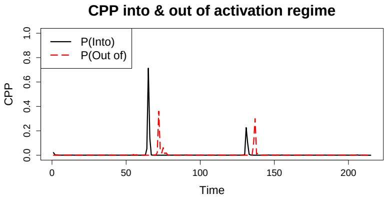
(s) VC time series



(t) AR(1), No detrending

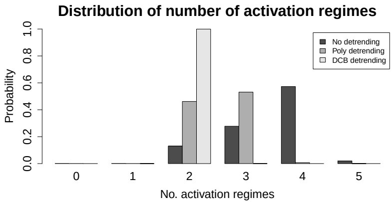


(u) AR(1), Poly detrending

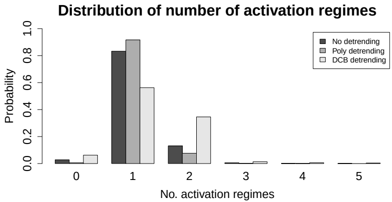


(v) AR(1), DCB detrending

Cluster 20, Distribution of Number Activation Regimes



(w) AR(0)



(x) AR(1)

Conclusions

- Proposed methodology quantifying the uncertainty for changepoint characteristics in light of model parameter uncertainty.

Conclusions

- Proposed methodology quantifying the uncertainty for changepoint characteristics in light of model parameter uncertainty.
- Combines recent work of exact changepoint distributions via FMCI in HMM.

Conclusions

- Proposed methodology quantifying the uncertainty for changepoint characteristics in light of model parameter uncertainty.
- Combines recent work of exact changepoint distributions via FMCI in HMM.
- Accounts for parameter uncertainty via SMC samplers.

Conclusions

- Proposed methodology quantifying the uncertainty for changepoint characteristics in light of model parameter uncertainty.
- Combines recent work of exact changepoint distributions via FMCI in HMM.
- Accounts for parameter uncertainty via SMC samplers.
- Application to fMRI data.

Conclusions

- Proposed methodology quantifying the uncertainty for changepoint characteristics in light of model parameter uncertainty.
- Combines recent work of exact changepoint distributions via FMCI in HMM.
- Accounts for parameter uncertainty via SMC samplers.
- Application to fMRI data.
- Effects of error process assumptions and types of detrending.

References

John A. D. Aston, J. Y. Peng, and Donald E. K. Martin. Implied distributions in multiple change point problems. CRiSM Research Report 08-26, University of Warwick, 2009.

Pierre Del Moral, Arnaud Doucet, and Ajay Jasra. Sequential Monte Carlo samplers. *Journal of the Royal Statistical Society Series B*, 68 (3):411–436, 2006.

Martin A. Lindquist, Christian Waugh, and Tor D. Wager. Modeling state-related fMRI activity using change-point theory. *NeuroImage*, 35 (3):1125–1141, 2007. doi: 10.1016/j.neuroimage.2007.01.004.

C. F. H. Nam, J. A. D. Aston, and A. M. Johansen. Quantifying the uncertainty in change points. *Journal of Time Series Analysis*, (in press).

Implementation of SMC samplers I

- Consider the model parameters for a 2-state Hamilton's MS-AR(r) model. $\theta = (\mathbf{P}, \mu_1, \mu_2, \sigma^2, \phi_1, \ldots, \phi_r)$
- Simple linear tempering schedule,
$\gamma_b = \frac{b-1}{B-1}, \qquad b = 1, \ldots, B.$

Implementation of SMC samplers I

- Consider the model parameters for a 2-state Hamilton's MS-AR(r) model. $\theta = (\mathbf{P}, \mu_1, \mu_2, \sigma^2, \phi_1, \ldots, \phi_r)$
- Simple linear tempering schedule, $\gamma_b = \frac{b-1}{B-1}, \qquad b = 1, \ldots, B.$
- Reparameterisation of variances to precisions, $\lambda = 1/\sigma^2$ -AR parameters to Partial Autocorrelation Coefficients (PAC).

Implementation of SMC samplers II

- **Initialisation:** Assume independence between the components of θ, and sample from Bayesian priors.

Approximation of posterior

$$p(\theta|y) \approx \{W_B^{(i)}, \theta_B^{(i)}\}_{i=1}^N = \{W^i, \theta^i\}_{i=1}^N$$

Implementation of SMC samplers II

- **Initialisation:** Assume independence between the components of θ, and sample from Bayesian priors.
- **Mutation:** Random Walk Metropolis Hastings (RWMH).

Approximation of posterior

$$p(\theta|y) \approx \{W_B^{(i)}, \theta_B^{(i)}\}_{i=1}^N = \{W^i, \theta^i\}_{i=1}^N$$

Implementation of SMC samplers II

- **Initialisation:** Assume independence between the components of θ, and sample from Bayesian priors.
- **Mutation:** Random Walk Metropolis Hastings (RWMH).
- **Selection:** Resample if $ESS = \{\sum_{i=1}^{N}(W_b^{(i)})^2\}^{-1} < N/2$. Re-weight resampled particles to $1/N$.

Approximation of posterior

$$p(\theta|y) \approx \{W_B^{(i)}, \theta_B^{(i)}\}_{i=1}^{N} = \{W^i, \theta^i\}_{i=1}^{N}$$

Implementation of SMC samplers II

- **Initialisation:** Assume independence between the components of θ, and sample from Bayesian priors.
- **Mutation:** Random Walk Metropolis Hastings (RWMH).
- **Selection:** Resample if $ESS = \{\sum_{i=1}^{N}(W_b^{(i)})^2\}^{-1} < N/2$. Re-weight resampled particles to $1/N$.
- **Intermediary Output:** Weighted cloud of N particles, $\{W_b^{(i)}, \theta_b^{(i)}\}_{i=1}^{N}$ approximates the distribution π_b.

Approximation of posterior

$$p(\theta|y) \approx \{W_B^{(i)}, \theta_B^{(i)}\}_{i=1}^{N} = \{W^i, \theta^i\}_{i=1}^{N}$$

Waiting Time Distributions for Runs in HMM context

For $t = 1, \ldots, n$

$$\Psi_t = \Psi_{t-1}\tilde{\mathbf{M}}_t$$
$$\psi_t^{(u)} \leftarrow \psi_t^{(u)} + \psi_{t-1}^{(u-1)}(\tilde{\mathbf{M}}_t - \mathbf{I})\Upsilon, \quad u = 2, \ldots, \lfloor n/k \rfloor$$
$$P(W(k, u|\theta^i, y) \leq t) = P(Z_t^{(u)} \in A) = \psi_t^{(u)} U(A)$$

where :$\psi_t^{(u)} = 1 \times |\Omega_Z|$ vector storing probabilities of the uth chain being in the corresponding state.

$\Psi_t = \lfloor n/k \rfloor \times |\Omega_Z|$ with $(\psi_t^{(1)}, \psi_t^{(2)}, \ldots)$ as row vectors.

$\mathbf{I} = |\Omega_Z| \times |\Omega_Z|$ identity matrix

$\Upsilon = |\Omega_Z| \times |\Omega_Z|$ matrix connecting absorption state to the corresponding continuation state

$U(A) = |\Omega_Z| \times 1$ vector with 1s in location of absorption states 0s elsewhere.