

Chapter 3

Structural breaks for models *with path dependence*

Chapter 3

- Path dependence (p. 3)
- Change-point models (p. 16)
- Markov-switching and Change-point models (p. 26)
 - PMCMC algorithm
 - IHMM-GARCH
- References (p. 43)

Path dependence

Chib's specification

Advantages

- Multiple breaks
- Recurrent or no recurrent states (Change-point/Markov-switching)
- MCMC with good mixing properties
- Allow to select an optimal number of regimes
- Forecast of structural breaks

—————→ **State of the art !**

Drawbacks

- Geometric distribution for the regime duration
- Many computation for selecting the number of regimes
- **Not applicable to models with path dependence**

Chib's specification

Why not applicable ?

- Simplification in the Forward-backward algorithm :

$$f(y_t | Y_{1:t-1}, S_{1:t}) = f(y_t | Y_{1:t-1}, s_t)$$

- If assumption does not hold :

$$\begin{aligned} \pi(s_t | Y_{1:T}, S_{t+1:T}) &\propto f(s_t | Y_{1:t}) f(S_{t+1:T}, Y_{t+1:T} | Y_{1:t}, s_t) \\ &\propto f(s_t | Y_{1:t}) f(S_{t+1:T} | Y_{1:t}, s_t) f(Y_{t+1:T} | Y_{1:t}, S_{t:T}) \\ &\neq f(s_t | Y_{1:t}) f(s_{t+1} | s_t) \end{aligned}$$

Chib's algorithm not available for

State-space model with structural breaks in parameters

Example : ARMA, GARCH

Path dependent models

CP- and MS-ARMA models

$$y_t = \mu_{s_t} + \theta_{s_t} y_{t-1} + \phi_{s_t} \epsilon_{t-1} + \epsilon_t$$

$$s_t | s_{t-1}, P \sim p_{s_{t-1}, s_t}$$

CP- and MS-GARCH models

$$y_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = \omega_{s_t} + \alpha_{s_t} y_{t-1}^2 + \beta_{s_t} \sigma_{t-1}^2$$

$$s_t | s_{t-1}, P \sim p_{s_{t-1}, s_t}$$

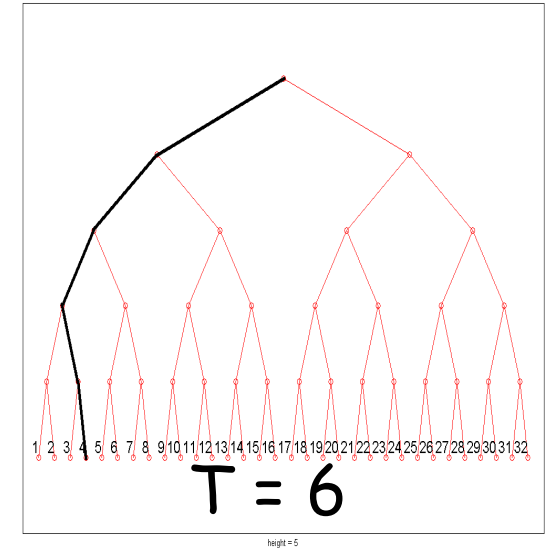
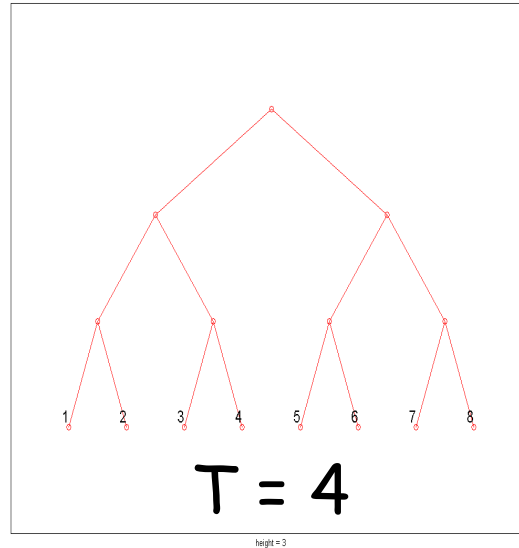
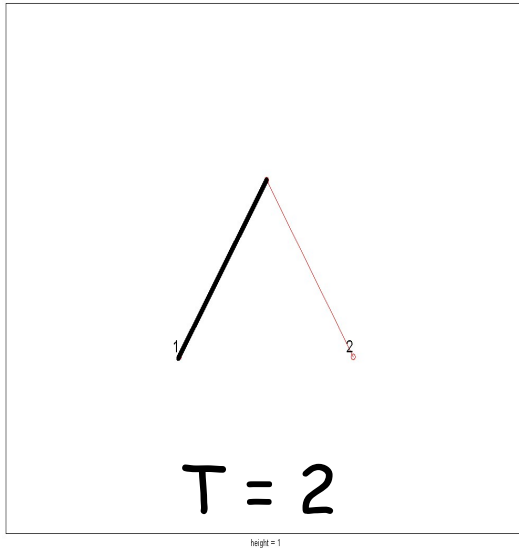
Change-point

$$P = \begin{pmatrix} p_{1,1} & 1 - p_{1,1} & 0 & \dots & 0 \\ 0 & p_{2,2} & 1 - p_{2,2} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Markov-switching

$$P = \begin{pmatrix} p_{1,1} & p_{1,2} & p_{1,3} & \dots & p_{1,K+1} \\ p_{2,1} & p_{2,2} & p_{2,3} & \dots & p_{2,K+1} \\ \dots & \dots & \dots & \dots & \dots \\ p_{K+1,1} & p_{K+1,2} & p_{K+1,3} & \dots & p_{K+1,K+1} \end{pmatrix}$$

Path dependence problem



Function of $S_{1:t-1}$

$$\left\{ \begin{array}{l} \text{ARMA } y_t = \mu_{s_t} + \theta_{s_t} y_{t-1} + \phi_{s_t} \epsilon_{t-1}(S_{1:t-1}) + \epsilon_t \\ \text{GARCH } \sigma_t^2(S_{1:t}) = \omega_{s_t} + \alpha_{s_t} y_{t-1}^2 + \beta_{s_t} \sigma_{t-1}^2(S_{1:t-1}) \end{array} \right.$$

Likelihood at time t depends on the whole path that has been followed so far

T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Poss. paths	1	2	4	8	16	32	64	128	256	512	1024	2048	4096	8192	16384

Path dependence problem

Solutions ?

- 1) Use of approximate models without path dependence
 - Gray (1996), Dueker (1997), Klaassen (2002)
 - Haas, Mittnik, Poella (2004)

$$y_t = \sigma_{t,s_t} \epsilon_t$$

$$\sigma_{t,1}^2 = \omega_1 + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1,1}^2$$

$$\sigma_{t,2}^2 = \omega_2 + \alpha_2 y_{t-1}^2 + \beta_2 \sigma_{t-1,2}^2$$

...

$$\sigma_{t,K+1}^2 = \omega_{K+1} + \alpha_{K+1} y_{t-1}^2 + \beta_{K+1} \sigma_{t-1,K+1}^2$$

Path dependence problem

Solutions ?

2) Stephens (1994) : Inference on multiple breaks

Drawbacks

- Time-consuming if T large
- Many MCMC iterations are required

→ May not converge in a finite amount of time!

3) Bauwens, Preminger, Rombouts (2011) :

- Single-move MCMC

Single-move MCMC

CP- and MS-GARCH models

$$y_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = \omega_{s_t} + \alpha_{s_t} y_{t-1}^2 + \beta_{s_t} \sigma_{t-1}^2$$

$$s_t | s_{t-1}, P \sim p_{s_{t-1}, s_t}$$

Change-point

$$P = \begin{pmatrix} p_{1,1} & 1 - p_{1,1} & 0 & \dots & 0 \\ 0 & p_{2,2} & 1 - p_{2,2} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Markov-switching

$$P = \begin{pmatrix} p_{1,1} & p_{1,2} & p_{1,3} & \dots & p_{1,K+1} \\ p_{2,1} & p_{2,2} & p_{2,3} & \dots & p_{2,K+1} \\ \dots & \dots & \dots & \dots & \dots \\ p_{K+1,1} & p_{K+1,2} & p_{K+1,3} & \dots & p_{K+1,K+1} \end{pmatrix}$$

Single-move MCMC

Metropolis-Hastings sampler : $\Theta = \{\omega_1, \alpha_1, \beta_1, \dots, \omega_{K+1}, \alpha_{K+1}, \beta_{K+1}\}$

$$\theta_i | Y_{1:T}, S_{1:T}, P, \Theta_{-i} \sim \text{Griddy-Gibbs}$$

$$P | Y_{1:T}, S_{1:T}, \Theta \sim \text{Dirichlet}(\eta + n_{i,1:K+1})$$

$$s_t | Y_{1:T}, P, \Theta, \underbrace{S_{1:t-1}, S_{t+1:T}} \sim \text{single-move}$$

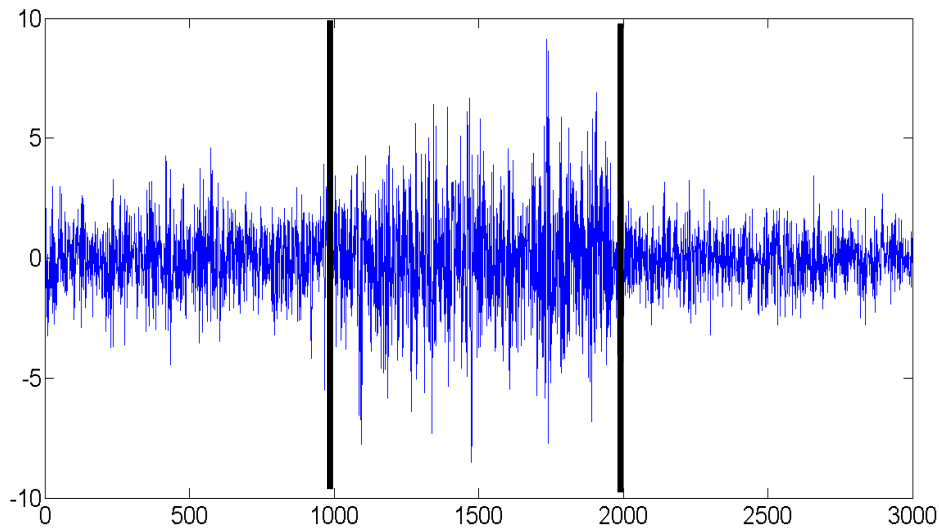
One state updated at a time !

$$\pi(s_t | Y_{1:T}, P, \Theta, S_{-t}) \propto f(Y_{1:T} | \Theta, S_{1:T}) f(s_t | s_{t-1}, P) f(s_{t+1} | s_t, P)$$

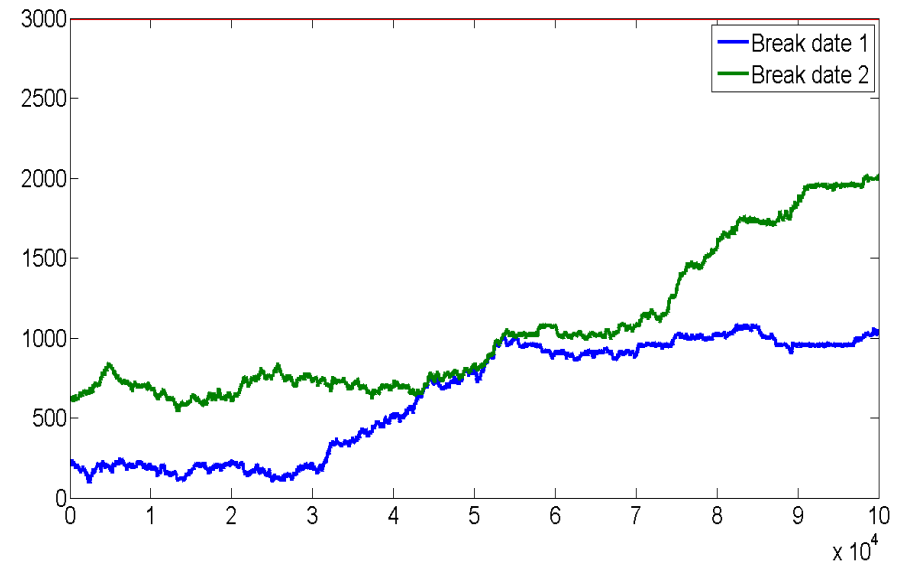
$$\propto \underbrace{f(Y_{1:T} | \Theta, S_{1:T})}_{\text{Likelihood}} \underbrace{p_{s_{t-1}, s_t} p_{s_t, s_{t+1}}}_{\text{Transition matrix}}$$

Likelihood | Transition matrix

Example



Simulated series : $T = 3000$



Initial state : $[200 \ 600]$

$$\text{Corr}(\tau_1^i, \tau_1^{i-200}) = 0.996$$

Convergence after 100.000 MCMC iterations !

Single-move

Advantages

- Generic method :
 - Works for many CP and MS models

Drawbacks

- No criterion for selecting the number of regimes
- Very Time-consuming if T large (especially for MS)
- Many MCMC iterations are required :

Very difficult to assess convergence

May not converge in a finite amount of time !

Questions ?

Change-point models

D-DREAM algorithm

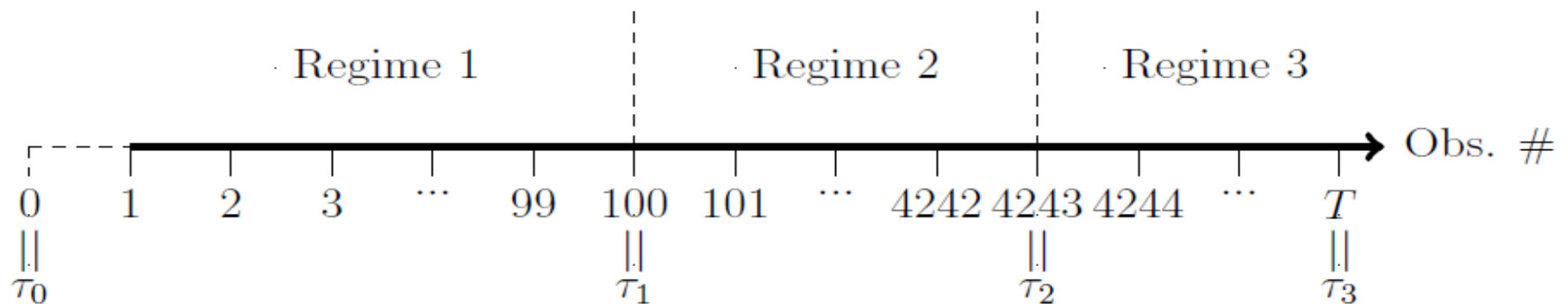
CP-GARCH models :

$$y_t = \mu_k + \epsilon_t,$$

$$\epsilon_t = \sigma_t \eta_t, \text{ with } \eta_t \sim i.i.d. N(0, 1),$$

$$\sigma_t^2 = \omega_k + \alpha_k \epsilon_{t-1}^2 + \beta_k \sigma_{t-1}^2,$$

if $t \in]\tau_{k-1}, \tau_k]$, with $k = \{1, 2, \dots, K + 1\}$ denoting the regime.



Come back to the Stephens' specification !

D-DREAM algorithm

Problem with Stephens' inference :

- Break dates sample one at a time (single-move)
 - MCMC mixing issue
- Very demanding if T is large

Discrete-DREAM MCMC :

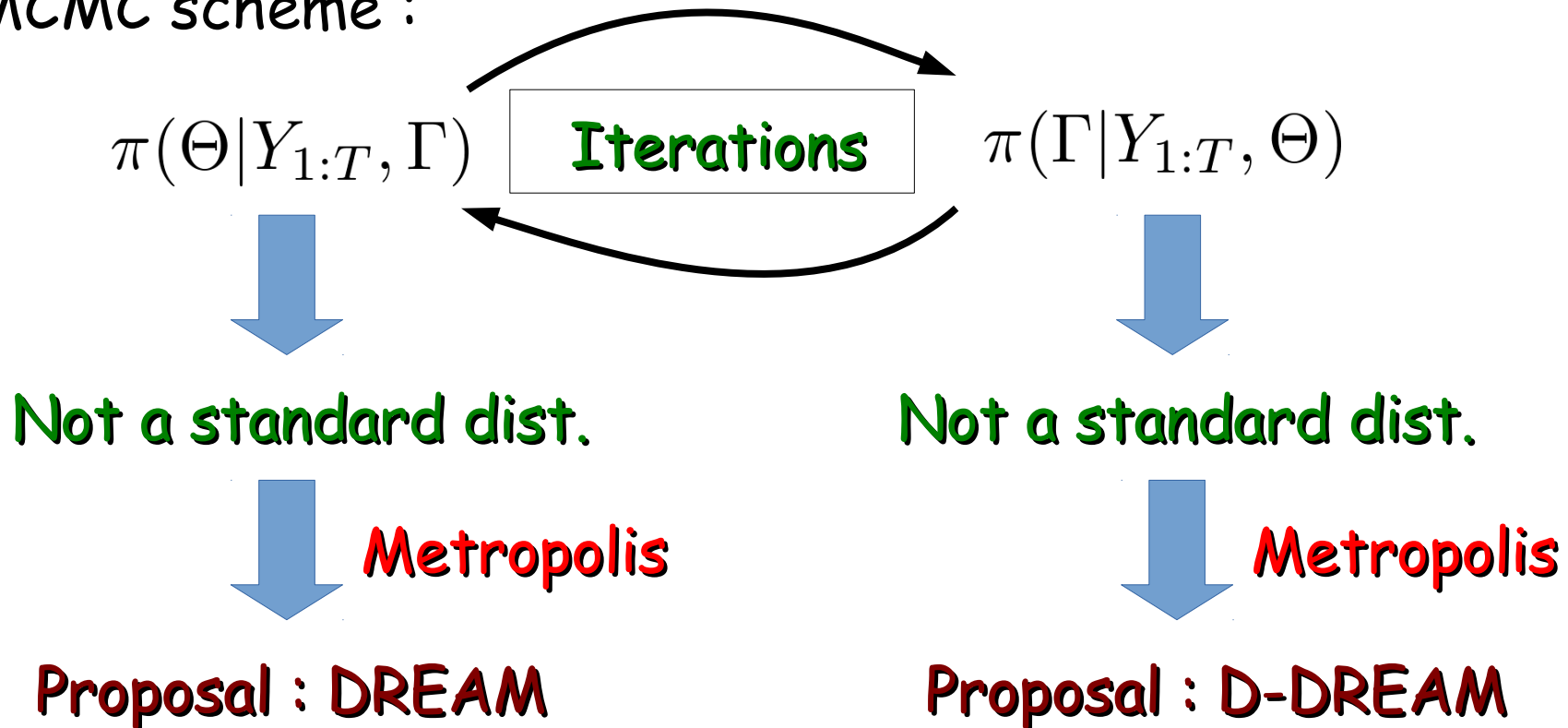
- Metropolis algorithm
 - Jointly sample the break dates
 - Very fast (faster than Forward-Backward)

D-DREAM algorithm

- Two sets of parameters to be estimated :

Continuous $\left\{ \begin{array}{l} \Theta = (\theta_1, \theta_2, \dots, \theta_{K+1})' \\ \theta_k = (\mu_k, \omega_k, \alpha_k, \beta_k) \end{array} \right.$ **Discrete** $\Gamma = (\tau_1, \tau_2, \dots, \tau_K)'$

- MCMC scheme :



D-DREAM algorithm

DiffeRential Adaptative Evolution Metropolis
(Vrugt et al. 2009)

- DREAM automatically determines the **size** of the jump.
- DREAM automatically determines the **direction** of the jump
- DREAM is well suited for **multi-modal** post. dist.
- DREAM is well suited for **high dimensional** sampling
- DREAM is **symmetric** : only a Metropolis ratio

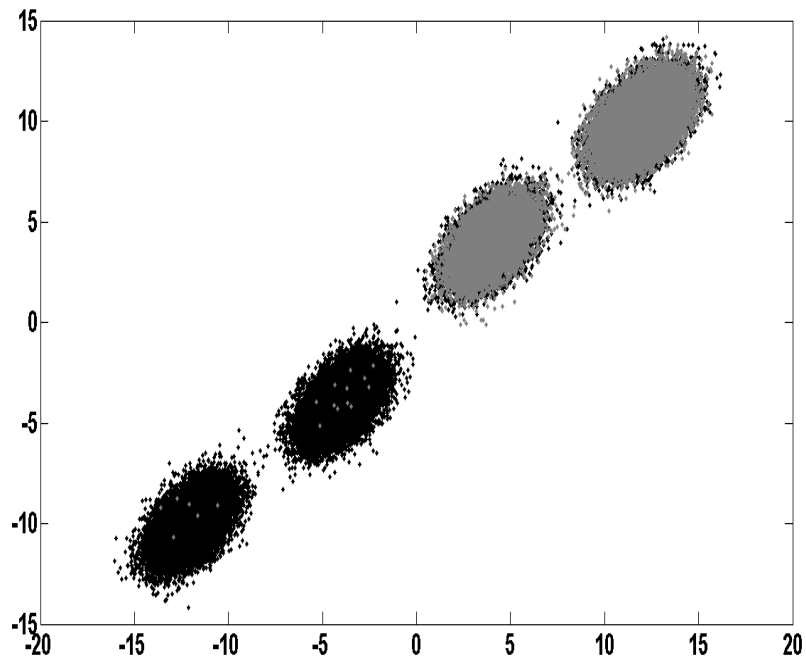
Nevertheless only applicable to continuous parameters



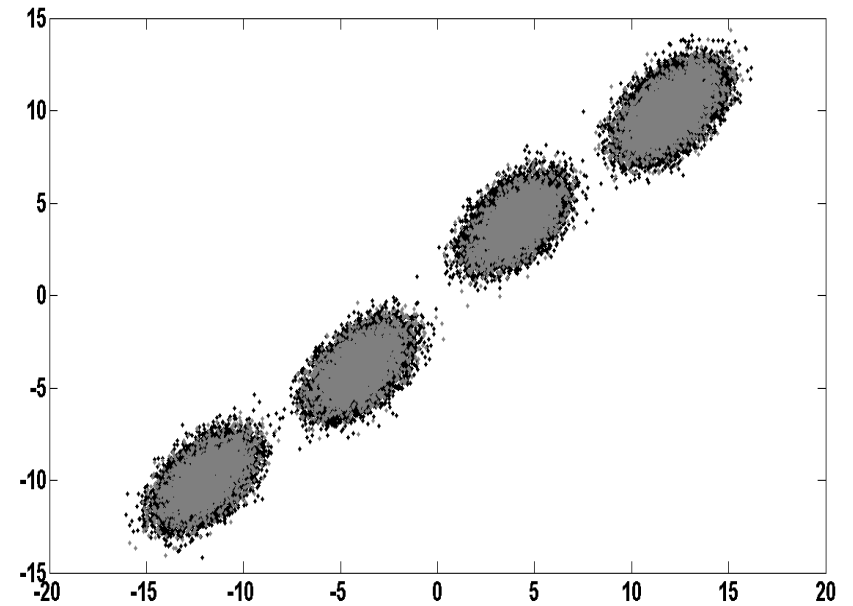
Extension for discrete parameter : Discrete-DREAM

DREAM : Example

Adaptive RW



DREAM



DREAM algorithm

M parallel MCMC chains :

$$\Theta_1^i \quad \Theta_2^i \quad \dots \quad \Theta_M^i$$

Proposal distribution :

$$\delta \in [1, \lfloor M/2 \rfloor]$$

$$\zeta \sim N(0, 10^{-8})$$

$$\tilde{\Theta}_j^i = \Theta_j^i + \gamma(\delta, d) \left(\sum_{g=1}^{\delta} \Theta_{r_1(g)}^i - \sum_{h=1}^{\delta} \Theta_{r_2(h)}^i \right) + \zeta$$

$$\gamma(\delta, d) = 2.38 / \sqrt{2\delta d}$$

$$r_1(g) \neq r_2(h) \neq j$$

Symmetric proposal dist :

- Accept/reject the draw according to the probability

$$\min \left[\frac{f(Y_{1:T} | \tilde{\Theta}_j^i, \Gamma_j^i) f(\tilde{\Theta}_j^i)}{f(Y_{1:T} | \Theta_j^i, \Gamma_j^i) f(\Theta_j^i)}, 1 \right]$$

D-DREAM algorithm

M parallel MCMC chains : $\{\Theta_j^i, \Gamma_j^i\}_{i=1, j=1}^{N, M}$

Discrete

Continuous

$$\Theta = (\theta_1, \theta_2, \dots, \theta_{K+1})'$$

$$\theta_k = (\mu_k, \omega_k, \alpha_k, \beta_k)$$

$$\Gamma = (\tau_1, \tau_2, \dots, \tau_K)'$$

Proposal distribution :

Proposal distribution :

$$\tilde{\Gamma}_j^i = \Gamma_j^i + \text{round}[\gamma(\delta, d) \left(\sum_{g=1}^{\delta} \Gamma_{r_1(g)}^i - \sum_{h=1}^{\delta} \Gamma_{r_2(h)}^i \right) + \zeta]$$

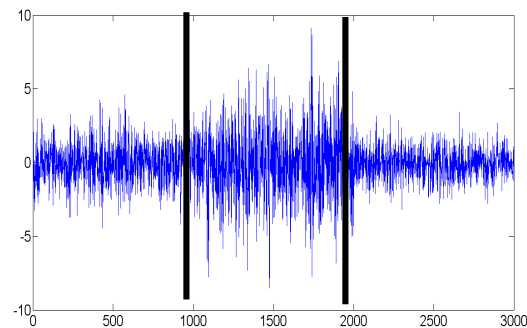
Accept with probability

Accept with probability

$$\min \left[\frac{f(Y_{1:T} | \tilde{\Theta}_j^i, \Gamma_j^i) f(\tilde{\Theta}_j^i)}{f(Y_{1:T} | \Theta_j^i, \Gamma_j^i) f(\Theta_j^i)}, 1 \right]$$

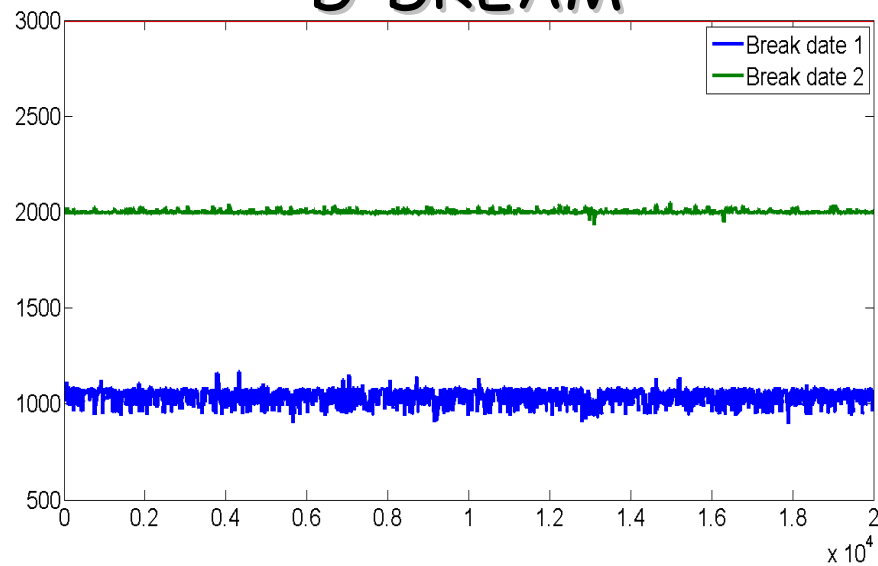
$$\min \left[\frac{f(Y_{1:T} | \Theta_j^i, \tilde{\Gamma}_j^i) f(\tilde{\Gamma}_j^i)}{f(Y_{1:T} | \Theta_j^i, \Gamma_j^i) f(\Gamma_j^i)}, 1 \right]$$

Example



$$T = 3000$$

D-DREAM



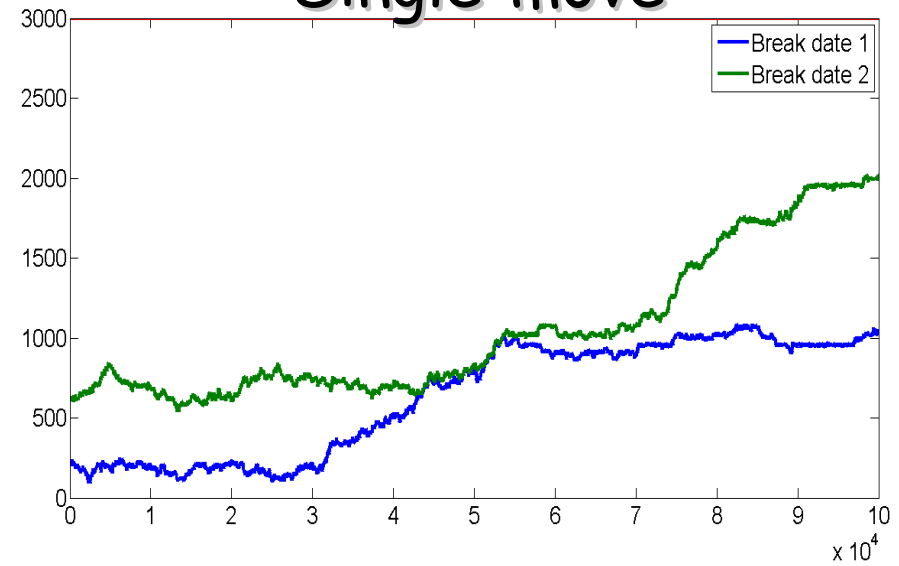
Initial states around [200 600]

$$\text{Corr}(\tau_1^i, \tau_1^{i-200}) = -0.005$$

$$\text{Corr}(\tau_1^i, \tau_1^{i-10}) = 0.54$$

**Convergence after 3.000
MCMC iterations !**

Single-move



Initial state : [200 600]

$$\text{Corr}(\tau_1^i, \tau_1^{i-200}) = 0.996$$

**Convergence after 100.000
MCMC iterations !**

D-DREAM (2014)

Advantages

- Generic method for CP models
- Inference on multiple breaks by marginal likelihood
- Very fast compared to existing algorithms

Drawbacks

- Model selection based on many estimations
- Only applicable to CP models and specific class of recurrent states

CP and MS models

Particle MCMC

CP- and MS-GARCH models

$$y_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = \omega_{s_t} + \alpha_{s_t} y_{t-1}^2 + \beta_{s_t} \sigma_{t-1}^2$$

$$\epsilon_t \sim \text{i.i.d. } N(0, 1)$$

$$s_t | s_{t-1}, P \sim p_{s_{t-1}, s_t}$$

Change-point

$$P = \begin{pmatrix} p_{1,1} & 1 - p_{1,1} & 0 & \dots & 0 \\ 0 & p_{2,2} & 1 - p_{2,2} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Markov-switching

$$P = \begin{pmatrix} p_{1,1} & p_{1,2} & p_{1,3} & \dots & p_{1,K+1} \\ p_{2,1} & p_{2,2} & p_{2,3} & \dots & p_{2,K+1} \\ \dots & \dots & \dots & \dots & \dots \\ p_{K+1,1} & p_{K+1,2} & p_{K+1,3} & \dots & p_{K+1,K+1} \end{pmatrix}$$

Particle MCMC

Sets of parameters :

$$\text{Continuous} \left\{ \begin{array}{l} \Theta = (\theta_1, \theta_2, \dots, \theta_{K+1})' \\ \theta_k = (\omega_k, \alpha_k, \beta_k) \\ P \end{array} \right. \quad \text{State var. } S_{1:T} = \{s_1, \dots, s_T\}$$

MCMC scheme :

- 1) $\Theta | Y_{1:T}, P, S_{1:T} \sim \text{Metropolis-Hastings}$
- 2) $P | Y_{1:T}, \Theta, S_{1:T} \sim \prod_{i=1}^{K+1} \text{Dir}(\eta_1 + n_{i,1}, \dots, \eta_1 + n_{i,K+1})$
- 3) $S_{1:T} | Y_{1:T}, P, \Theta \sim \text{Particle-Gibbs}$

Sampling a full state vector is unfeasible
due to the path dependence issue

Particle MCMC

3) $S_{1:T} | Y_{1:T}, P, \Theta \sim \text{Particle-Gibbs}$

Idea : Approximate the distribution with a SMC algorithm

→ Does not keep invariant the posterior distribution

Andrieu, Doucet and Holenstein (2010)

- Show how to incorporate the SMC into an MCMC
- Allow for Metropolis and Gibbs algorithms
- Introduce the concept of conditional SMC

→ With a conditional SMC, the MCMC exhibits the posterior distribution as invariant one.

Particle MCMC

3) $S_{1:T} | Y_{1:T}, P, \Theta \sim$ Particle-Gibbs

$$f(S_{1:t} | Y_{1:t}) \propto f(y_t | Y_{1:t-1}, S_{1:t}) f(s_t | S_{1:t-1}) \overbrace{f(S_{1:t-1} | Y_{1:t-1})}^{\text{Previous value}}$$

SMC :

1) Initialisation of the particles and weights: $\begin{cases} \{s_1^r\}_{r=1}^R \\ w_r \propto f(y_1 | s_1^r) \end{cases}$

Iterations

$\forall t \in [2, T]$

- Re-sample the particles

$$\forall r \in [1, R]; A_t^r \sim \text{Mult}(W_{t-1})$$

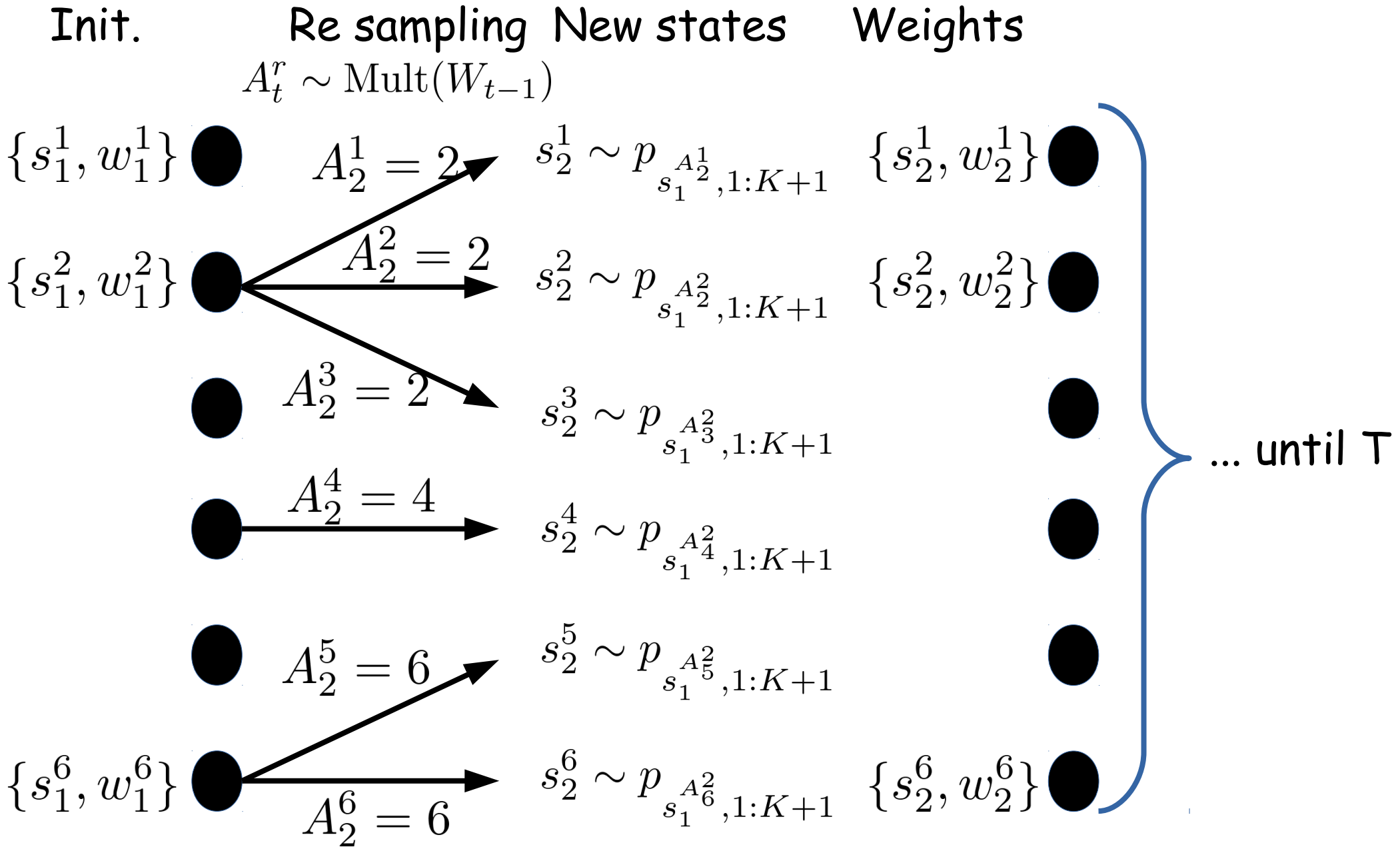
- Generate new states

$$\forall r \in [1, R]; s_t^r \sim \text{Mult}(p_{s_{t-1}, s_t}^{A_t^r})$$

- Compute new weights

$$\forall r \in [1, R]; w_t^r \propto f(y_t | S_{1:t-1}^{A_t^r}, s_t^r) \text{ and } W_t^r = w_t^r / \left(\sum_{i=1}^R w_t^i \right)$$

SMC



Particle Gibbs

- **Conditional SMC** : SMC where the previous MCMC state vector is ensured to survive during the entire SMC sequence.

3) $S_{1:T} | Y_{1:T}, P, \Theta \sim \text{Particle-Gibbs}$

- Launch a conditional SMC
- Sample a state vector as follows :

1) $r \sim \text{Mult}(W_T^1, \dots, W_T^R)$ and set $s_T = s_T^r$

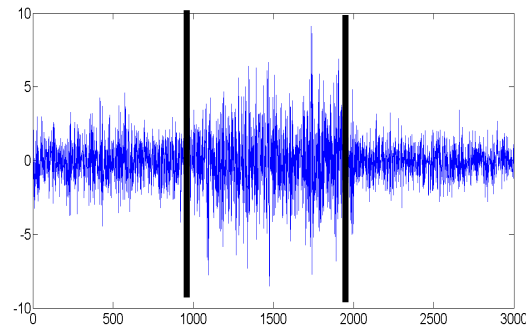
2) From T-1 until 1, retrieve the path of the state : $s_t = s_t^{A_{t+1}^r}$

- Improvements :

1) Incorporation of the APF in the conditional SMC

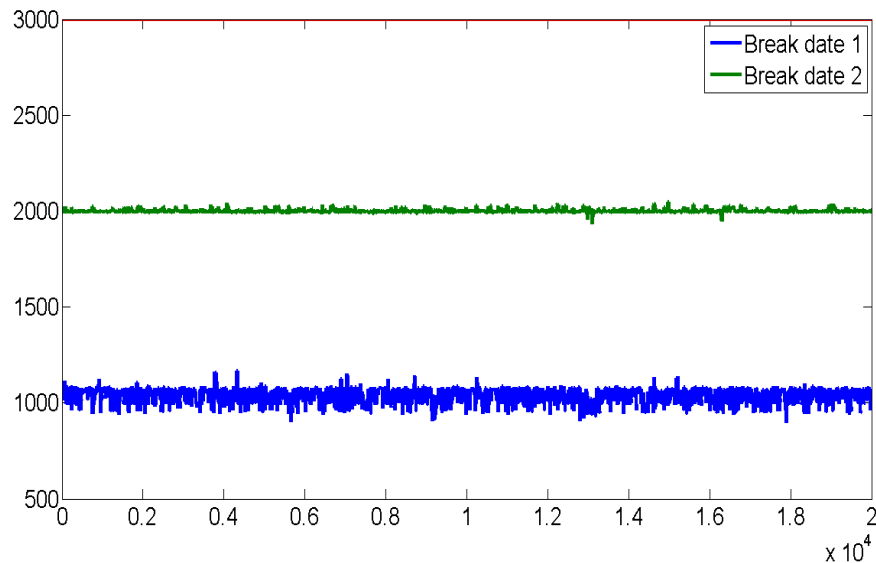
2) Backward sampling as Godsill, Doucet and West (2004)

Example

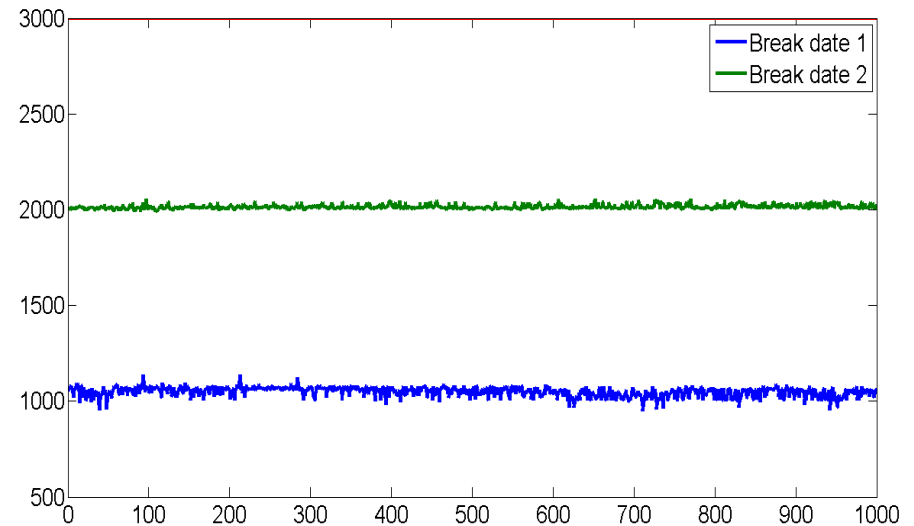


$$T = 3000$$

D-DREAM



PMCMC



Initial states around [200 600]

$$\text{Corr}(\tau_1^i, \tau_1^{i-200}) = -0.005$$

$$\text{Corr}(\tau_1^i, \tau_1^{i-10}) = 0.54$$

Initial state : [200 600]

$$\text{Corr}(\tau_1^i, \tau_1^{i-200}) = 0.05$$

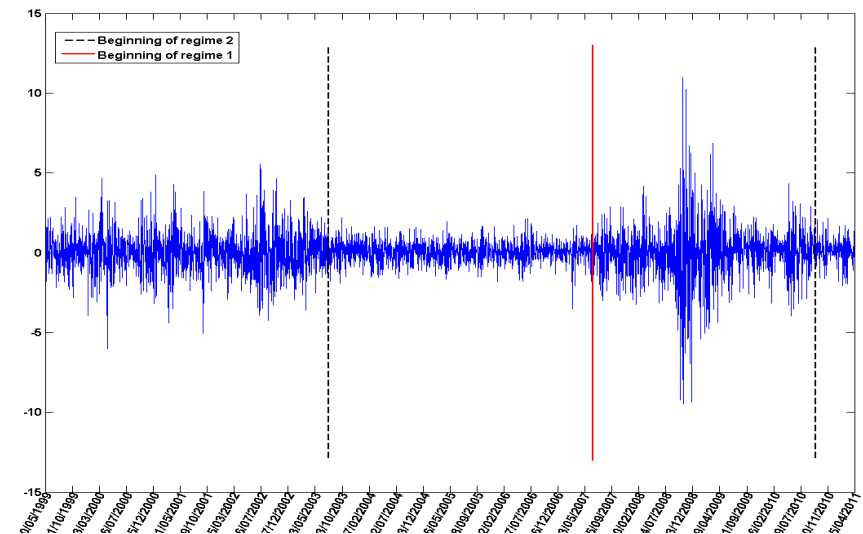
$$\text{Corr}(\tau_1^i, \tau_1^{i-10}) = 0.21$$

PMCMC

S&P 500 daily percentage returns
from May 20, 1999 to April 25, 2011

Table: Marginal log-likelihood values for S&P 500 data

Regimes	1	2	3	4
Change-Point				
BS	-4505.33	-4505.83	-4503.05	-4519.23
Chib	-4504.95	-4505.93	-4502.97	-4516.16
Markov-switching				
BS	-4505.31	-4497.99	-4502.74	
Chib	-4505.08	-4496.04	-4497.73	



Regime	GARCH			CP-GARCH			MS-GARCH		
	σ^2	α	β	σ^2	α	β	σ^2	α	β
1	1.55	0.075	0.915	1.95	0.0849	0.868	2.32	0.089	0.891
2				0.45	0.023	0.931	0.46	0.031	0.901
3				2.75	0.098	0.890			

PMCMC

Various financial time series

Series	Spline-GARCH		GARCH	CP-GARCH		MS-GARCH		
	knots	log-BF	MLL	K+1	log-BF	K+1	log-BF	nswitch
S&P 500	3	5.21	-4505.33	3	2.28	2	7.34	3
DJIA	3	2.99	-4333.43	1	0	2	4.7	3
NASDAQ	3	3.20	-5429.84	1	0	2	1.94	7
NYSE	3	2.40	-4380.62	1	0	2	3.91	13
BAC	4	16.62	-6127.39	3	50.12	3	79.49	11
BA	4	9.10	-6174.57	2	8.9	2	11.48	6
JPM	3	8.82	-6400.27	3	5.17	3	7.22	9
MRK	5	48.78	-6209.73	5	215.39	3	335.23	56
PG	4	16.34	-4842.02	3	24.23	2	33.6	9
Metals	2	6.66	-5267.44	2	11.33	2	14.68	5
Yen/Dollar	1	-3.34	-2982.33	1	0	2	3.05	7

PMCMC (2013)

Advantages

- Generic method for CP and MS models
- Inference on multiple breaks by marginal likelihood
- Very good mixing properties

Drawbacks

- Model selection based on many estimations
- Very computationally demanding
- Difficult to calibrate the number of particles
- Difficult to implement

IHMM-GARCH

CP- and MS-GARCH models

$$y_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = \omega_{s_t} + \alpha_{s_t} y_{t-1}^2 + \beta_{s_t} \sigma_{t-1}^2$$

$$\epsilon_t \sim \text{i.i.d. } N(0, 1)$$

$$s_t | s_{t-1}, P \sim p_{s_{t-1}, s_t}$$

Change-point

$$P = \begin{pmatrix} p_{1,1} & 1 - p_{1,1} & 0 & \dots & 0 \\ 0 & p_{2,2} & 1 - p_{2,2} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Markov-switching

$$P = \begin{pmatrix} p_{1,1} & p_{1,2} & p_{1,3} & \dots & p_{1,K+1} \\ p_{2,1} & p_{2,2} & p_{2,3} & \dots & p_{2,K+1} \\ \dots & \dots & \dots & \dots & \dots \\ p_{K+1,1} & p_{K+1,2} & p_{K+1,3} & \dots & p_{K+1,K+1} \end{pmatrix}$$

IHMM-GARCH

Sets of parameters :

$$\text{Continuous } \left\{ \begin{array}{l} \Theta = (\theta_1, \theta_2, \dots, \theta_{K+1})' \\ \theta_k = (\omega_k, \alpha_k, \beta_k) \\ P \end{array} \right. \quad \text{State var. } S_{1:T} = \{s_1, \dots, s_T\}$$

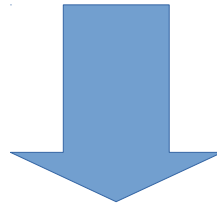
MCMC scheme :

- 1) $\Theta | Y_{1:T}, P, S_{1:T} \sim \text{Metropolis-Hastings}$
- 2) $P | Y_{1:T}, \Theta, S_{1:T} \sim \prod_{i=1}^{K+1} \text{Dir}(\eta_1 + n_{i,1}, \dots, \eta_1 + n_{i,K+1})$
- 3) $S_{1:T} | Y_{1:T}, P, \Theta \sim \text{M-H based on approximate models}$

IHMM-GARCH

3) $S_{1:T} | Y_{1:T}, P, \Theta \sim$ M-H based on approximate models

Sampling a full state vector is infeasible
due to the path dependence issue



Sampling a full state vector from an approximate model
Klaassen or Haas, Mittnik and Paolela



Accept/reject according to the Metropolis-hastings ratio

IHMM-GARCH

Moreover, **Hierarchical dirichlet processes** are used

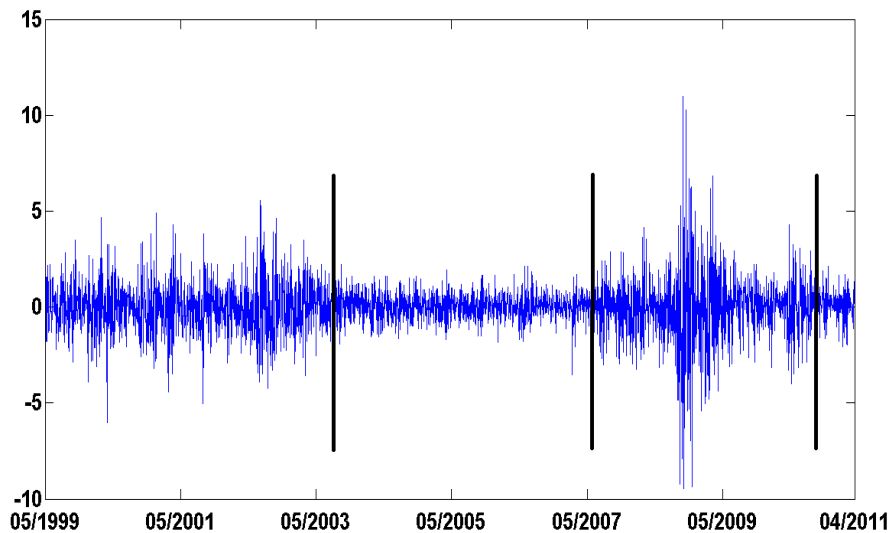
- To infer the number of regime in one estimation
- To include both CP and MS specification in one model

$$P = \begin{pmatrix} p_{1,1} & p_{1,2} & p_{1,3} & \dots & \dots \\ p_{2,1} & p_{2,2} & p_{2,3} & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ p_{K+1,1} & p_{K+1,2} & p_{K+1,3} & \dots & \dots \end{pmatrix}$$

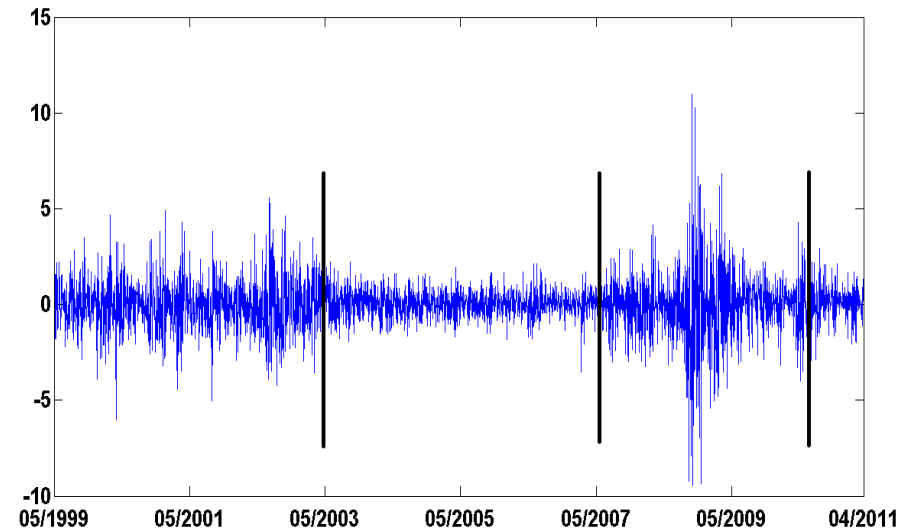
IHMM-GARCH

S&P 500 daily percentage returns
from May 20, 1999 to April 25, 2011

PMCMC



IHMM-GARCH



	Regime 1	Regime 2	Regime 3	Regime 4	Regime 5	Regime 6	Regime 7
Prob.	0	0.6046	0.2075	0.1455	0.0224	0.0196	0.0004

Table: Posterior probabilities of the number of regimes for the S&P500 daily index

IHMM-GARCH (2014)

Advantages

- Generic method for CP and MS models
- Self-determination of the number of breaks
- Self-determination of the specification (CP and/or MS)
- Predictions of breaks
- Very good mixing properties
- Fast MCMC estimation

Drawbacks

- Difficult to implement

References

References

- Bauwens, Preminger, Rombouts (2010)

Bauwens, L.; Preminger, A. & Rombouts, J. '*Theory and Inference for a Markov-switching GARCH Model*', *Econometrics Journal*, 2010, 13, 218-244

- Bauwens, Dufays, De Backer (2011)

Bauwens, L.; Dufays, A. & De Backer, B. '*Estimating and forecasting structural breaks in financial time series*', *CORE discussion paper*, 2011/55, 2011

- Bauwens, Dufays, Rombouts (2013)

Bauwens, L.; Dufays, A. & Rombouts, J. '*Marginal Likelihood for Markov Switching and Change-point GARCH Models*', *Journal of Econometrics*, 2013, 178 (3), 508-522

References

- Dufays (2012)

- Dufays, A. '*Infinite-State Markov-switching for Dynamic Volatility and Correlation Models*', CORE discussion paper, 2012/43, 2012

SMC algorithm :

- He, Maheu (2010)

- He, Z. & Maheu, J. M. '*Real time detection of structural breaks in GARCH models*', Computational Statistics & Data Analysis, 2010, 54, 2628-2640

Without path dependence

- Gray (1996)

Gray, S. *'Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process'*, Journal of Financial Economics, 1996, 42, 27-62

- Dueker (1997)

Dueker, M. *'Markov Switching in GARCH Processes in Mean Reverting Stock Market Volatility'*, Journal of Business and Economics Statistics, 1997, 15, 26-34

- Klaassen (2002)

Klaassen, F. *'Improving GARCH volatility forecasts with regime-switching GARCH'*, Empirical Economics, 2002, 27, 363-394

- Haas, Mittnik and Paolella (2004)

Haas, M.; Mittnik, S. & Paolella, M. *'A New Approach to Markov-Switching GARCH Models'*, Journal of Financial Econometrics, 2004, 2, 493-530