

# Une introduction aux méthodes particulières : suivi de la trajectoire d'un mobile

Alexis Huet

23 mai 2014

# Plan

- 1 Définitions et problématique
  - Chaînes de Markov
  - Chaînes de Markov cachées
  - Problématique
- 2 Algorithmes particuliers
  - Algorithme SIS
  - Algorithme SISR
  - Comparaison des algorithmes
- 3 Applications réelles

# Plan

- 1 Définitions et problématique
  - Chaînes de Markov
  - Chaînes de Markov cachées
  - Problématique
- 2 Algorithmes particuliers
  - Algorithme SIS
  - Algorithme SISR
  - Comparaison des algorithmes
- 3 Applications réelles

# Chaînes de Markov

On se place dans  $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n))$ .

## Définition

Une suite de variables aléatoires  $(X_k)_{k \in \mathbb{N}}$  à valeurs dans  $\mathbb{R}^n$  est une chaîne de Markov si pour tous  $k \geq 1$ ,  $x_0, \dots, x_{k-1} \in \mathbb{R}^n$  et  $A \in \mathcal{L}(\mathbb{R}^n)$  :

$$P(X_k \in A | X_{k-1} = x_{k-1}, \dots, X_0 = x_0) = P(X_k \in A | X_{k-1} = x_{k-1}).$$

## Hypothèse

On suppose que la chaîne de Markov est homogène et admet une famille de densités  $(p(\cdot|x))_{x \in \mathbb{R}^n}$  :

$$P(X_k \in A | X_{k-1} = x_{k-1}) = \int_A p(x_k | x_{k-1}) dx_k.$$

# Chaînes de Markov

On se place dans  $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n))$ .

## Définition

Une suite de variables aléatoires  $(X_k)_{k \in \mathbb{N}}$  à valeurs dans  $\mathbb{R}^n$  est une chaîne de Markov si pour tous  $k \geq 1$ ,  $x_0, \dots, x_{k-1} \in \mathbb{R}^n$  et  $A \in \mathcal{L}(\mathbb{R}^n)$  :

$$P(X_k \in A | X_{k-1} = x_{k-1}, \dots, X_0 = x_0) = P(X_k \in A | X_{k-1} = x_{k-1}).$$

## Hypothèse

On suppose que la chaîne de Markov est homogène et admet une famille de densités  $(p(\cdot|x))_{x \in \mathbb{R}^n}$  :

$$P(X_k \in A | X_{k-1} = x_{k-1}) = \int_A p(x_k | x_{k-1}) dx_k.$$

## Exemple

On se place dans  $\mathbb{R}^2$ . Pour tout  $x \in \mathbb{R}^2$ ,  $B(x, 1)$  boule unité centrée en  $x$ .

### Définition

Pour tout  $x \in \mathbb{R}^2$ , la loi  $Unif(B(x, 1))$  est définie par sa densité :

$$\frac{1}{\pi} \mathbf{1}(\cdot \in B(x, 1)).$$

On choisit comme condition initiale :

$$X_0 \rightsquigarrow Unif(B(0, 1)).$$

Et comme transitions :

$$X_k | (X_{k-1} = x_{k-1}) \rightsquigarrow Unif(B(x_{k-1}, 1)).$$

## Exemple

On se place dans  $\mathbb{R}^2$ . Pour tout  $x \in \mathbb{R}^2$ ,  $B(x, 1)$  boule unité centrée en  $x$ .

### Définition

Pour tout  $x \in \mathbb{R}^2$ , la loi  $Unif(B(x, 1))$  est définie par sa densité :

$$\frac{1}{\pi} \mathbf{1}(\cdot \in B(x, 1)).$$

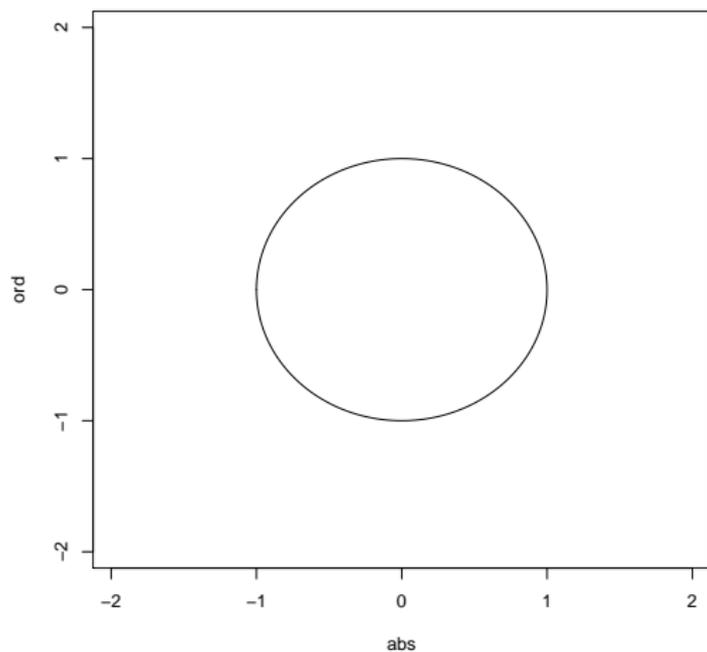
On choisit comme condition initiale :

$$X_0 \rightsquigarrow Unif(B(0, 1)).$$

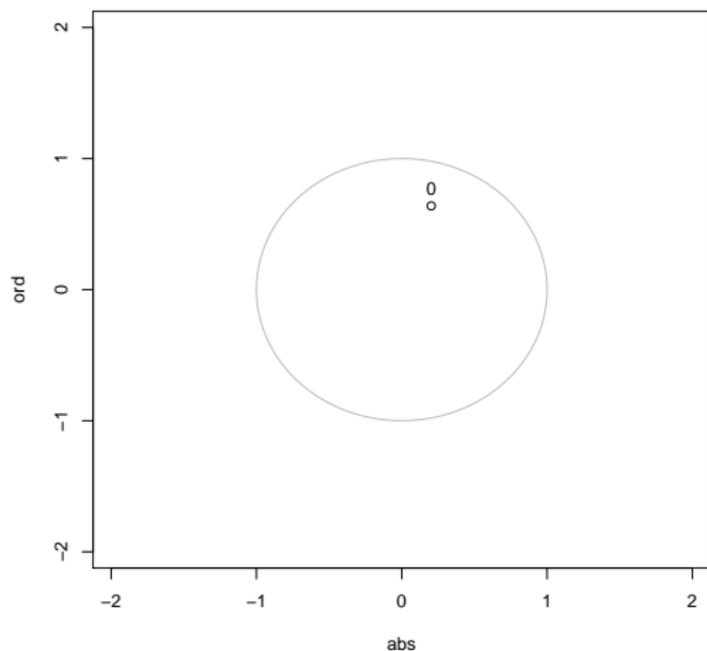
Et comme transitions :

$$X_k | (X_{k-1} = x_{k-1}) \rightsquigarrow Unif(B(x_{k-1}, 1)).$$

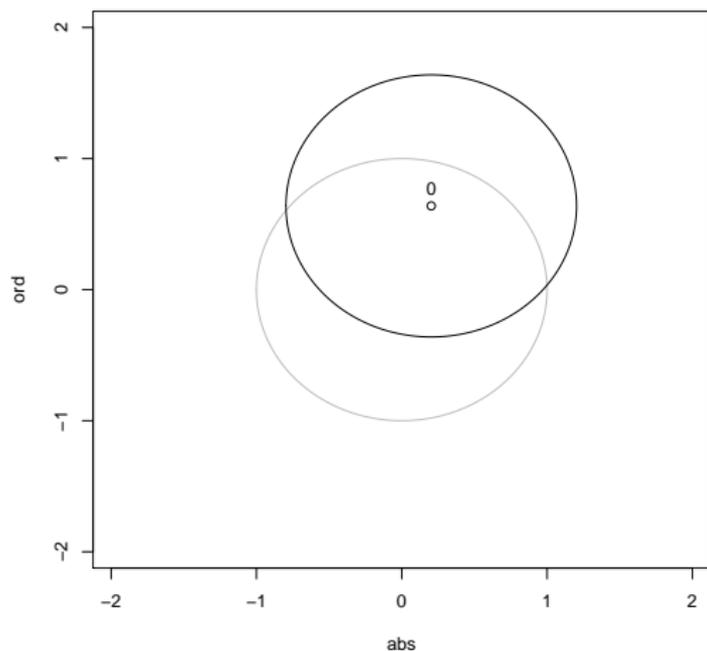
# Exemple



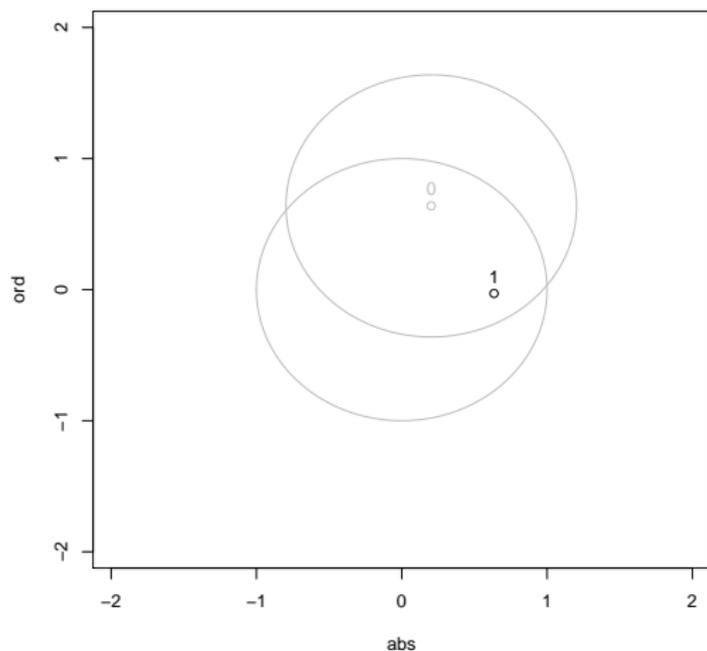
# Exemple



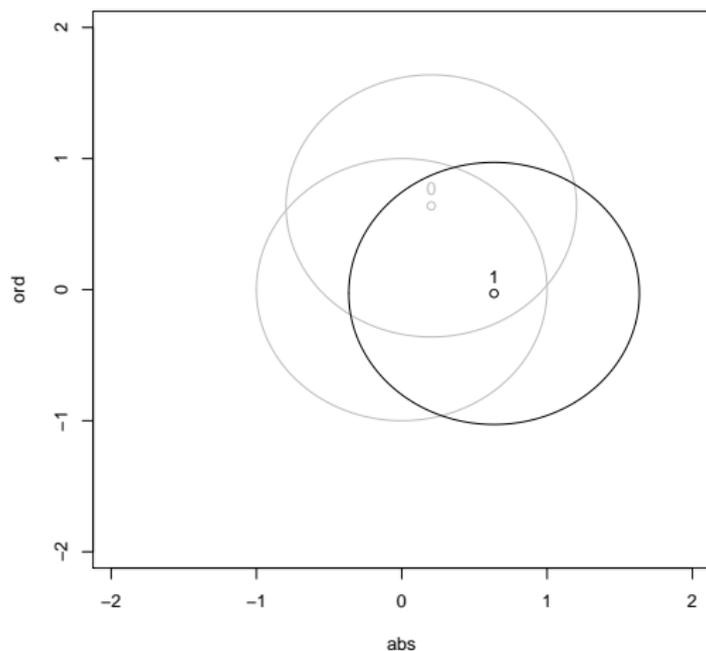
# Exemple



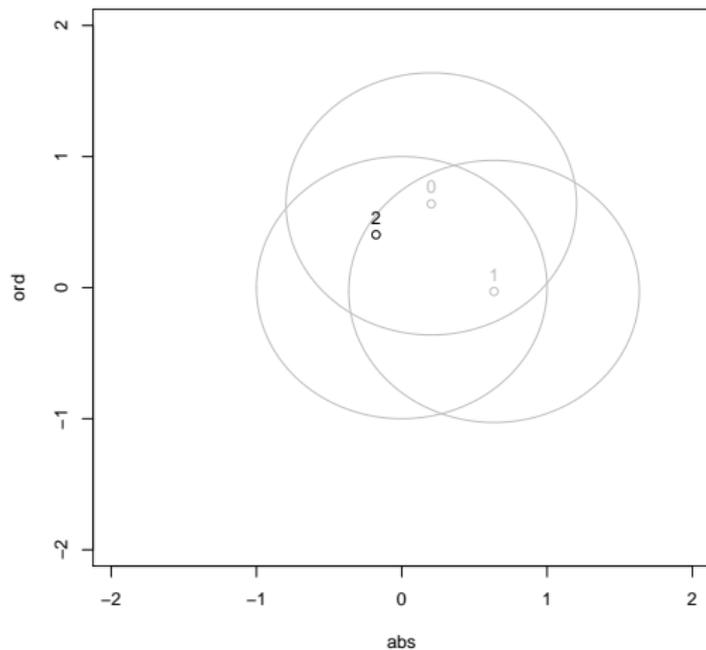
# Exemple



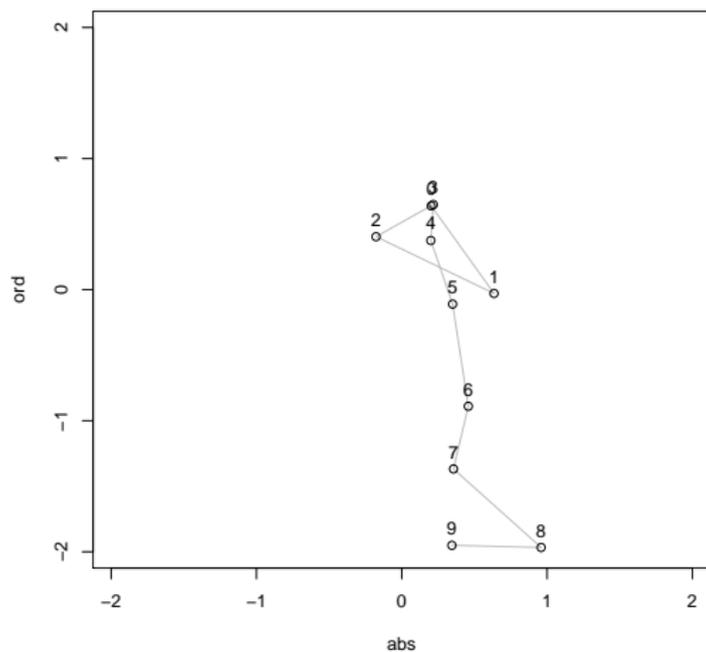
# Exemple



# Exemple



# Exemple

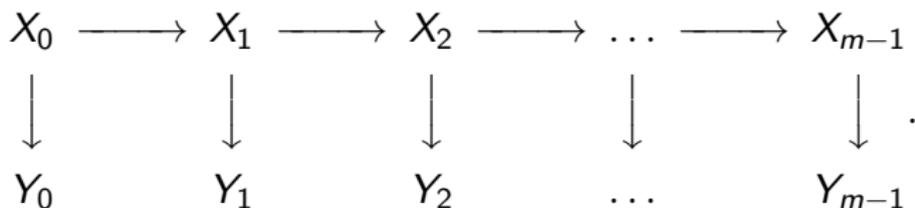


# Chaînes de Markov cachées

## Définition

$(X_k, Y_k)_{k \in 0:m-1}$  est une chaîne de Markov cachée si  $(X_k)_{k \in 0:m-1}$  est une chaîne de Markov, si  $(Y_k)_{k \in 0:m-1}$  sont indépendantes conditionnellement à  $(X_k)_{k \in 0:m-1}$  et si pour tout  $k$ ,  $Y_k$  dépend uniquement de  $X_k$ .

Schématiquement, on a :



# Exemple

On choisit  $(Y_k)$  à valeurs dans  $] -\pi, \pi]$ . Conditionnellement à  $X_k = x_k$ , on définit :

$$Y_k = \text{Arg}(x_k) + \varepsilon$$

où  $\varepsilon \rightsquigarrow \mathcal{N}(0, \sigma^2)$ .

Densité associée :  $y_k \mapsto p(y_k | x_k)$ .

# Exemple

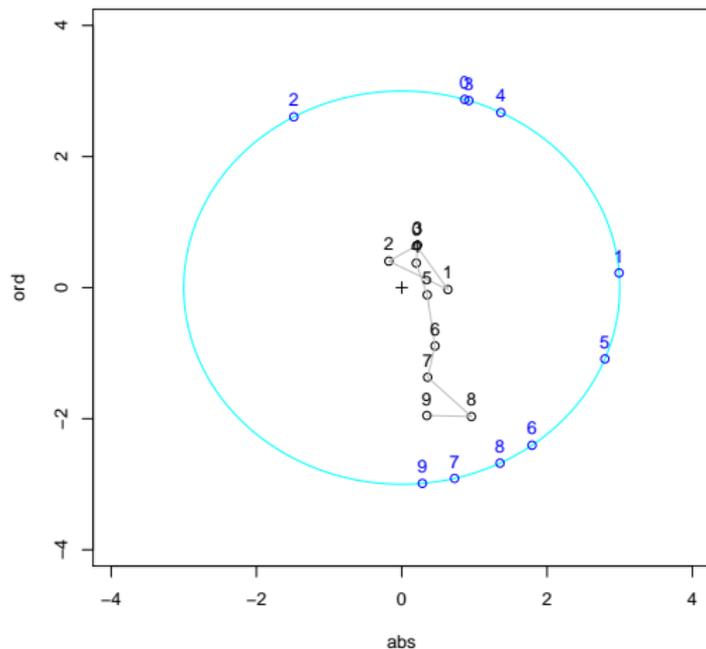
On choisit  $(Y_k)$  à valeurs dans  $] -\pi, \pi]$ . Conditionnellement à  $X_k = x_k$ , on définit :

$$Y_k = \text{Arg}(x_k) + \varepsilon$$

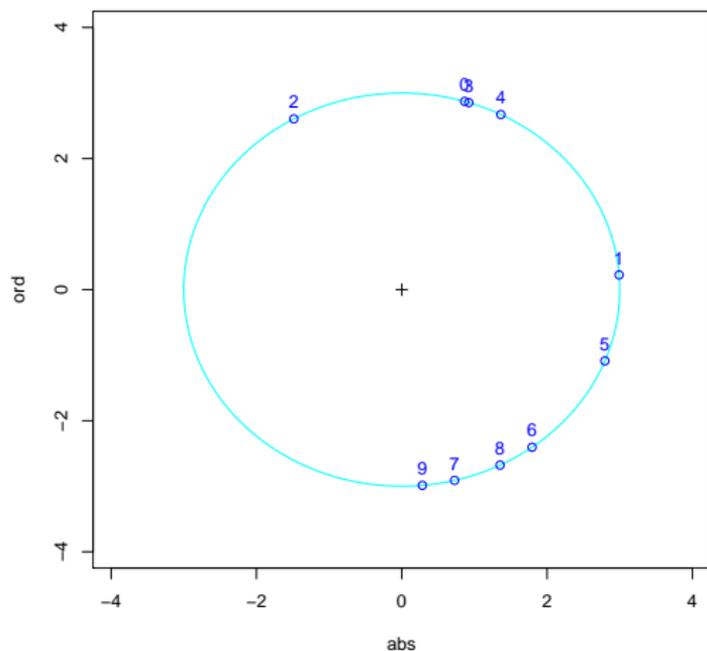
où  $\varepsilon \rightsquigarrow \mathcal{N}(0, \sigma^2)$ .

Densité associée :  $y_k \mapsto p(y_k | x_k)$ .

# Exemple ( $\sigma = 0.1$ )



# Exemple ( $\sigma = 0.1$ )



# Problématique

Dans la suite, la chaîne cachée  $(x_k)$  est inconnue et on dispose uniquement des observations  $(y_k)$ .

But : Reconstituer  $x_{0:m-1} := (x_0, \dots, x_{m-1})$  conditionnellement aux observations.

Plus précisément, simuler un échantillon selon la densité

$$p(x_{0:m-1} | y_{0:m-1}) dx_{0:m-1}.$$

# Plan

- 1 Définitions et problématique
  - Chaînes de Markov
  - Chaînes de Markov cachées
  - Problématique
- 2 Algorithmes particuliers
  - Algorithme SIS
  - Algorithme SISR
  - Comparaison des algorithmes
- 3 Applications réelles

# Algorithme SIS

On cherche à simuler un échantillon de densité  $p(x_0|y_0)dx_0$ .

$$p(x_0|y_0) = \frac{p(x_0, y_0)}{p(y_0)} = \frac{1}{p(y_0)} p(y_0|x_0)p(x_0).$$

- On sait simuler selon  $p(x_0)dx_0$ .
- On sait calculer  $x_0 \mapsto p(y_0|x_0)$ .

# Algorithme SIS

On cherche à simuler un échantillon de densité  $p(x_0|y_0)dx_0$ .

$$p(x_0|y_0) = \frac{p(x_0, y_0)}{p(y_0)} = \frac{1}{p(y_0)} p(y_0|x_0)p(x_0).$$

- On sait simuler selon  $p(x_0)dx_0$ .
- On sait calculer  $x_0 \mapsto p(y_0|x_0)$ .

$$p(x_0|y_0)dx_0 = \frac{1}{p(y_0)}p(y_0|x_0)p(x_0)dx_0.$$

Algorithme d'échantillonnage séquentiel par importance (SIS) :

- Simuler un échantillon de longueur  $N$  selon  $p(x_0)dx_0$  :

$$x_0^{(1)}, \dots, x_0^{(N)}.$$

- Calculer pour chaque particule  $j$  :

$$w_0^{(j)} = \frac{1}{p(y_0)}p(y_0|x_0^{(j)}).$$

- L'échantillon  $(\tilde{x}_0^{(j)})_{j \in 1:N}$  approchant  $p(x_0|y_0)dx_0$  est :

$$\sum_{j=1}^N \frac{w_0^{(j)}}{\sum_{j'=1}^N w_0^{(j')}} \mathbf{1}_{x_0^{(j)}}(dx_0).$$

$$p(x_0|y_0)dx_0 = \frac{1}{p(y_0)}p(y_0|x_0)p(x_0)dx_0.$$

Algorithme d'échantillonnage séquentiel par importance (SIS) :

- Simuler un échantillon de longueur  $N$  selon  $p(x_0)dx_0$  :

$$x_0^{(1)}, \dots, x_0^{(N)}.$$

- Calculer pour chaque particule  $j$  :

$$w_0^{(j)} = p(y_0|x_0^{(j)}).$$

- L'échantillon  $(\tilde{x}_0^{(j)})_{j \in 1:N}$  approchant  $p(x_0|y_0)dx_0$  est :

$$\sum_{j=1}^N \frac{w_0^{(j)}}{\sum_{j'=1}^N w_0^{(j')}} \mathbf{1}_{x_0^{(j)}}(dx_0).$$



$$p(x_{0:i}|y_{0:i})dx_{0:i} = \frac{1}{p(y_{0:i})} p(y_0|x_0) \dots p(y_i|x_i) p(x_0) p(x_1|x_0) \dots p(x_i|x_{i-1}) dx_{0:i}.$$

Algorithme d'échantillonnage séquentiel par importance (pas  $i$ ).

Pour toute particule  $j \in 1 : N$  :

- Conditionnellement à  $x_{i-1}^{(j)}$ , simuler un selon  $p(x_i|x_{i-1})dx_i$  :

$$x_i^{(1)}, \dots, x_i^{(N)}.$$

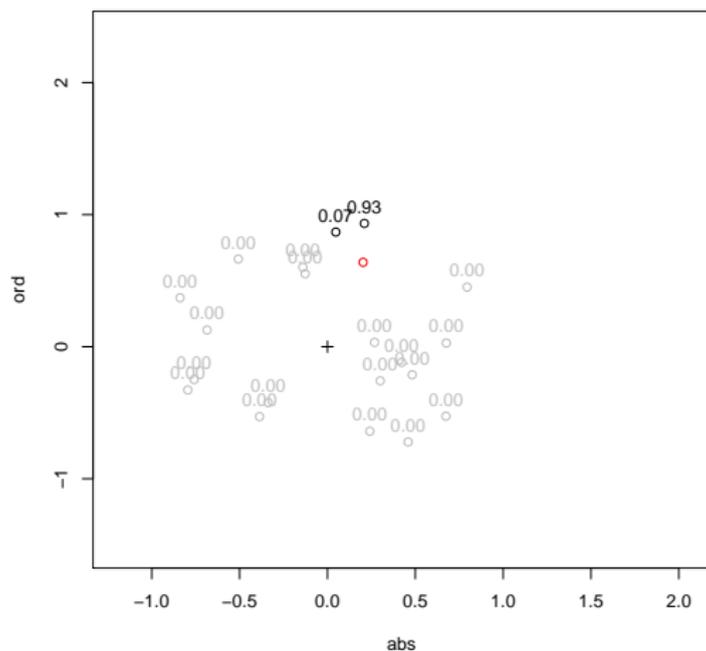
- Calculer pour chaque particule  $j$  :

$$w_i^{(j)} = w_{i-1}^{(j)} \times p(y_i|x_i^{(j)}).$$

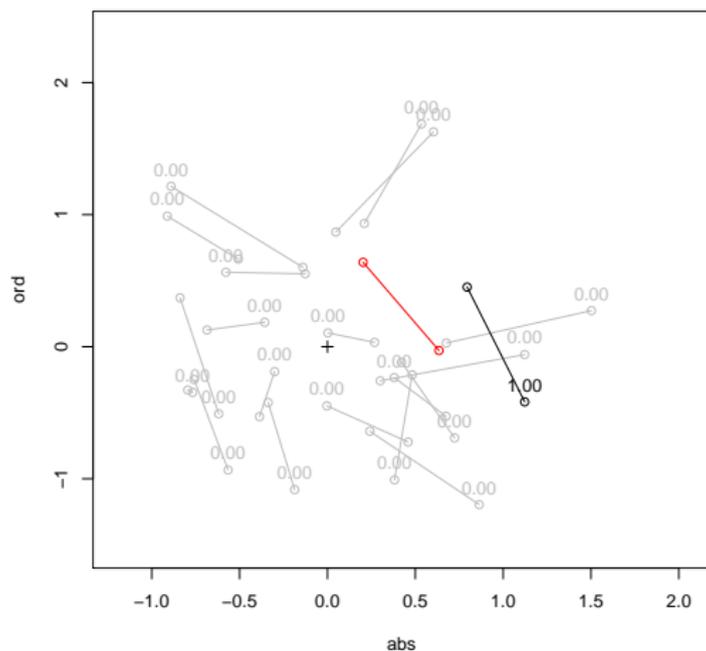
- L'échantillon  $(\tilde{x}_{1:i}^{(j)})_{j \in 1:N}$  approchant  $p(x_{1:i}|y_i)dx_{1:i}$  est :

$$\sum_{j=1}^N \frac{w_i^{(j)}}{\sum_{j'=1}^N w_i^{(j')}} \mathbf{1}_{x_{1:i}^{(j)}}(dx_{1:i}).$$

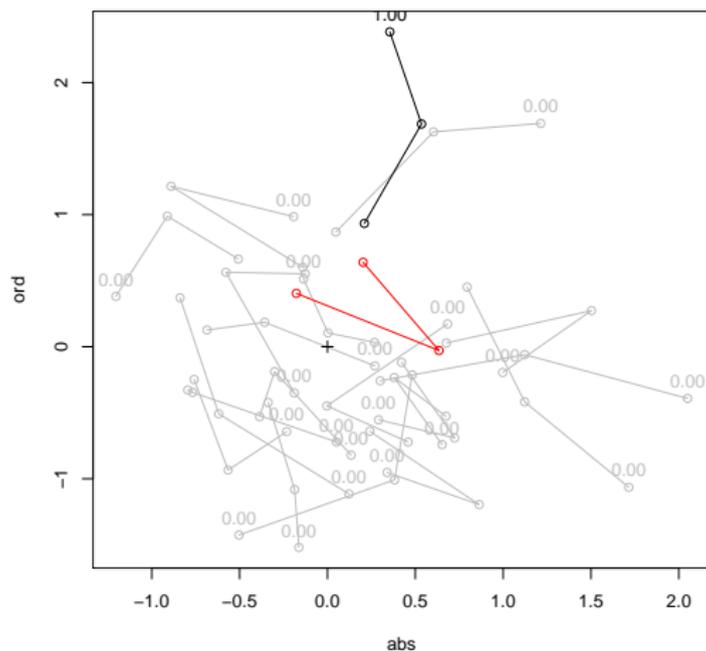
# Exemple (SIS) $N = 20$



# Exemple (SIS) $N = 20$



# Exemple (SIS) $N = 20$



# Convergence et problèmes de dégénérescence (SIS)

Convergence :

- à  $m$  fixé, théorème central limite quand  $N \rightarrow +\infty$ .

Problèmes de dégénérescence :

- beaucoup de particules avec un poids relatif très petit,
- à  $N$  fixé, à partir d'un certain nombre de sites, une seule particule à un poids relatif important.

→ Rééchantillonnage des particules.

# Convergence et problèmes de dégénérescence (SIS)

Convergence :

- à  $m$  fixé, théorème central limite quand  $N \rightarrow +\infty$ .

Problèmes de dégénérescence :

- beaucoup de particules avec un poids relatif très petit,
- à  $N$  fixé, à partir d'un certain nombre de sites, une seule particule à un poids relatif important.

→ Rééchantillonnage des particules.

$$p(x_0|y_0)dx_0 = \frac{1}{p(y_0)}p(y_0|x_0)p(x_0)dx_0.$$

### Algorithme SISR (premier site)

- Simuler un échantillon de longueur  $N$  selon  $p(x_0)dx_0$  :

$$x_0^{(1)}, \dots, x_0^{(N)}.$$

- Calculer pour chaque particule  $j$  :

$$w_0^{(j)} = p(y_0|x_0^{(j)}).$$

- Simuler  $(\tilde{x}_0^{(j)})_{j \in 1:N}$  selon :

$$\sum_{j=1}^N \frac{w_0^{(j)}}{\sum_{j'=1}^N w_0^{(j')}} \mathbf{1}_{x_0^{(j)}}(dx_0)$$

et poser  $(\tilde{w}_0^{(j)})_{j \in 1:N} \equiv 1$ .

Algorithme SISR (site  $i$ ). Pour toute particule  $j \in 1 : N$  :

- Conditionnellement à  $\tilde{x}_{i-1}^{(j)}$ , simuler un selon  $p(x_i | \tilde{x}_{i-1}) dx_i$  :

$$x_i^{(1)}, \dots, x_i^{(N)}.$$

- Calculer pour chaque particule  $j$  :

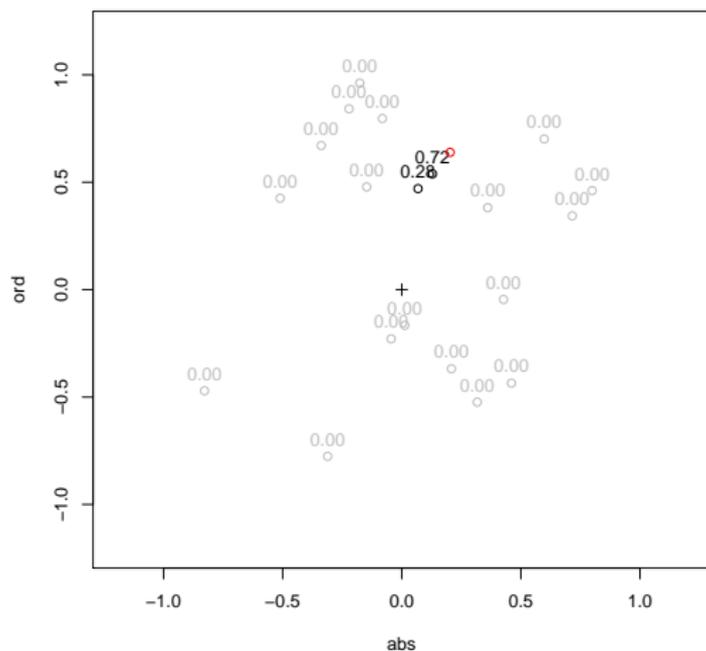
$$w_i^{(j)} = \tilde{w}_{i-1}^{(j)} \times p(y_i | x_i^{(j)}) = p(y_i | x_i^{(j)}).$$

- Simuler  $(\tilde{x}_i^{(j)})_{j \in 1:N}$  selon :

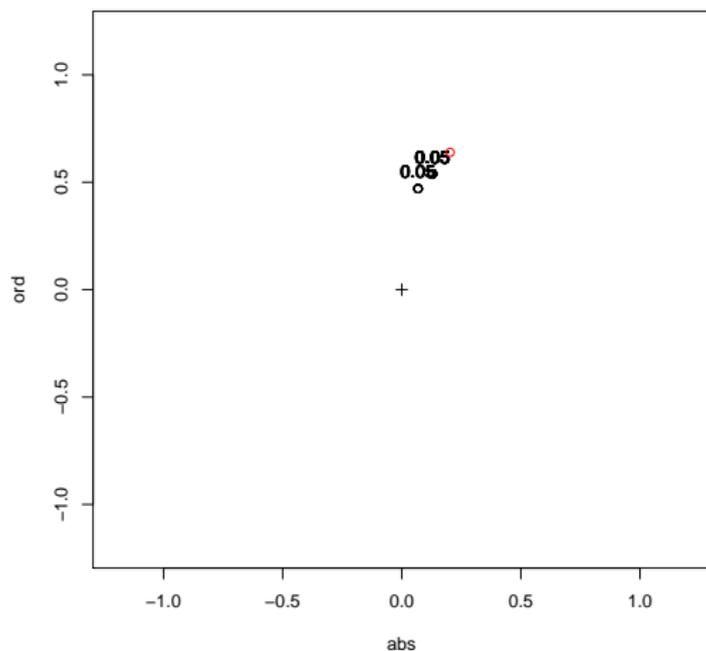
$$\sum_{j=1}^N \frac{w_i^{(j)}}{\sum_{j'=1}^N w_i^{(j')}} \mathbf{1}_{x_i^{(j)}}(dx_i)$$

et poser  $(\tilde{w}_i^{(j)})_{j \in 1:N} \equiv 1$ .

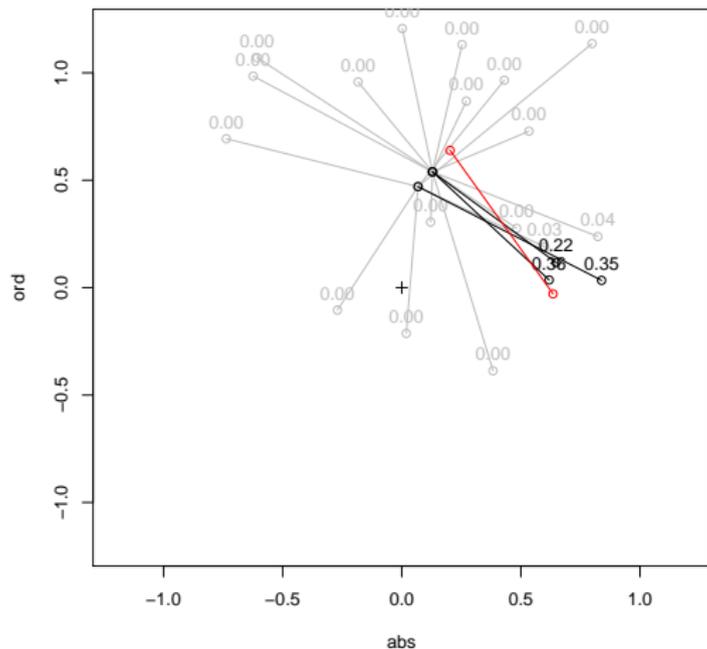
# Exemple (SISR)



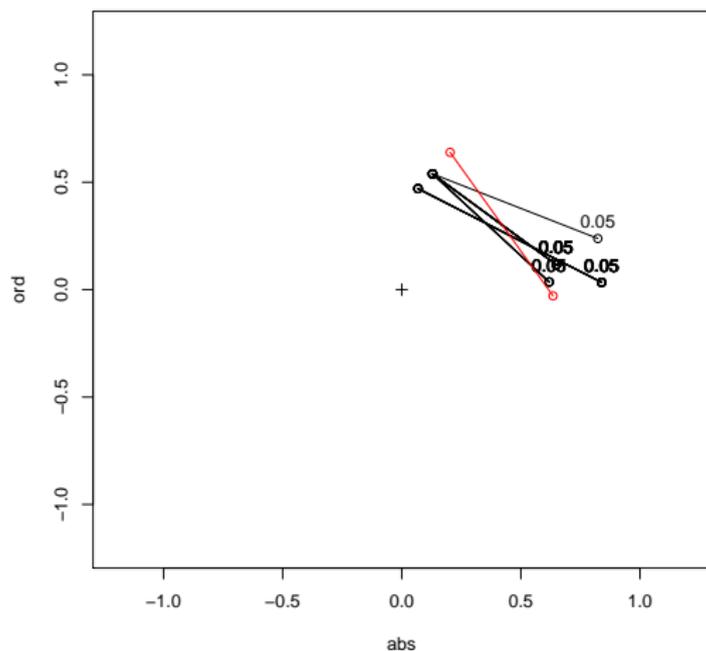
# Exemple (SISR)



# Exemple (SISR)



# Exemple (SISR)



# Convergence (SISR)

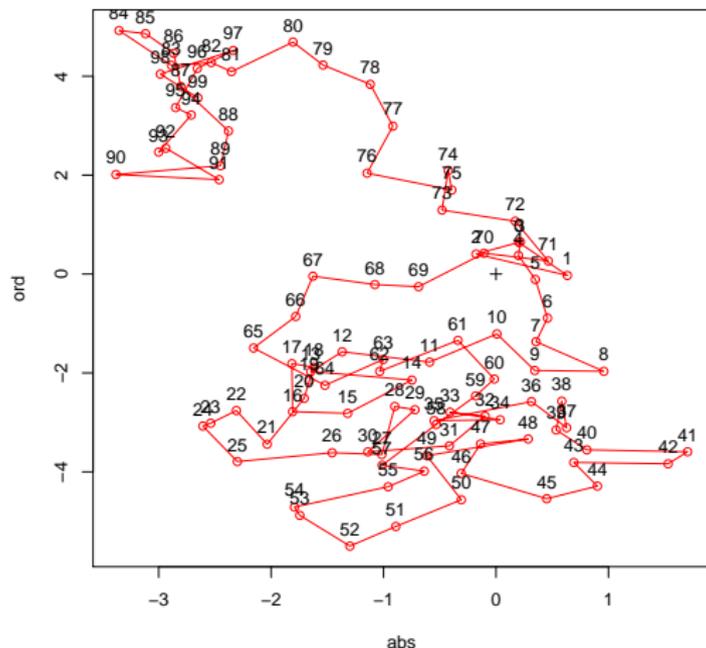
Élimination du problème de dégénérescence des poids.

Convergence :

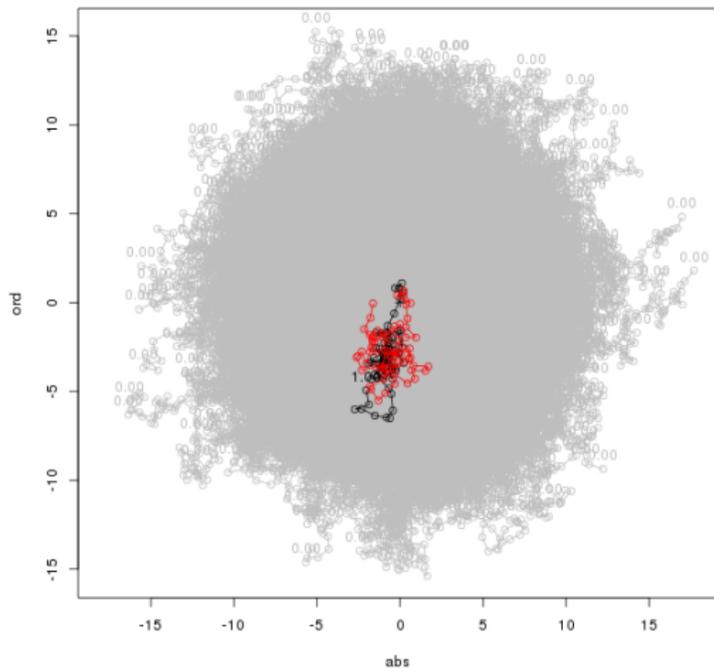
- à  $m$  fixé, théorème central limite quand  $N \rightarrow +\infty$ .

# Exemple

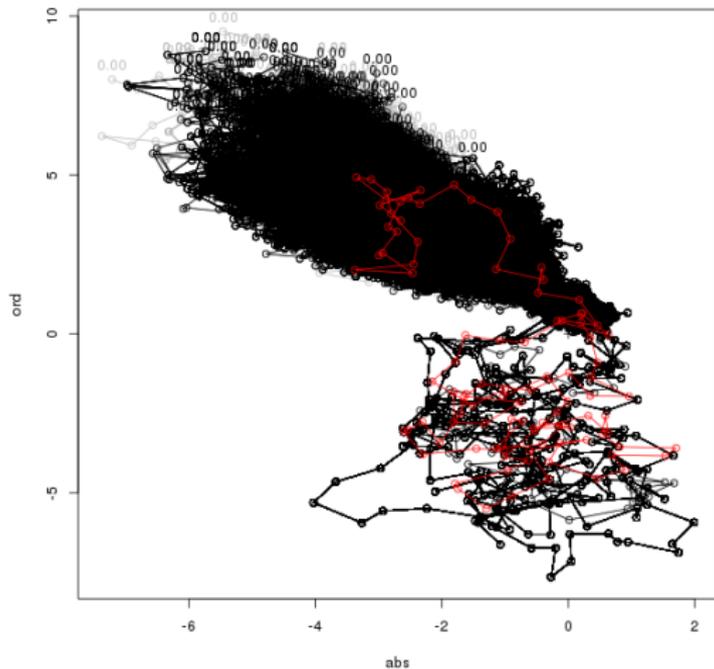
Comparaison avec  $m = 100$  sites et  $N = 10000$  particules.



# Exemple (SIS)

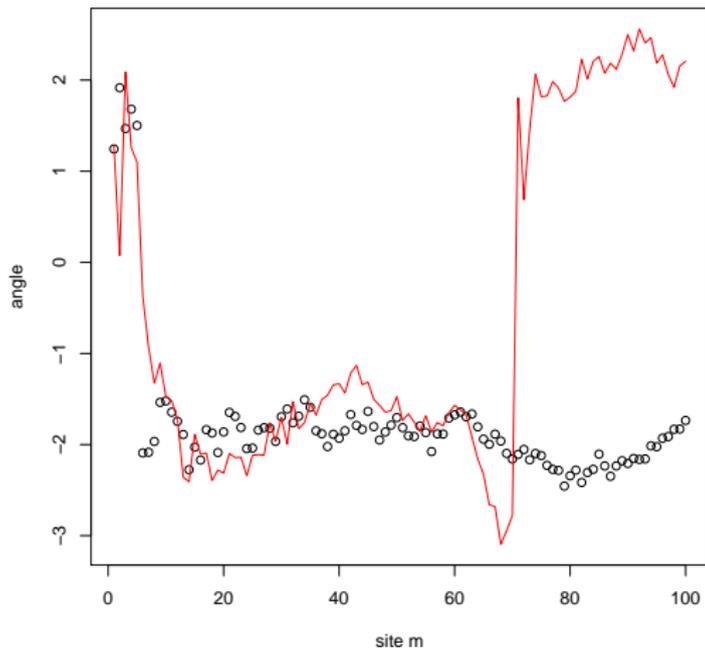


# Exemple (SISR)



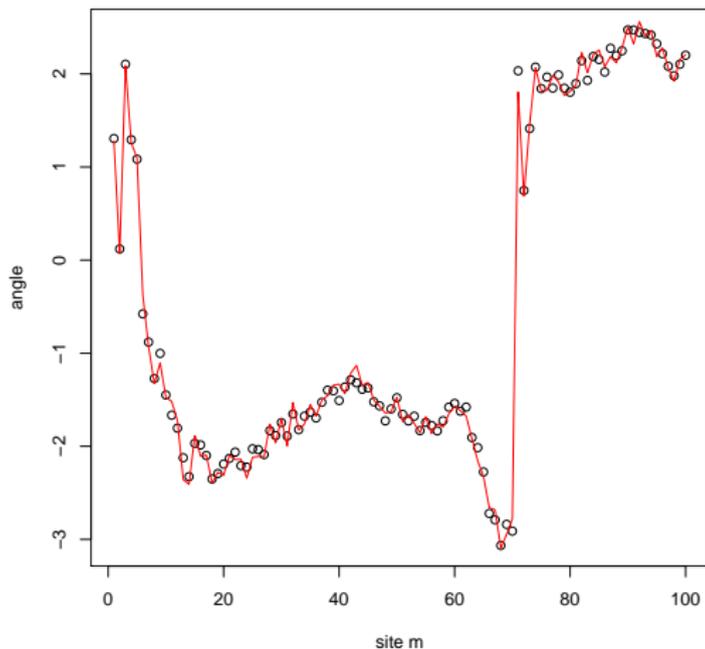
# Exemple (SIS)

Angle observé (rouge) et de la 'meilleure' particule (noir)



# Exemple (SISR)

Angle observé (rouge) et de la 'meilleure' particule (noir)



# Plan

- 1 Définitions et problématique
  - Chaînes de Markov
  - Chaînes de Markov cachées
  - Problématique
- 2 Algorithmes particuliers
  - Algorithme SIS
  - Algorithme SISR
  - Comparaison des algorithmes
- 3 Applications réelles

# Suivi de trajectoire

- Observations : position bruitée.
- Etats cachés : position réelle.
- Paramètres : comportement du mobile.

# Reconnaissance vocale

- Observations : un mot prononcé, découpé toutes les 15ms.
- Etats cachés : phonèmes qui ont conduit au mot prononcé.
- Paramètres : ensemble des mots d'un dictionnaire.

# Évolution moléculaire

- Observations : séquence d'ADN des différentes espèces aux feuilles de l'arbre.
- Etats cachés : séquence d'ADN du temps de l'ancêtre commun au temps présent.
- Paramètres : taux de sauts de mutation, longueur des branches de l'arbre.

# Évolution moléculaire

$t = -1$     ACGAGGTGA  
              ACAAGGTGA  
              ACAGGGTGA  
              ACAGGGCGA  
              ACAGGGCAA

$t = 0$       ACAGGGCAA

# Évolution moléculaire

$i$	1	2	3	4	5	6	7	8	9
$t = -1$	A	C	G	A	G	G	T	G	A
	A	C	A	A	G	G	T	G	A
	A	C	A	G	G	G	T	G	A
	A	C	A	G	G	G	C	G	A
	A	C	A	G	G	G	C	A	A
$t = 0$	A	C	A	G	G	G	C	A	A

# Évolution moléculaire

$i$	1	2	3	4	5	6	7	8	9
$t = -1$	?	?	?	?	?	?	?	?	?
	?	?	?	?	?	?	?	?	?
	?	?	?	?	?	?	?	?	?
	?	?	?	?	?	?	?	?	?
	?	?	?	?	?	?	?	?	?
$t = 0$	A	C	A	G	G	G	C	A	A

- [1] Olivier Cappé, Eric Moulines, and Tobias Rydén. *Inference in hidden Markov models*. Springer, 2005
- [2] Neil Gordon, David Salmond, and Adrian Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEEE Proceedings F (Radar and Signal Processing)*, 1993.
- [3] Lawrence Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 1989.