Introduction to particle filters: a trajectory tracking example

Alexis Huet

23 May 2014

Outline

1. Definitions and issues
   - Markov chains
   - Hidden Markov models
   - Issues
2. Particle filtering algorithms
   - Sequential Importance Sampling algorithm
   - Sequential Importance Sampling Resampling algorithm
   - Comparison of algorithms
3. Real applications

Outline

1. Definitions and issues
   - Markov chains
   - Hidden Markov models
   - Issues
2. Particle filtering algorithms
   - Sequential Importance Sampling algorithm
   - Sequential Importance Sampling Resampling algorithm
   - Comparison of algorithms
3. Real applications

Markov chains

We consider the following state space : $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n))$.

Definition

A sequence of random variables $(X_k)_{k\in\mathbb{N}}$ taking values in $\mathbb{R}^n$ is a Markov chain if for all $k \geq 1$, $x_0, \ldots x_{k-1} \in \mathbb{R}^n$ and $A \in \mathcal{L}(\mathbb{R}^n)$:

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

Hypothesis

In the sequel, the Markov chain is homogeneous and admits a family of density functions $(p(.|x))_{x\in\mathbb{R}^n}$ such that :

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

Example

The state space is here $\mathbb{R}^2$. For all $x \in \mathbb{R}^2$, let $B(x,1)$ be the unit ball centered in x.

Definition

For all $x \in \mathbb{R}^2$, the distribution $Unif(B(x,1))$ is defined by its density :

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

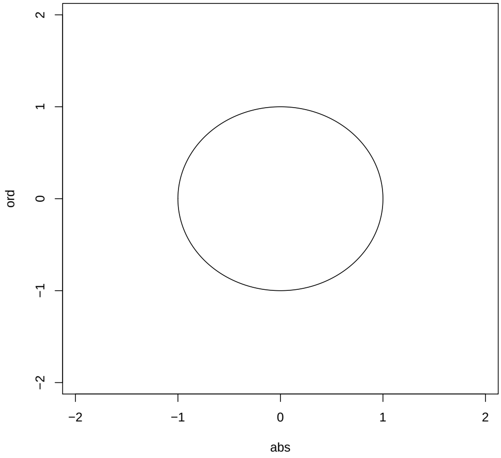
For the initial condition, we take :

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

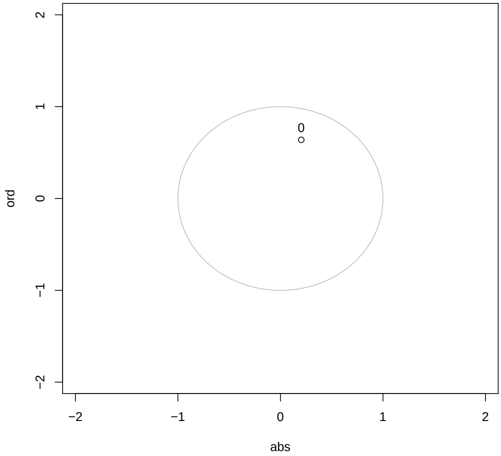
And for transition distributions :

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

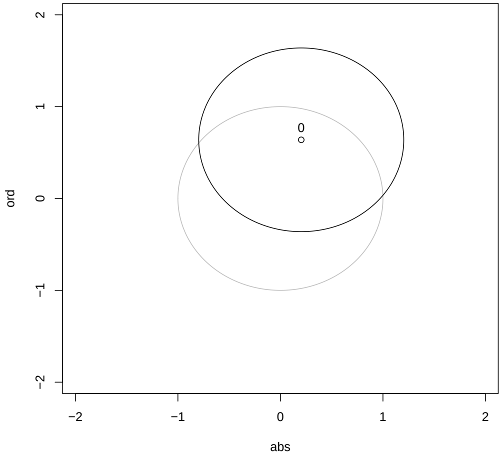
Example



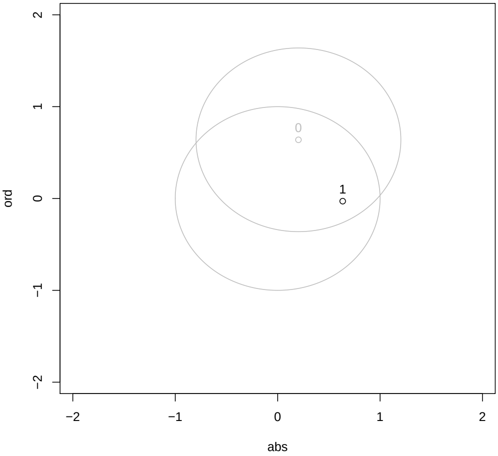
Example



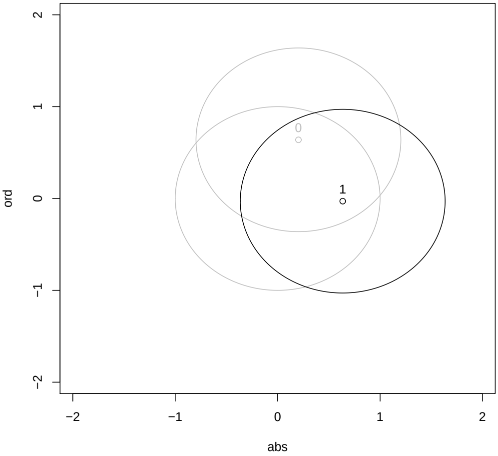
Example



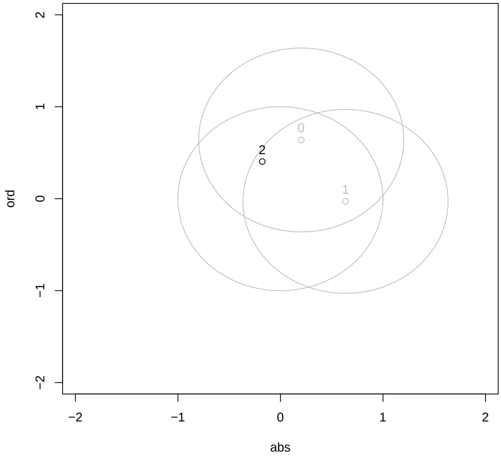
Example



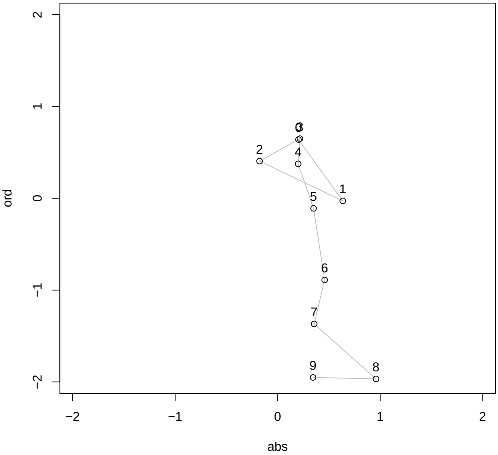
Example



Example



Example



Hidden Markov models

Definition

$(X_k, Y_k)_{k \in 0:m-1}$ is a hidden Markov model if : $(X_k)_{k \in 0:m-1}$ is a Markov chain, $(Y_k)_{k \in 0:m-1}$ are independent conditionally to $(X_k)_{k \in 0:m-1}$ and for all k, Y_k depends only on X_k.

Schematically, we have :

$$
\begin{array}{ccccccccc}
X_0 & \longrightarrow & X_1 & \longrightarrow & X_2 & \longrightarrow & \dots & \longrightarrow & X_{m-1} \\
\downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
Y_0 & & Y_1 & & Y_2 & & \dots & & Y_{m-1}
\end{array}
.
$$

Example

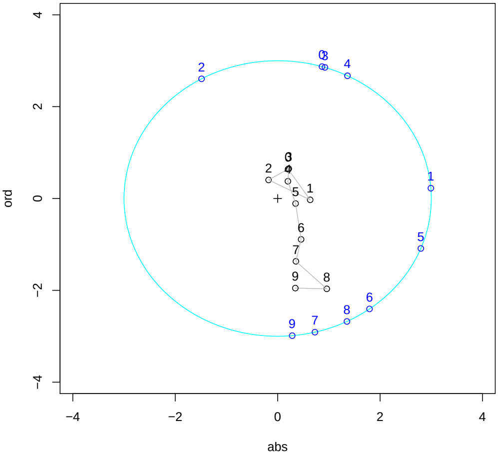
We choose (Y_k) taking values in $]-\pi,\pi]$. Conditionally to $X_k = x_k$, we define :

$$Y_k = Arg(x_k) + \varepsilon$$

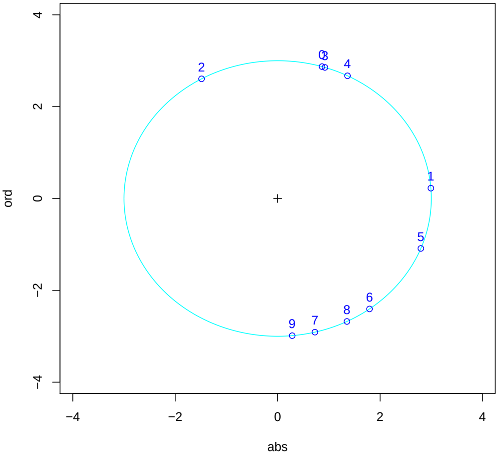
where $\varepsilon \rightsquigarrow \mathcal{N}(0,\sigma^2)$.

Related density : $y_k \longmapsto p(y_k|x_k)$.

Example ($\sigma = 0.1$)



Example ($\sigma = 0.1$)



Issues

Now, the hidden chain (x_k) is unknown and we only have the observations (y_k).

Aim : reconstitute $x_{0:m-1} := (x_0, \ldots, x_{m-1})$ conditionally to the observations.

Specifically, to generate a sample according to the following density :

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

Outline

1. Definitions and issues
   - Markov chains
   - Hidden Markov models
   - Issues
2. Particle filtering algorithms
   - Sequential Importance Sampling algorithm
   - Sequential Importance Sampling Resampling algorithm
   - Comparison of algorithms
3. Real applications

Sequential Importance Sampling (SIS) algorithm

We try to simulate a sample from the density $p(x_0|y_0)dx_0$.

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

- We can generate a sample from $p(x_0)dx_0$.
- We can compute $x_0 \mapsto p(y_0|x_0)$.

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

SIS algorithm (first step) :

- Generate a sample of length N from $p(x_0)dx_0$:

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

- Compute for each particle j :

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

- The sample $(\tilde{x}_0^{(j)})_{j\in 1:N}$ which approximates $p(x_0|y_0)dx_0$ is :

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

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

SIS algorithm (first step) :

- Generate a sample of length N from $p(x_0)dx_0$:

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

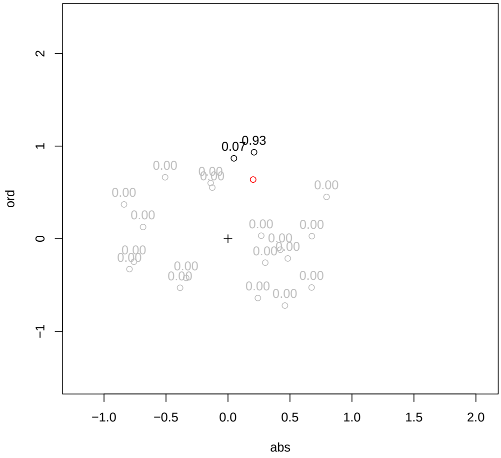
- Compute for each particle j :

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

- The sample $(\tilde{x}_0^{(j)})_{j \in 1:N}$ which approximates $p(x_0|y_0)dx_0$ is :

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

Example (SIS) $N = 20$

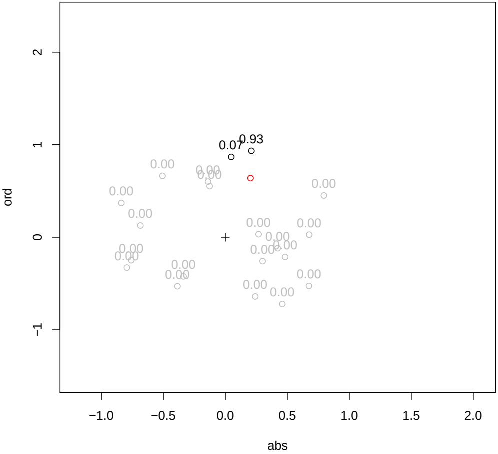


$$p(x_{0:i}|y_{0:i})dx_{0:i} = \frac{1}{p(y_{0:i})} p(y_0|x_0)\dots p(y_i|x_i)p(x_0)p(x_1|x_0)\dots p(x_i|x_{i-1})dx_{0:i}.$$

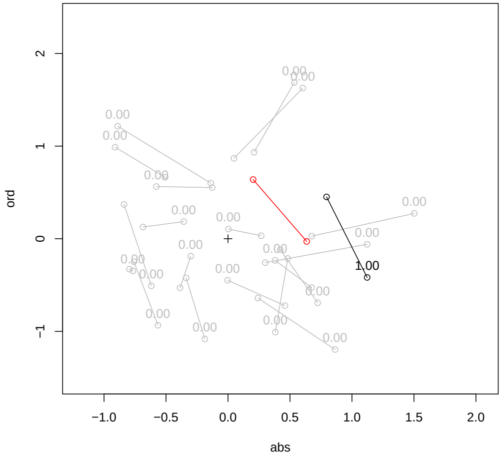
SIS algorithm (step i). For all particle $j \in 1:N$:

- Conditionally to $x_{i-1}^{(j)}$, generate a sample according to $p(x_i|x_{i-1})dx_i$:
$$x_i^{(1)}, \dots, x_i^{(N)}.$$
- Compute for each particle j :
$$w_i^{(j)} = w_{i-1}^{(j)} \times p(y_i|x_i^{(j)}).$$
- The sample $(\tilde{x}_{1:i}^{(j)})_{j \in 1:N}$ which approximates $p(x_{1:i}|y_i)dx_{1:i}$ is :
$$\sum_{j=1}^{N} \frac{w_i^{(j)}}{\sum_{j'=1}^{N} w_i^{(j')}} \mathbf{1}_{x_{1:i}^{(j)}}(dx_{1:i}).$$

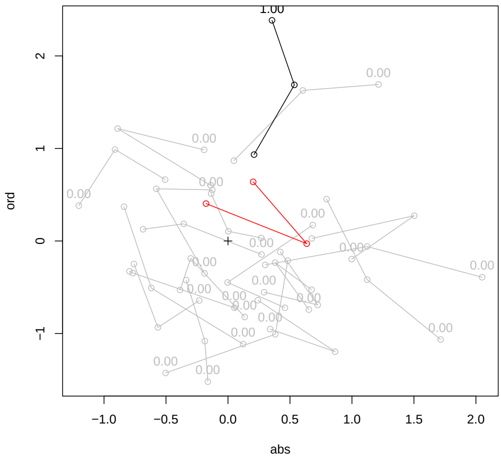
Example (SIS) $N = 20$



Example (SIS) $N = 20$



Example (SIS) $N = 20$



Convergence and degeneracy problem (SIS)

Convergence :

- for m the length of the chain fixed, central limit theorem when the number of particles $N \to +\infty$.

Degeneracy problem :

- many particles have a relative weight close to 0%,
- for N fixed, from a certain number of sites, only one particle has an important relative weight.

$\to$ Resampling of the particles.

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

SISR algorithm (first step)

- Generate a sample of length N according to $p(x_0)dx_0$:

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

- Compute for each particle j :

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

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

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

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

SISR algorithm (step i). For all particle $j \in 1:N$:

- Conditionally to $\tilde{x}_{i-1}^{(j)}$, generate a sample according to $p(x_i|\tilde{x}_{i-1})dx_i$:

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

- Compute for each particle j :

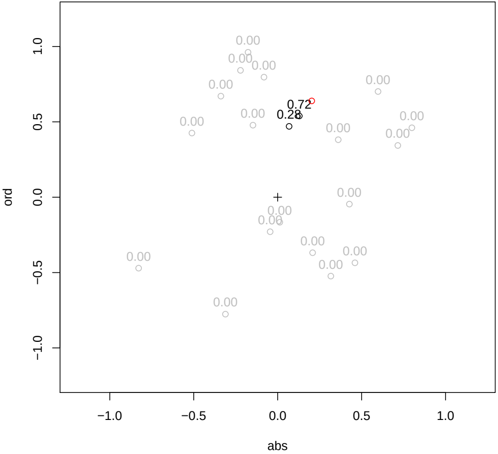
$$w_i^{(j)} = \tilde{w}_{i-1}^{(j)} \times p(y_i|x_i^{(j)}) = p(y_i|x_i^{(j)}).$$

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

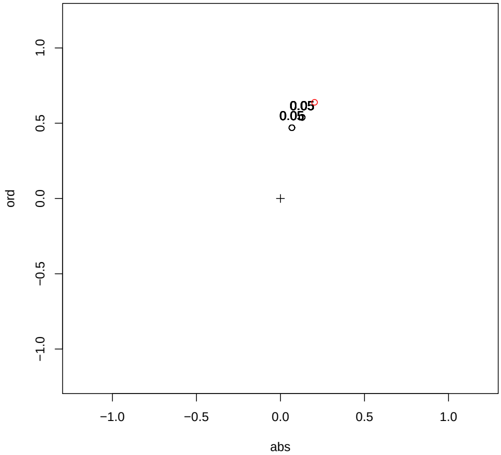
$$\sum_{j=1}^{N} \frac{w_i^{(j)}}{\sum_{j'=1}^{N} w_i^{(j')}} \mathbf{1}_{x_i^{(j)}}(dx_i)$$

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

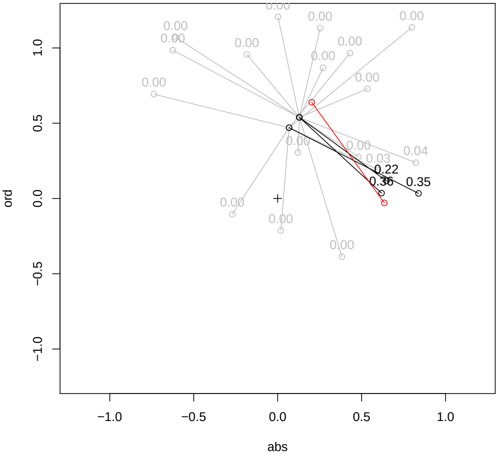
Example (SISR)



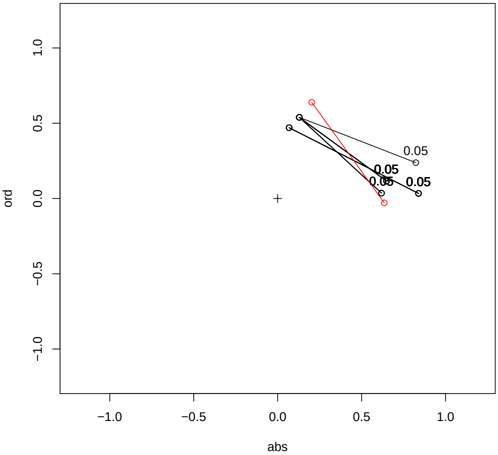
Example (SISR)



Example (SISR)



Example (SISR)



Convergence (SISR)

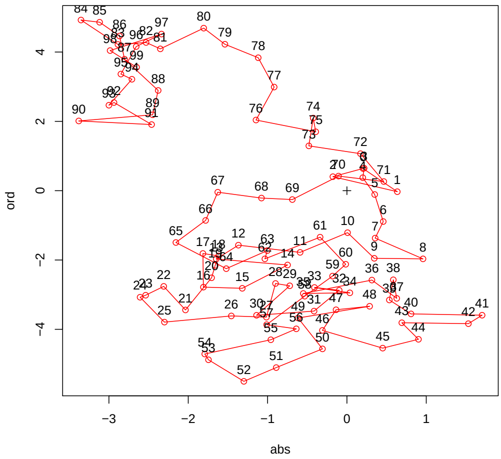
Elimination of the weights degeneracy problem.

Convergence :

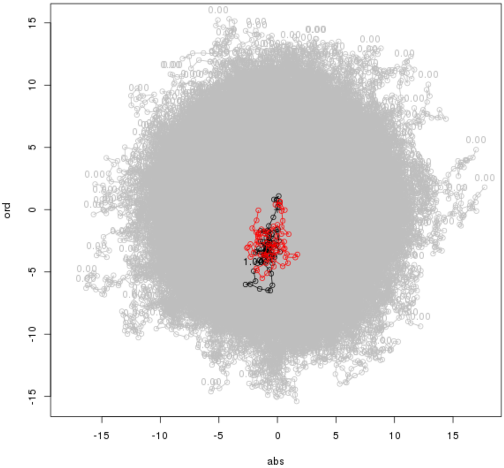
- for m the length of the chain fixed, central limit theorem when the number of particles $N \to +\infty$.

Example

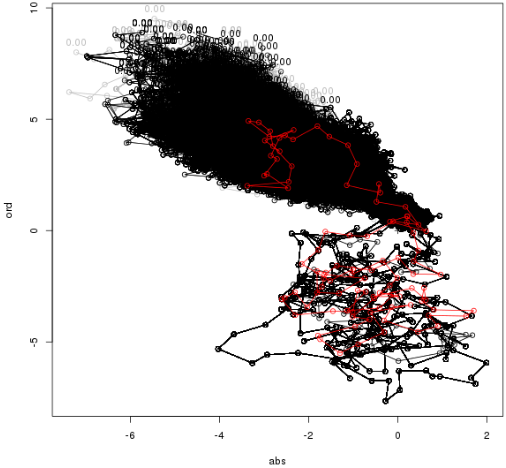
Comparison with $m = 100$ sites and $N = 10000$ particles.



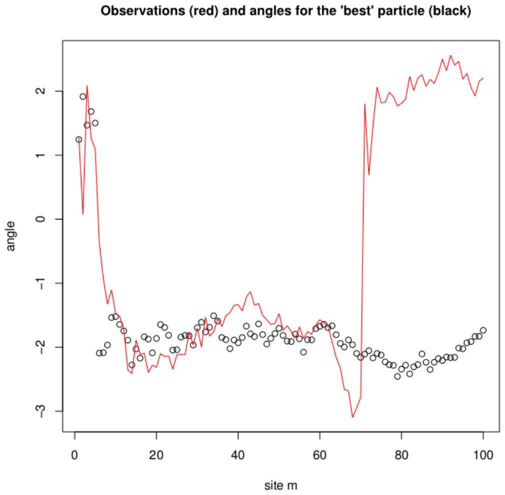
Example (SIS)



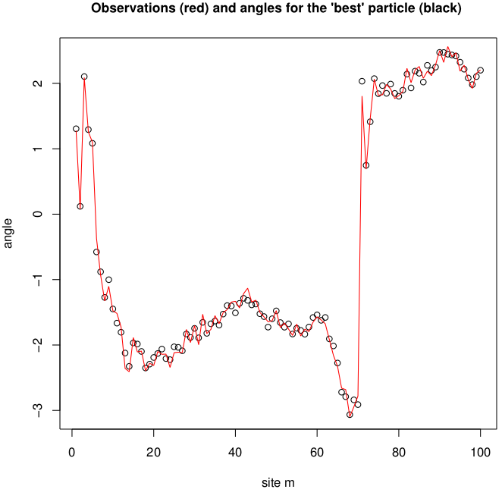
Example (SISR)



Example (SIS)


Observations (red) and angles for the 'best' particle (black)
angle
2
1
0
−1
−2
−3
0
20
40
60
80
100
site m

Example (SISR)



Outline

1. Definitions and issues
   - Markov chains
   - Hidden Markov models
   - Issues
2. Particle filtering algorithms
   - Sequential Importance Sampling algorithm
   - Sequential Importance Sampling Resampling algorithm
   - Comparison of algorithms
3. Real applications

Path tracking

- Observations : noisy position.
- Hidden states : real position.
- Parameters : behavior of the moving body.

Voice recognition system

- Observations : a word is pronounced, cut every 15ms.
- Hidden states : phonemes that led to this pronounced word.
- Parameters : the set of all dictionary words.

Phylogenetic analysis

- Observations : DNA sequences of several species at the leafs of a tree graph.
- Hidden states : all DNA sequences from the common ancestry sequence to the present time.
- Parameters : mutation parameters, lengths of the tree branches.

Phylogenetic analysis

$t = -1$	ACGAGGTGA
	ACAAGGTGA
	ACAGGGTGA
	ACAGGGCGA
	ACAGGGCAA
$t = 0$	ACAGGGCAA

Phylogenetic analysis

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

Phylogenetic analysis

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

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