Adaptive approximate Bayesian computation for complex models

Maxime Lenormand, Franck Jabot, Guillaume Deffuant

Adaptive approximate Bayesian computation for complex models

Maxime Lenormand, Franck Jabot & Guillaume Deffuant

Laboratory of Engineering for Complex Systems
Irstea of Clermont-Ferrand

8th World Congress in Probability and Statistics

July 12th 2012

Prototypical Policy Impacts on Multifunctional Activities in rural municipalities

A collaborative project under the EU Seventh Framework Programme

This publication has been funded under the PRIMA (Prototypical policy impacts on multifunctional activities in rural municipalities) collaborative project, EU 7th Framework Programme (ENV 2007-1), contract no. 212345.

Plan

1. Motivation
2. Approximate Bayesian Computation (ABC)
3. Approximate Bayesian Computation Sequential Monte Carlo
4. Adaptive Population Monte Carlo Approximate Bayesian Computation
5. Comparison of the algorithms
6. The *SimVillages* Model
7. Conclusion

Motivation

Complex social model

- Individual-based model
- Stochastic
- High dimensional parameter space
- High computational cost by simulation

Estimate the parameter values

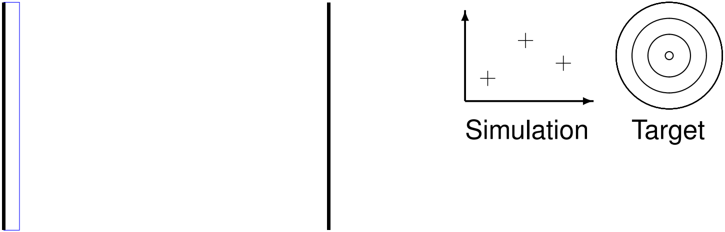
- Calibrate the model
- Understand the model behaviour
- Uncertainty analysis
- Validation

Plan

1. Motivation
2. Approximate Bayesian Computation (ABC)
3. Approximate Bayesian Computation Sequential Monte Carlo
4. Adaptive Population Monte Carlo Approximate Bayesian Computation
5. Comparison of the algorithms
6. The *SimVillages* Model
7. Conclusion

Approximate Bayesian Computation

1. Sample $\theta^* \sim \pi(\theta)$.
2. Simulate $x \sim f(x|\theta^*)$.
3. If $\rho(S(x), S(y)) \leq \epsilon$, accept θ^*, otherwise reject.
4. Repeat until a sample of the desired size is obtained



Prior distribution
$\pi(\theta)$

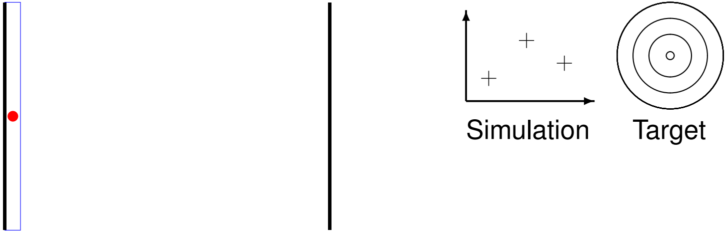
Posterior distribution
$\pi(\theta)P_\theta\{f(x|\theta) = y\}$

(Pritchard et al., 1999)

Derived from T. Toni 2011

Approximate Bayesian Computation

1. Sample $\theta^* \sim \pi(\theta)$.
2. Simulate $x \sim f(x|\theta^*)$.
3. If $\rho(S(x), S(y)) \leq \epsilon$, accept θ^*, otherwise reject.
4. Repeat until a sample of the desired size is obtained



Prior distribution
$\pi(\theta)$

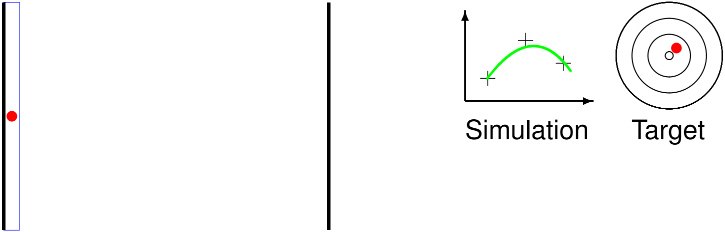
Posterior distribution
$\pi(\theta)P_\theta\{f(x|\theta) = y\}$

(Pritchard et al., 1999)

Derived from T. Toni 2011

Approximate Bayesian Computation

1. Sample $\theta^* \sim \pi(\theta)$.
2. Simulate $x \sim f(x|\theta^*)$.
3. If $\rho(S(x), S(y)) \leq \epsilon$, accept θ^*, otherwise reject.
4. Repeat until a sample of the desired size is obtained



Prior distribution
$\pi(\theta)$

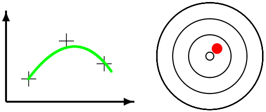
Posterior distribution
$\pi(\theta)P_\theta\{f(x|\theta) = y\}$

(Pritchard et al., 1999)

Derived from T. Toni 2011

Approximate Bayesian Computation

1. Sample $\theta^* \sim \pi(\theta)$.
2. Simulate $x \sim f(x|\theta^*)$.
3. If $\rho(S(x), S(y)) \leq \epsilon$, accept θ^*, otherwise reject.
4. Repeat until a sample of the desired size is obtained



Prior distribution
$\pi(\theta)$

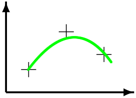
Posterior distribution
$\pi(\theta) P_\theta\{f(x|\theta) = y\}$

(Pritchard et al., 1999)

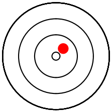
Derived from T. Toni 2011

Approximate Bayesian Computation

1. Sample $\theta^* \sim \pi(\theta)$.
2. Simulate $x \sim f(x|\theta^*)$.
3. If $\rho(S(x), S(y)) \leq \epsilon$, accept θ^*, otherwise reject.
4. Repeat until a sample of the desired size is obtained



Simulation



Target

Prior distribution
$\pi(\theta)$

Posterior distribution
$\pi(\theta)P_\theta\{f(x|\theta) = y\}$

(Pritchard et al., 1999)

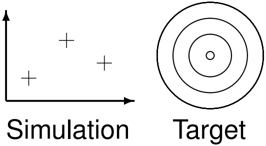
Derived from T. Toni 2011

Approximate Bayesian Computation

1. Sample $\theta^* \sim \pi(\theta)$.
2. Simulate $x \sim f(x|\theta^*)$.
3. If $\rho(S(x), S(y)) \leq \epsilon$, accept θ^*, otherwise reject.
4. Repeat until a sample of the desired size is obtained

Prior distribution
$\pi(\theta)$

Posterior distribution
$\pi(\theta) P_\theta\{f(x|\theta) = y\}$

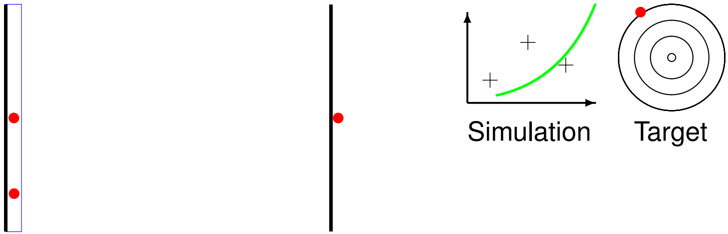


(Pritchard et al., 1999)

Derived from T. Toni 2011

Approximate Bayesian Computation

1. Sample $\theta^* \sim \pi(\theta)$.
2. Simulate $x \sim f(x|\theta^*)$.
3. If $\rho(S(x), S(y)) \leq \epsilon$, accept θ^*, otherwise reject.
4. Repeat until a sample of the desired size is obtained



Prior distribution
$\pi(\theta)$

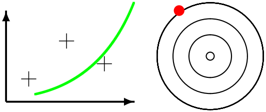
Posterior distribution
$\pi(\theta)P_\theta\{f(x|\theta) = y\}$

(Pritchard et al., 1999)

Derived from T. Toni 2011

Approximate Bayesian Computation

1. Sample $\theta^* \sim \pi(\theta)$.
2. Simulate $x \sim f(x|\theta^*)$.
3. If $\rho(S(x), S(y)) \leq \epsilon$, accept θ^*, otherwise reject.
4. Repeat until a sample of the desired size is obtained



Prior distribution
$\pi(\theta)$

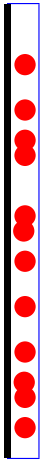
Posterior distribution
$\pi(\theta)P_\theta\{f(x|\theta) = y\}$

(Pritchard et al., 1999)

Derived from T. Toni 2011

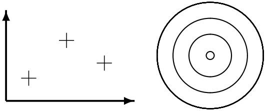
Approximate Bayesian Computation

1. Sample $\theta^* \sim \pi(\theta)$.
2. Simulate $x \sim f(x|\theta^*)$.
3. If $\rho(S(x), S(y)) \leq \epsilon$, accept θ^*, otherwise reject.
4. Repeat until a sample of the desired size is obtained



Prior distribution
$\pi(\theta)$

Posterior distribution
$\pi(\theta)P_\theta\{f(x|\theta) = y\}$



Simulation Target

(Pritchard et al., 1999)

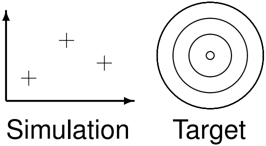
Derived from T. Toni 2011

Approximate Bayesian Computation

1. Sample $\theta^* \sim \pi(\theta)$.
2. Simulate $x \sim f(x|\theta^*)$.
3. If $\rho(S(x), S(y)) \leq \epsilon$, accept θ^*, otherwise reject.
4. Repeat until a sample of the desired size is obtained

Prior distribution
$\pi(\theta)$

Posterior distribution
$\pi(\theta) P_\theta\{f(x|\theta) = y\}$



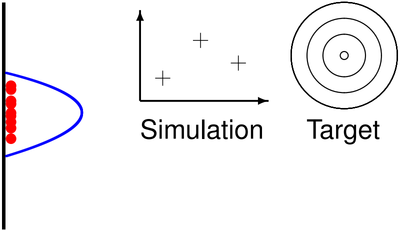
(Pritchard et al., 1999)

Derived from T. Toni 2011

Approximate Bayesian Computation

1. Sample $\theta^* \sim \pi(\theta)$.
2. Simulate $x \sim f(x|\theta^*)$.
3. If $\rho(S(x), S(y)) \leq \epsilon$, accept θ^*, otherwise reject.
4. Repeat until a sample of the desired size is obtained

Prior distribution
$\pi(\theta)$



Posterior distribution
$\pi(\theta) P_\theta\{f(x|\theta) = y\}$

(Pritchard et al., 1999)

Derived from T. Toni 2011

Plan

1. Motivation
2. Approximate Bayesian Computation (ABC)
3. Approximate Bayesian Computation Sequential Monte Carlo
4. Adaptive Population Monte Carlo Approximate Bayesian Computation
5. Comparison of the algorithms
6. The *SimVillages* Model
7. Conclusion

ABC SMC
Algorithm

Prior Posterior

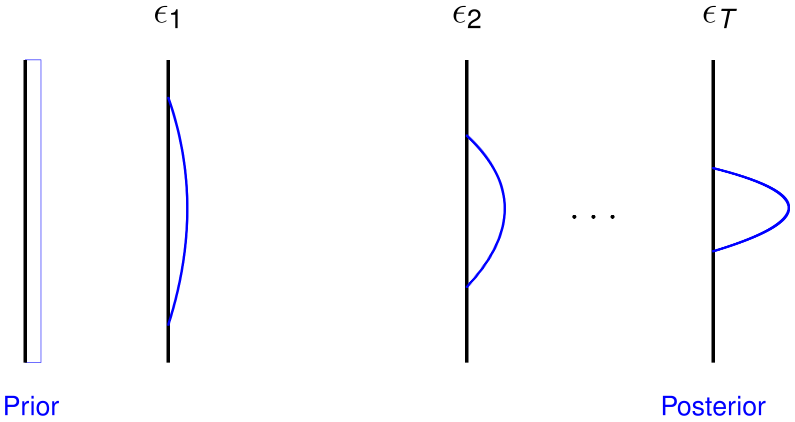
(Beaumont et al., 2009)
(Toni et al., 2009)

ABC SMC
Algorithm

(Beaumont et al., 2009)

(Toni et al., 2009)

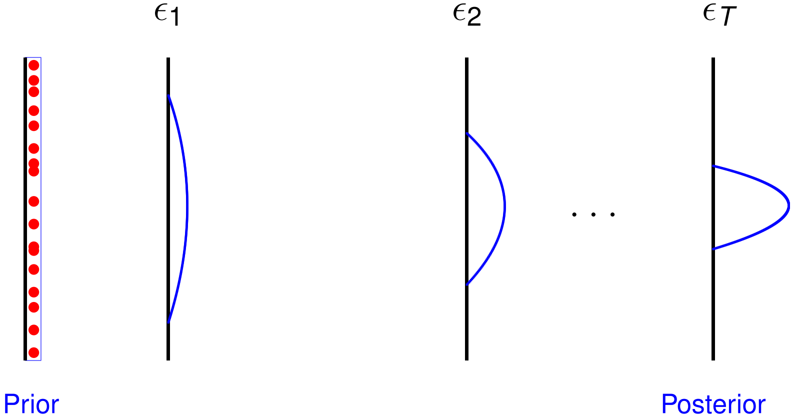
ABC SMC
Algorithm



(Beaumont et al., 2009)

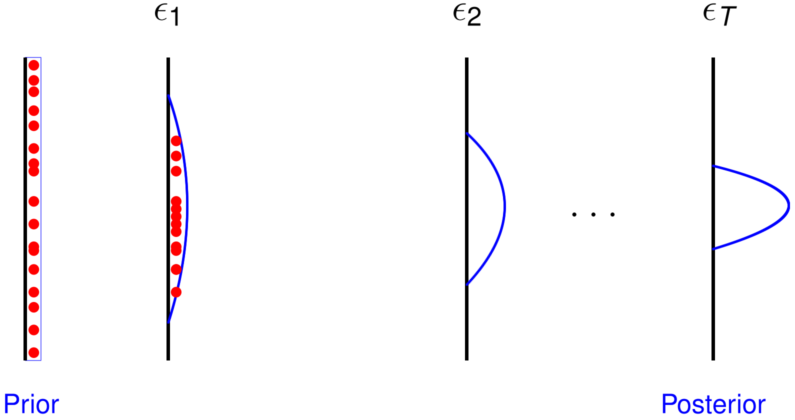
(Toni et al., 2009)

ABC SMC
Algorithm



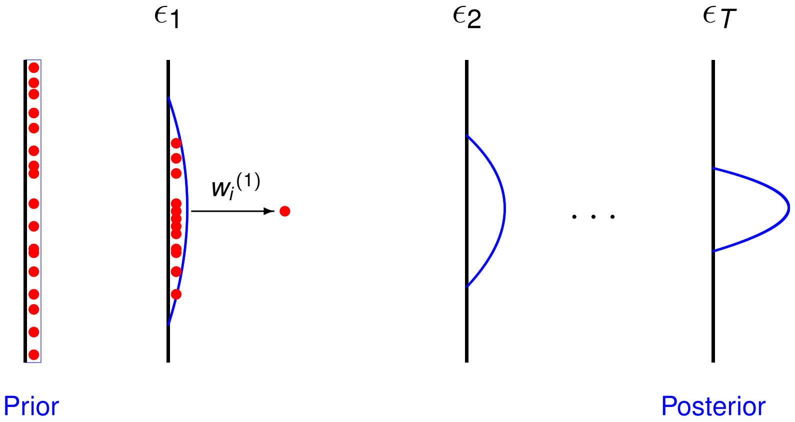
(Beaumont et al., 2009)
(Toni et al., 2009)

ABC SMC
Algorithm



(Beaumont et al., 2009)
(Toni et al., 2009)

ABC SMC
Algorithm

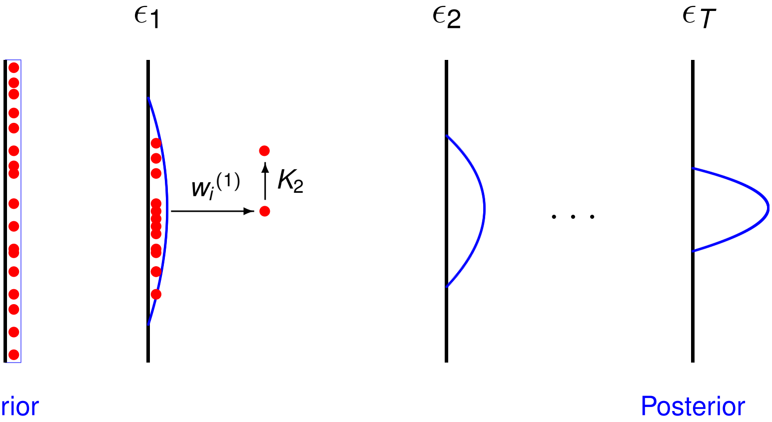


$$w_i^{(t)} = \frac{\pi(\theta_i^{(t)})}{\sum_{j=1}^{N} w_j^{(t-1)} \sigma_{t-1}^{-1} \varphi\left(\sigma_{t-1}^{-1}(\theta_i^{(t)} - \theta_j^{(t-1)})\right)}$$

(Beaumont et al., 2009)

(Toni et al., 2009)

ABC SMC
Algorithm

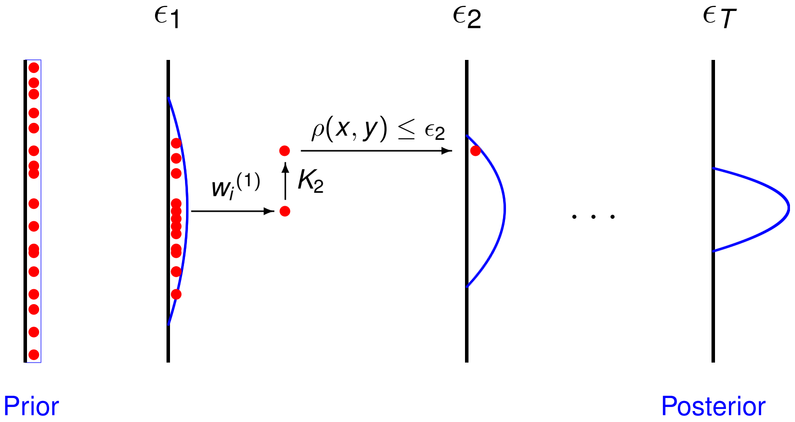


(Beaumont et al., 2009)
(Toni et al., 2009)

$$w_i^{(t)} = \frac{\pi(\theta_i^{(t)})}{\sum_{j=1}^{N} w_j^{(t-1)} \sigma_{t-1}^{-1} \varphi\left(\sigma_{t-1}^{-1}(\theta_i^{(t)} - \theta_j^{(t-1)})\right)}$$

$$K_t(\theta^* | \theta) = \sigma_{t-1}^{-1} \varphi\left(\sigma_{t-1}^{-1}(\theta^* - \theta)\right)$$

ABC SMC
Algorithm



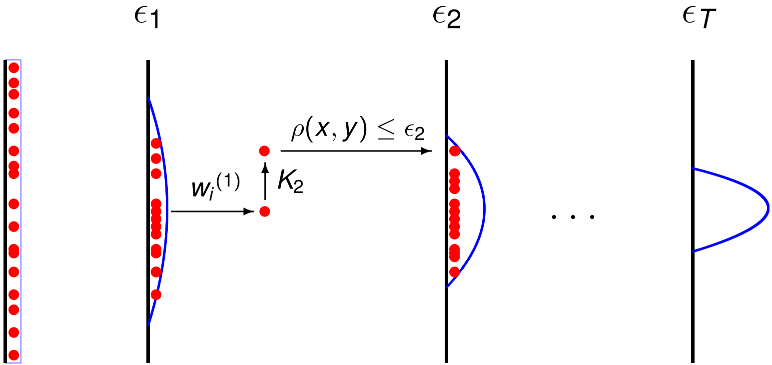
(Beaumont et al., 2009)

(Toni et al., 2009)

$$w_i^{(t)} = \frac{\pi(\theta_i^{(t)})}{\sum_{j=1}^{N} w_j^{(t-1)} \sigma_{t-1}^{-1} \varphi\left(\sigma_{t-1}^{-1}(\theta_i^{(t)} - \theta_j^{(t-1)})\right)}$$

$$K_t(\theta^* | \theta) = \sigma_{t-1}^{-1} \varphi\left(\sigma_{t-1}^{-1}(\theta^* - \theta)\right)$$

ABC SMC
Algorithm



Prior

Posterior

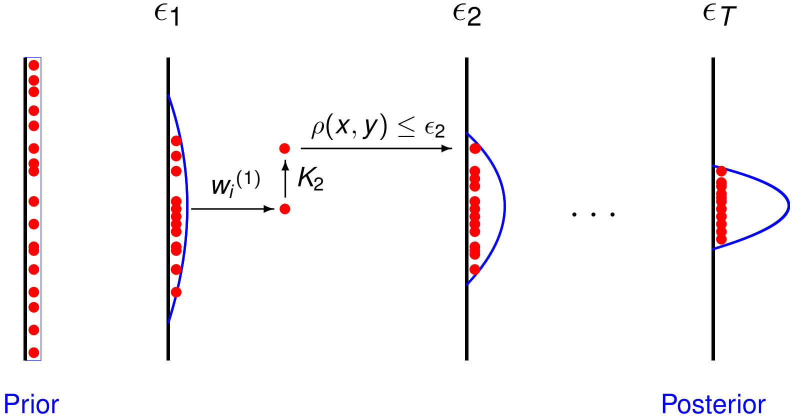
(Beaumont et al., 2009)

(Toni et al., 2009)

$$w_i^{(t)} = \frac{\pi(\theta_i^{(t)})}{\sum_{j=1}^{N} w_j^{(t-1)} \sigma_{t-1}^{-1} \varphi\left(\sigma_{t-1}^{-1}(\theta_i^{(t)} - \theta_j^{(t-1)})\right)}$$

$$K_t(\theta^* | \theta) = \sigma_{t-1}^{-1} \varphi\left(\sigma_{t-1}^{-1}(\theta^* - \theta)\right)$$

ABC SMC
Algorithm



(Beaumont et al., 2009)
(Toni et al., 2009)

$$w_i^{(t)} = \frac{\pi(\theta_i^{(t)})}{\sum_{j=1}^{N} w_j^{(t-1)} \sigma_{t-1}^{-1} \varphi\left(\sigma_{t-1}^{-1}(\theta_i^{(t)} - \theta_j^{(t-1)})\right)}$$

$$K_t(\theta^* | \theta) = \sigma_{t-1}^{-1} \varphi\left(\sigma_{t-1}^{-1}(\theta^* - \theta)\right)$$

ABC SMC
Recent Improvements

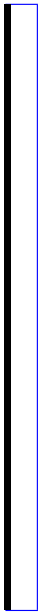
- Determination "on-line" of the decreasing sequence of tolerance levels $\{\epsilon_1, ..., \epsilon_T\}$. (DelMoral et al. (2011)) and (Drovandi and Pettitt (2011).
- From a quadratic complexity to a linear complexity for the computation of the weights thanks to a MCMC kernel. (DelMoral et al. (2011)) and (Drovandi and Pettitt (2011).
- Realisation of *M* pseudo-observations (DelMoral et al. (2011))

$\Longrightarrow$ The MCMC kernel leads to the problem of particles duplication!!!

Plan

1. Motivation
2. Approximate Bayesian Computation (ABC)
3. Approximate Bayesian Computation Sequential Monte Carlo
4. Adaptive Population Monte Carlo Approximate Bayesian Computation
5. Comparison of the algorithms
6. The *SimVillages* Model
7. Conclusion

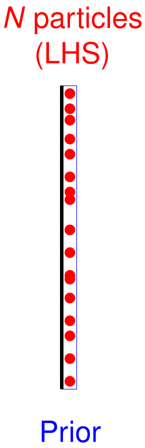
Adaptive PMC ABC
Algorithm



Prior

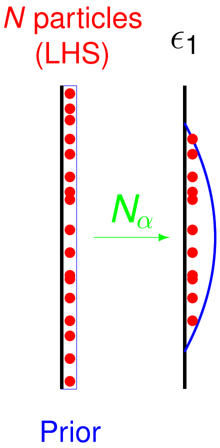
(Lenormand et al.)

Adaptive PMC ABC
Algorithm



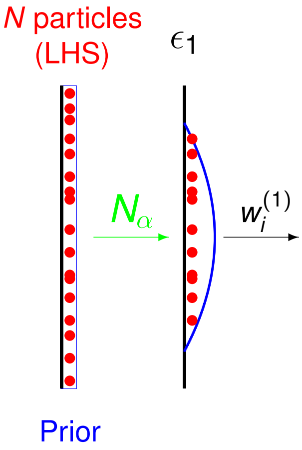
(Lenormand et al.)

Adaptive PMC ABC
Algorithm



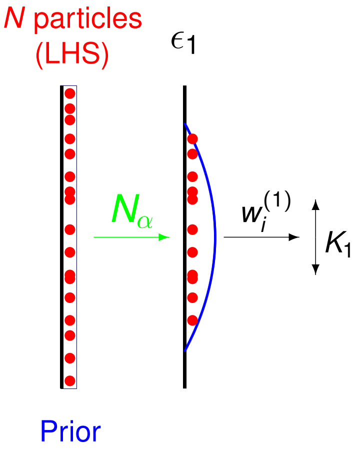
(Lenormand et al.)

Adaptive PMC ABC
Algorithm



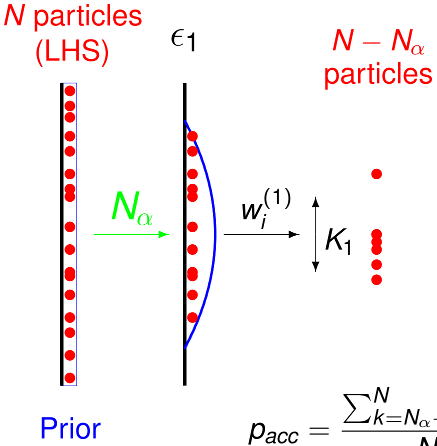
(Lenormand et al.)

Adaptive PMC ABC
Algorithm



(Lenormand et al.)

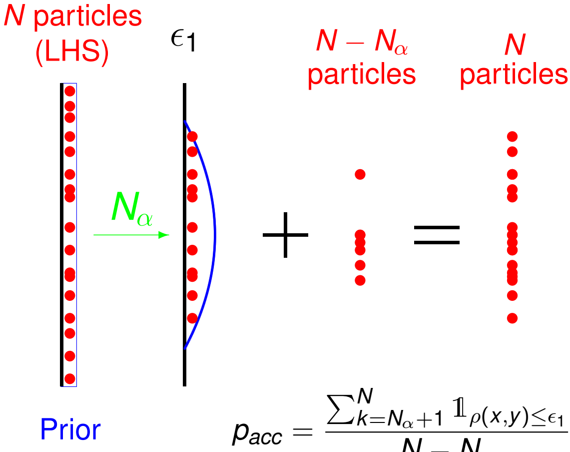
Adaptive PMC ABC
Algorithm



$$p_{acc} = \frac{\sum_{k=N_\alpha+1}^{N} \mathbb{1}_{\rho(x,y)\leq\epsilon_1}}{N - N_\alpha}$$

(Lenormand et al.)

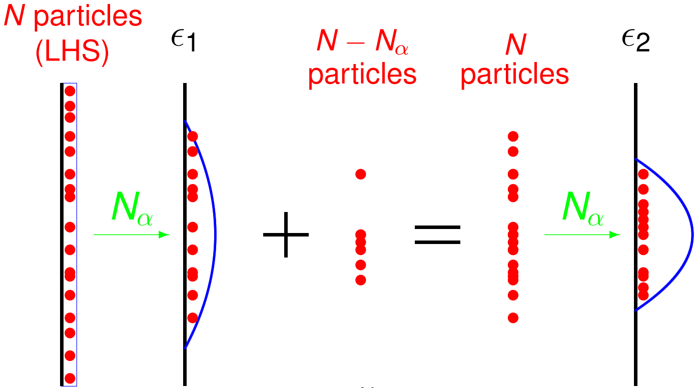
Adaptive PMC ABC
Algorithm



$$p_{acc} = \frac{\sum_{k=N_\alpha+1}^{N} \mathbb{1}_{\rho(x,y)\leq\epsilon_1}}{N - N_\alpha}$$

(Lenormand et al.)

Adaptive PMC ABC
Algorithm

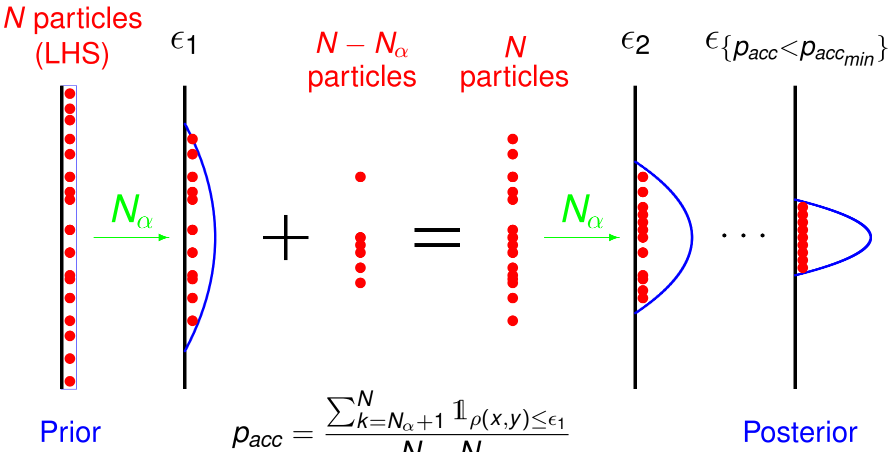


Prior

$$p_{acc} = \frac{\sum_{k=N_\alpha+1}^{N} \mathbb{1}_{\rho(x,y)\leq\epsilon_1}}{N - N_\alpha}$$

(Lenormand et al.)

Adaptive PMC ABC
Algorithm



$$p_{acc} = \frac{\sum_{k=N_\alpha+1}^{N} \mathbb{1}_{\rho(x,y) \leq \epsilon_1}}{N - N_\alpha}$$

(Lenormand et al.)

Adaptive PMC ABC
Pros and Cons

Pros

- Control the number of simulations at each iteration ($N - N_\alpha$ simulations).

Adaptive PMC ABC
Pros and Cons

Pros

- Control the number of simulations at each iteration ($N - N_\alpha$ simulations).
- Determination "on-line" of the decreasing sequence of tolerance levels.

Adaptive PMC ABC
Pros and Cons

Pros

- Control the number of simulations at each iteration ($N - N_\alpha$ simulations).
- Determination "on-line" of the decreasing sequence of tolerance levels.
- No problem of particles duplication.

Adaptive PMC ABC
Pros and Cons

Pros

- Control the number of simulations at each iteration ($N - N_\alpha$ simulations).
- Determination "on-line" of the decreasing sequence of tolerance levels.
- No problem of particles duplication.
- Stopping criterion

Adaptive PMC ABC
Pros and Cons

Pros

- Control the number of simulations at each iteration ($N - N_\alpha$ simulations).
- Determination "on-line" of the decreasing sequence of tolerance levels.
- No problem of particles duplication.
- Stopping criterion

Cons

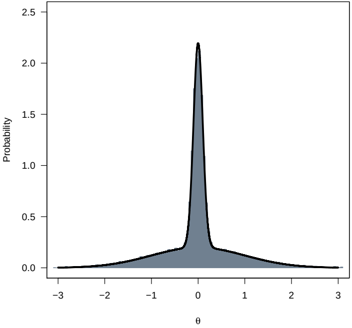
- Complexity $O({N_\alpha}^2)$ for the computation of the weights

Plan

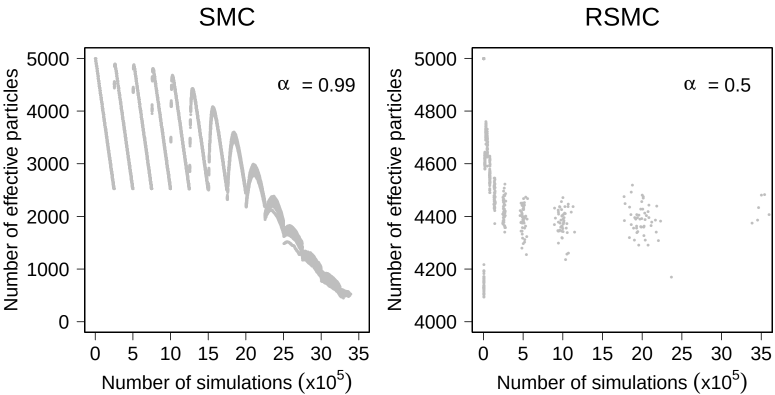
1. Motivation
2. Approximate Bayesian Computation (ABC)
3. Approximate Bayesian Computation Sequential Monte Carlo
4. Adaptive Population Monte Carlo Approximate Bayesian Computation
5. Comparison of the algorithms
6. The *SimVillages* Model
7. Conclusion

Comparison of the algorithms
Toy Example : Presentation

$$f(x|\theta) \sim \frac{1}{2}\phi\left(\theta, \frac{1}{100}\right) + \frac{1}{2}\phi(\theta, 1) \text{ and } \theta \sim \mathcal{U}_{[-10,10]}$$

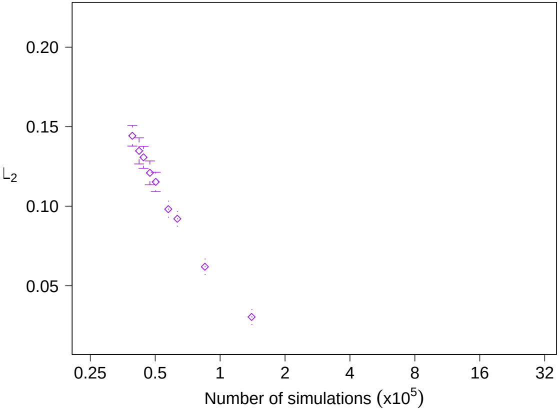


Comparison of the algorithms
Toy Example : Particles Duplication



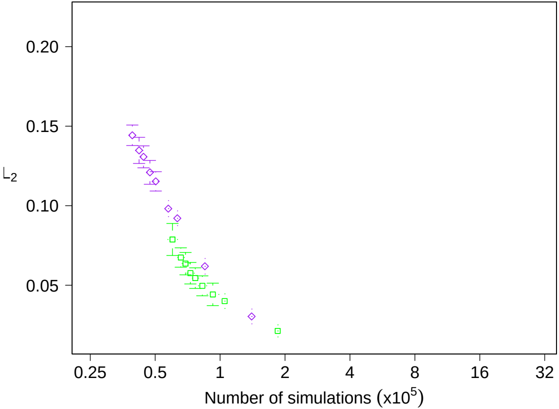
Comparison of the algorithms
Toy Example : Comparison

$N_\alpha = 5000$; α from 0.9 to 0.1 corresponding to $N = 5555$ to 50000



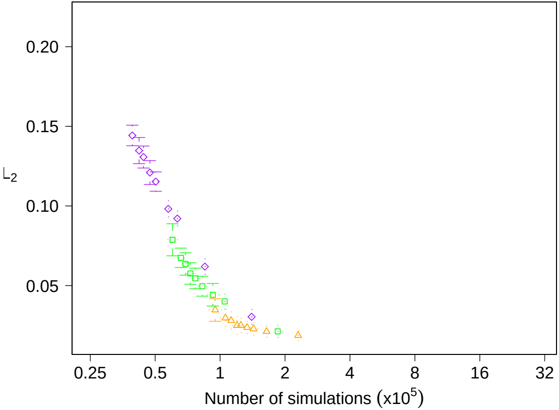
Comparison of the algorithms
Toy Example : Comparison

$N_\alpha = 5000$; α from 0.9 to 0.1 corresponding to $N = 5555$ to 50000



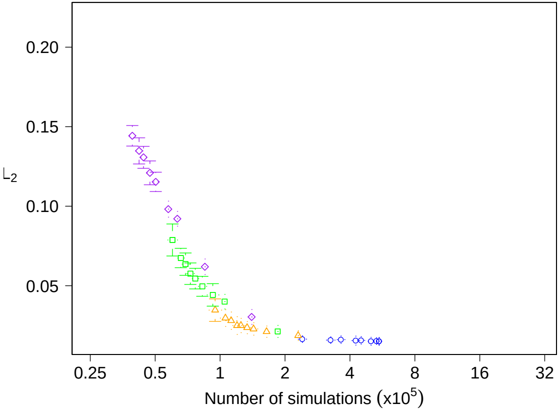
Comparison of the algorithms
Toy Example : Comparison

$N_\alpha = 5000$; α from 0.9 to 0.1 corresponding to $N = 5555$ to 50000



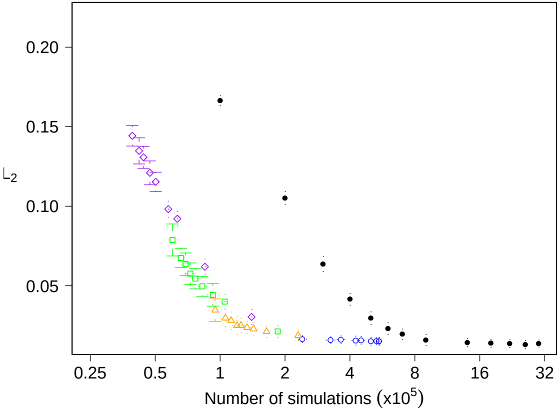
Comparison of the algorithms
Toy Example : Comparison

$N_\alpha = 5000$; α from 0.9 to 0.1 corresponding to $N = 5555$ to 50000



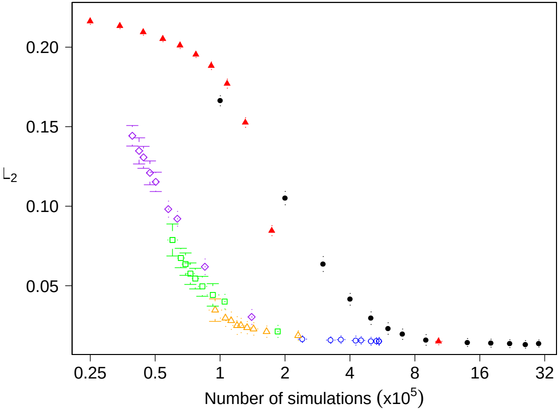
Comparison of the algorithms
Toy Example : Comparison

$N_\alpha = 5000$; α from 0.9 to 0.1 corresponding to $N = 5555$ to 50000



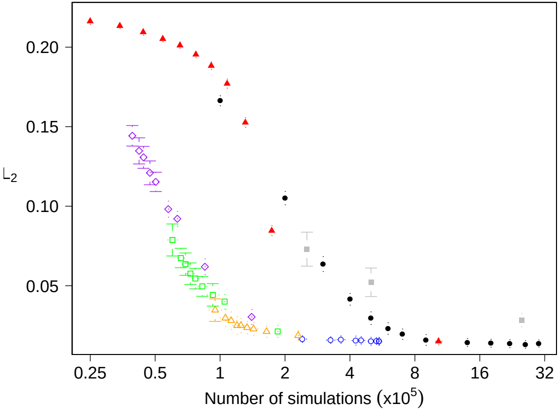
Comparison of the algorithms
Toy Example : Comparison

$N_\alpha = 5000$; α from 0.9 to 0.1 corresponding to $N = 5555$ to 50000



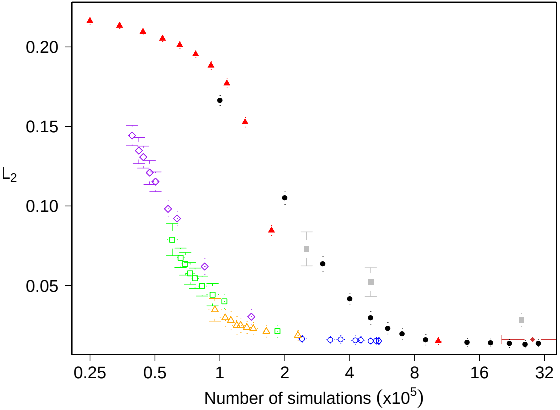
Comparison of the algorithms
Toy Example : Comparison

$N_\alpha = 5000$; α from 0.9 to 0.1 corresponding to $N = 5555$ to 50000



Comparison of the algorithms
Toy Example : Comparison

$N_\alpha = 5000$; α from 0.9 to 0.1 corresponding to $N = 5555$ to 50000



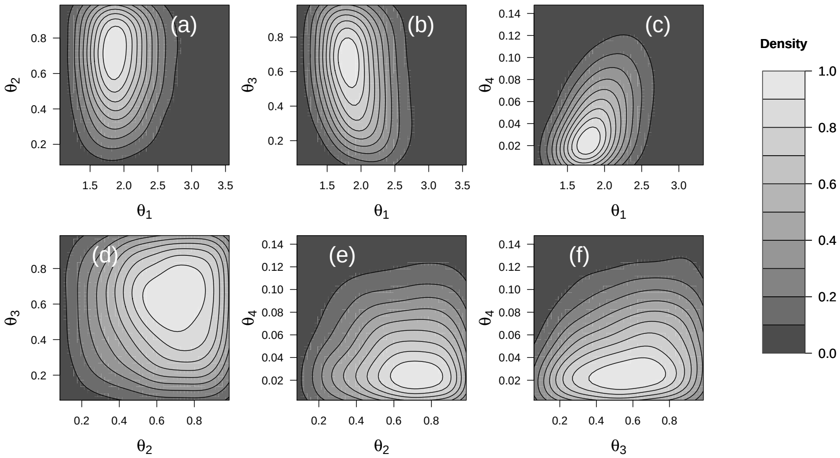
Plan

1. Motivation
2. Approximate Bayesian Computation (ABC)
3. Approximate Bayesian Computation Sequential Monte Carlo
4. Adaptive Population Monte Carlo Approximate Bayesian Computation
5. Comparison of the algorithms
6. The *SimVillages* Model
7. Conclusion

The *SimVillages* Model
Parameter and Summary Statistics

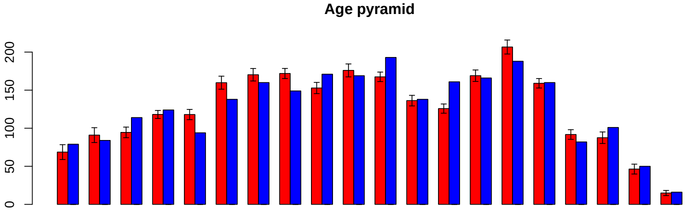
- 4 parameters
- 8 statistics
- $\|(\rho_m(S_m, S_m^{'}))_{1 \leq m \leq M}\|_\infty = \sup_{1 \leq m \leq M} |\rho_m(S_m, S_m^{'})|$

The *SimVillages* Model
Posterior Density

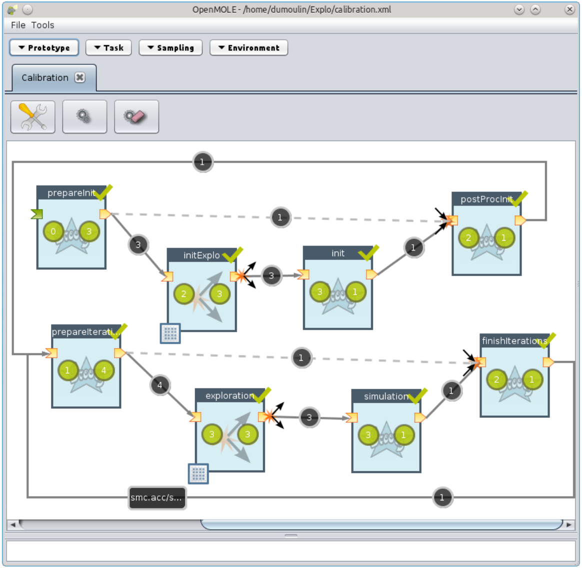


The *SimVillages* Model
Concrete Results

- 1.4 second by simulation
- 400 000 simulations
- 6 days



The *SimVillages* Model
Open Mole



Plan

1. Motivation
2. Approximate Bayesian Computation (ABC)
3. Approximate Bayesian Computation Sequential Monte Carlo
4. Adaptive Population Monte Carlo Approximate Bayesian Computation
5. Comparison of the algorithms
6. The *SimVillages* Model
7. Conclusion

Conclusion

- We have developed a new algorithm to reduce the number of simulations in SMC ABC.
- Comparison with three methods.
- Calibration of a complex social model.

Bibliography

Beaumont, M. A., Cornuet, J., Marin, J., Robert, C. P.:
Adaptive approximate Bayesian computation.
Biometrika. 96(4),983–990 (2009)

Toni, T., Welch, D., Strelkowa, N., Ipsen, A., Stumpf, M. P. H.:
Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems
JOURNAL OF THE ROYAL SOCIETY INTERFACE. 6,187 (2009)

Drovandi, C. C., Pettitt, A. N.:
Estimation of parameters for macroparasite population evolution using approximate Bayesian computation.
Biometrics. 67(1),225–233 (2011)

Del Moral, P., Doucet, A., Jasra, A.:
An adaptive sequential Monte Carlo method for approximate Bayesian computation.
Statistics and Computing (In press) (2012)

Lenormand, M., Jabot, F., Deffuant, G.:
Adaptive approximate Bayesian computation for complex models.
arXiv:1111.1308v2. (Submit) (2012)