



Adaptive approximate Bayesian computation for complex models

Maxime Lenormand, Jabot Franck, Guillaume Deffuant

► To cite this version:

Maxime Lenormand, Jabot Franck, Guillaume Deffuant. Adaptive approximate Bayesian computation for complex models. The 11th World Meeting of The International Society for Bayesian Analysis, 2012, Kyoto, Japan. hal-02889634

HAL Id: hal-02889634

<https://hal.inrae.fr/hal-02889634>

Submitted on 4 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ADAPTIVE APPROXIMATE BAYESIAN COMPUTATION FOR COMPLEX MODELS

MAXIME LENORMAND, FRANCK JABOT & GUILLAUME DEFFUANT
LISC, IRSTEA, Clermont-Ferrand, FRANCE (maxime.lenormand@irstea.fr)



CONTRIBUTION

- **What?** Improve the sequential Monte Carlo methods in ABC to minimize the number of model runs to reach a given quality of the posterior approximation.
- **Why?** In the PMC algorithm proposed by Beaumont et al. (2009) the sequence of tolerance levels is not automatically determined. The recent SMC algorithm proposed by Del Moral et al. (2012) and the RSMC proposed by Drovandi and Pettitt (2011) use a MCMC kernel which leads to a problem of particle duplication.
- **How?** We propose an Adaptive Population Monte Carlo algorithm where the sequence of tolerance levels is determined automatically and which makes use of an easily interpretable stopping criterion.

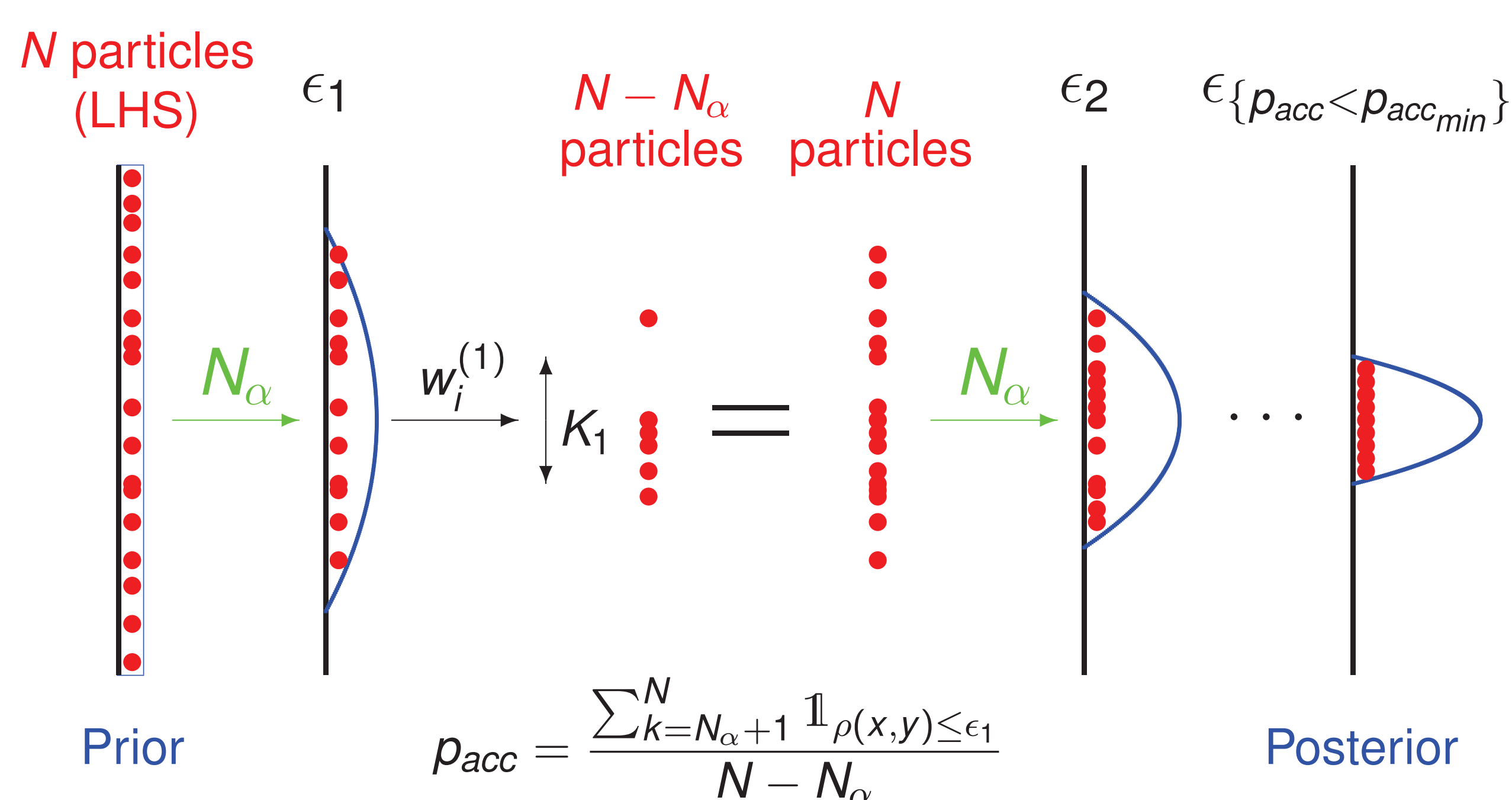
ABC ALGORITHM

1. Simulate a particle θ from the prior distribution of the model parameters and use it to simulate the model realisation $x \sim f(x|\theta)$.
2. If the distance $\rho(x, y)$ to the data y is below a tolerance level ϵ , accept the particle otherwise reject it.
3. Repeat until a particle sample of the desired size is obtained.

APMC ALGORITHM

The Adaptive Population Monte Carlo algorithm is a sequential Monte Carlo method in ABC. The purpose of this method is to sequentially improve the parameter posterior distributions represented by a sample of particles of length N_α .

1. Simulate N particles from the prior distribution π using a *Latin Hypercube Sampling*, simulate the N associated realizations of the model and compute the N associated distances to the data.
2. Keep the $N_\alpha = \lfloor N\alpha \rfloor$ closest to data particles and update their associated tolerance level (the maximum of the N_α distances).
3. Move $N - N_\alpha$ particles from the previous weighted sample with a jumping kernel and simulate the $N - N_\alpha$ associated realizations of the model.
4. Merge the two samples to obtain a new weighted sample of length N .
5. Go to step 2 until the proportion of "accepted" particles p_{acc} among the $N - N_\alpha$ particles is below the stopping criterion $p_{acc_{min}}$.



$W_i^{(t)} = \pi(\theta_i^{(t)})/d_i^{(t)}$ where π is the prior distribution and $d_i^{(t)}$ is given by the sum of the probabilities to reach $\theta_i^{(t)}$ from one of the N_α particles of the previous step.

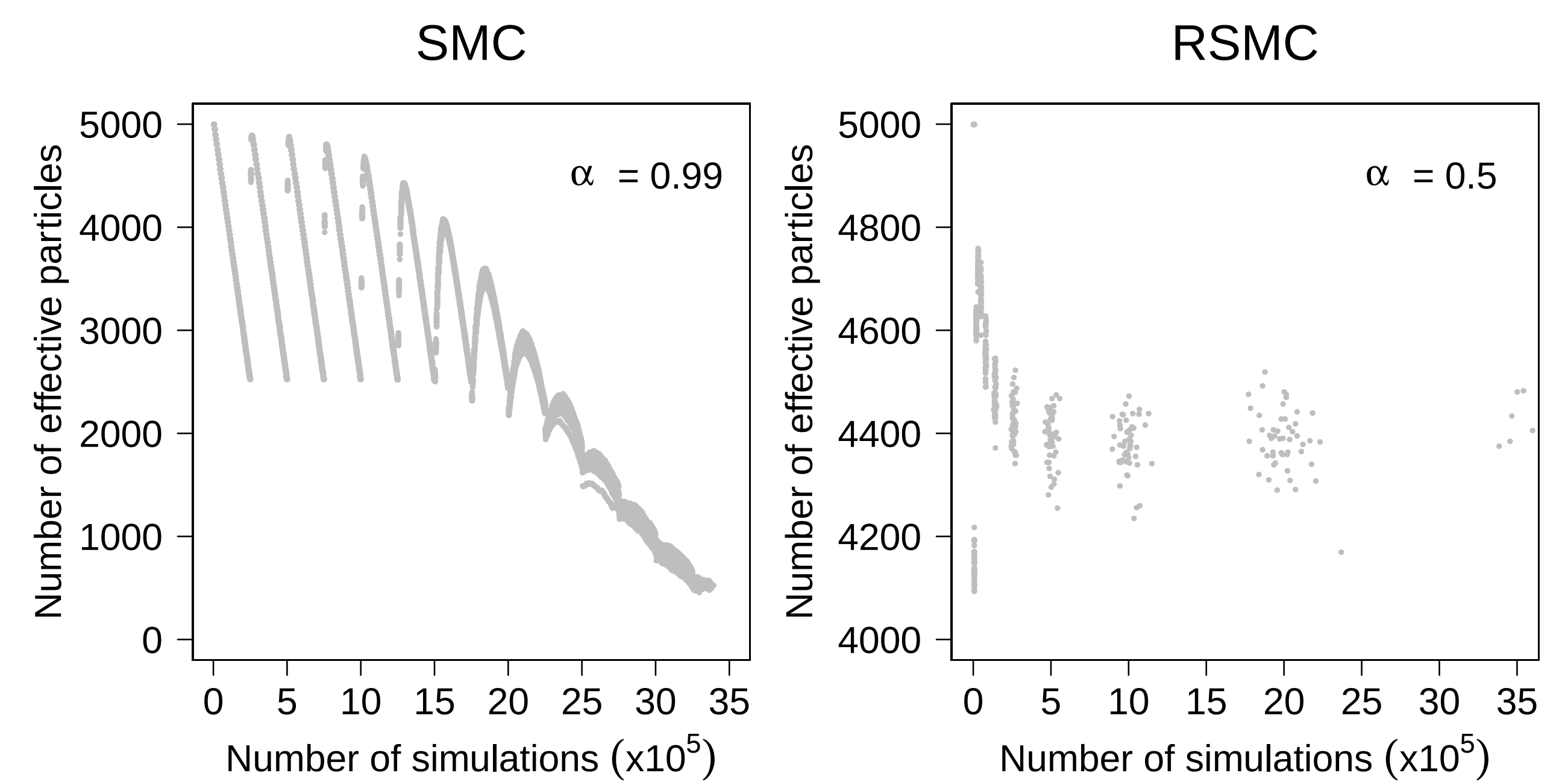
The kernel transition K_t is a multivariate normal distribution parameterized with twice the weighted variance of the previous sample.

EXPERIMENTS ON A TOY EXAMPLE

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

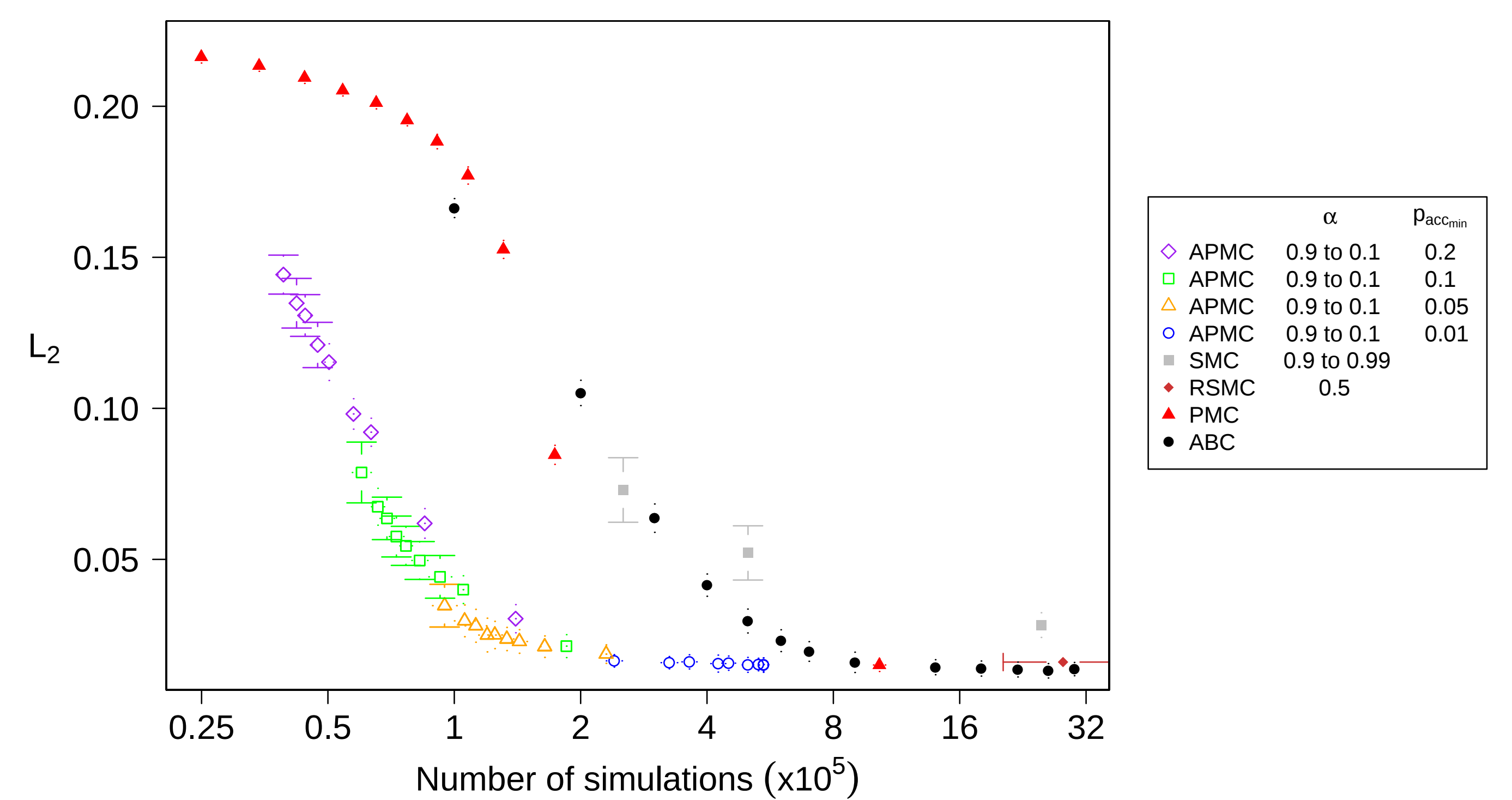
$\phi(\mu, \sigma^2)$ is the normal density of mean μ and variance σ^2

The problem of particle duplication

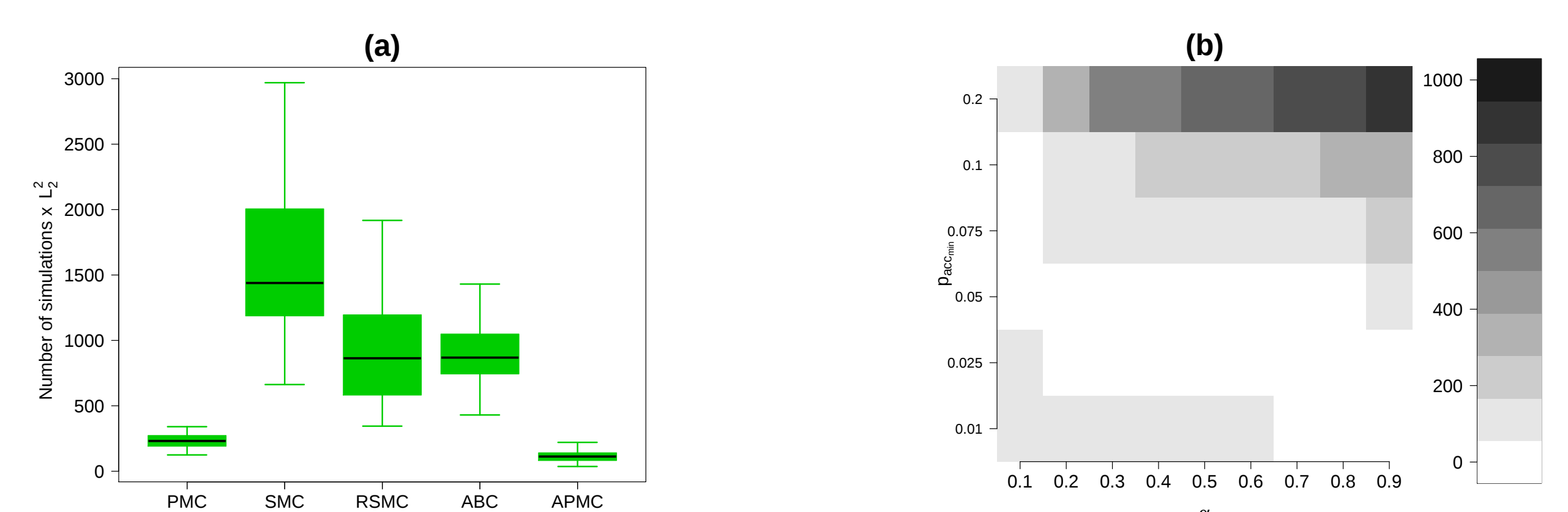


⇒ The number of distinct particles decreases during the course of the algorithms.

Comparison of algorithms



⇒ The APMC algorithm is 2 to 8 times faster than the other algorithms for an equivalent $\mathbb{L}_2^2$ distance between the simulated and the observed posterior distribution of model parameters.



⇒ The number of simulations $\times \mathbb{L}_2^2$ is smaller for the APMC. $0.3 \leq \alpha \leq 0.7$ and $0.01 \leq p_{acc_{min}} \leq 0.05$ present a good compromise

REFERENCES

- Beaumont, M. A., Cornuet, J., Marin, J., Robert, C. P.: Adaptive approximate Bayesian computation. *Biometrika*. 96(4),983–990 (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. *Stat. Comput.* (In press) (2012)

FUNDING

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