TAMIS : Tempered Anti-truncated Multiple Importance Sampling

Grégoire Aufort, Pierre Pudlo

gregoire.aufort@univ-amu.fr

Aix-Marseille Université

Abstract

We propose a Tempered Anti-truncated Adaptive Multiple Importance Sampling (TAMIS) algorithm to solve the initialization difficulty of the AIS algorithms, without introducing too many evaluations of the target density. Combining a tempering scheme and a new nonlinear transformation of the weights. As a result, our proposal is an automatically tuned sequential algorithm that is robust to many initial proposals, doesn't require gradient computations and scales well with the dimension.

Anti-Truncation

Let π be the target distribution, q the proposal and $w_i = \dfrac{\pi(x_i)}{q(x_i)}$ (1) the importance weights

Then the modified weights

$$\widehat{w}_i = \max(s; w_i)$$

Can be used to approximate the distribution

$$\widehat{\pi}(x) \propto sq(x)\mathbf{1}\{x \in E\} + \pi(x)\mathbf{1}\{x \notin E\}, \quad \text{where } E = \{x : \pi(x)/q(x) \le s\}.$$

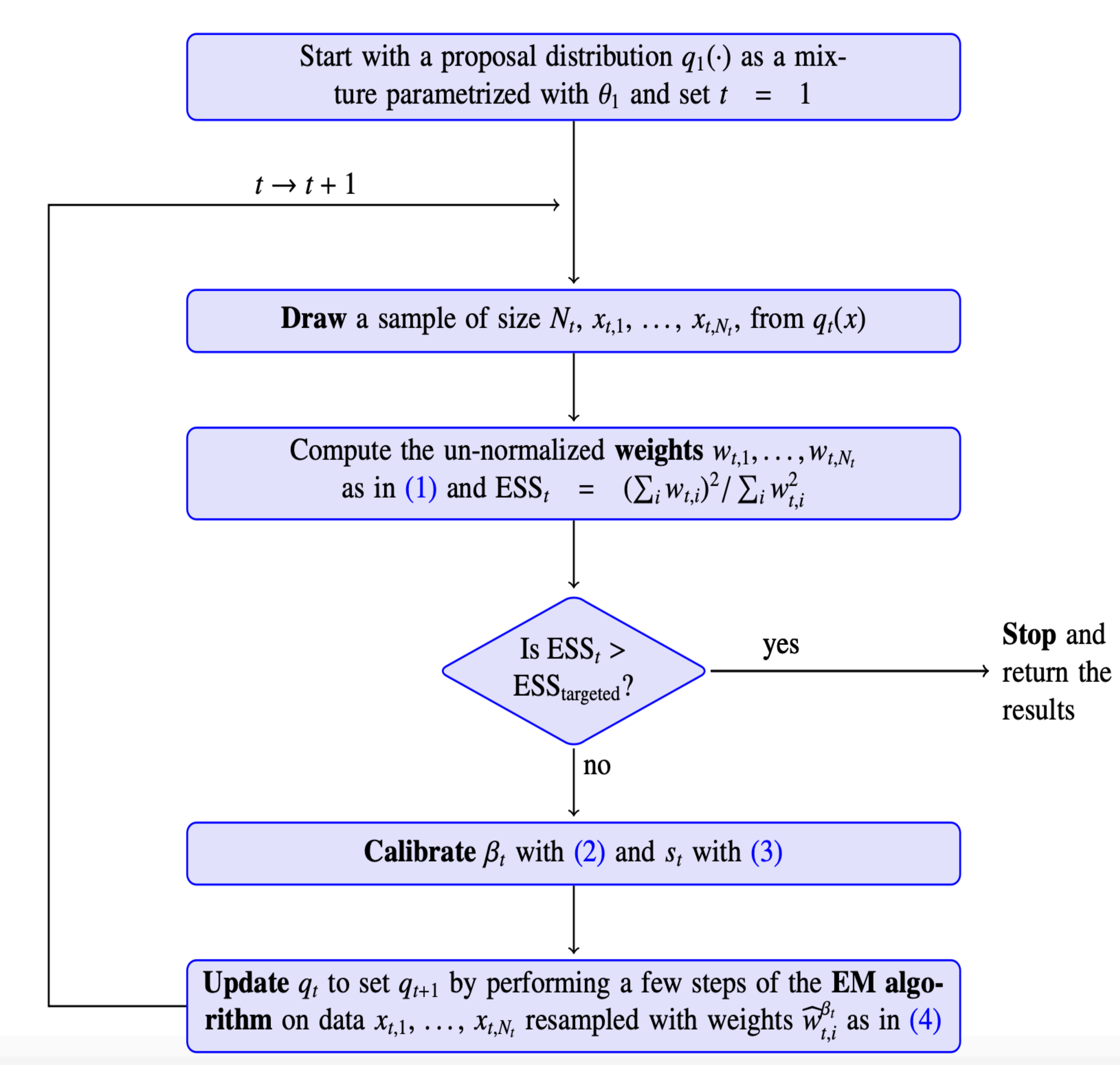
And, if $s \le 1$ we have :

$$0 = \mathrm{KL}(\pi \| \pi) \le \mathrm{KL}(\pi \| \widehat{\pi}) \le \mathrm{KL}(\pi \| q)$$

TAMIS

Idea :

Combine auto-tuned tempering [1] and anti-truncation to create a sequence of auxiliary distribution bridging from the proposal to the target



$$\frac{\pi_{\beta,t}(x_{t,i})}{q_t(x_{t,i})} \propto \frac{\pi(x)^\beta q_t(x)^{1-\beta}}{q_t(x_{t,i})} = \left(\frac{\pi(x)}{q_t(x_{t,i})}\right)^\beta = w_{t,i}^\beta.$$

$$\beta_t = \sup\{\beta \in (0,1) : \mathrm{ESS}(\beta) > \alpha\}$$

$$\text{where } \mathrm{ESS}(\beta) = \left(\sum_{i=1}^{N_t} w_{t,i}^\beta\right)^2 \Big/ \sum_{i=1}^{N_t} w_{t,i}^{2\beta} \qquad (2)$$

$$s_t = \text{quantile}_{\text{order}=\tau}\left(w_{t,1}^\beta, \ldots, w_{t,N_t}^\beta\right). \qquad (3)$$

$$\widehat{w}_{t,i}^\beta = \min(s_t; w_{t,i}^\beta) \qquad (4)$$

Numerical Results

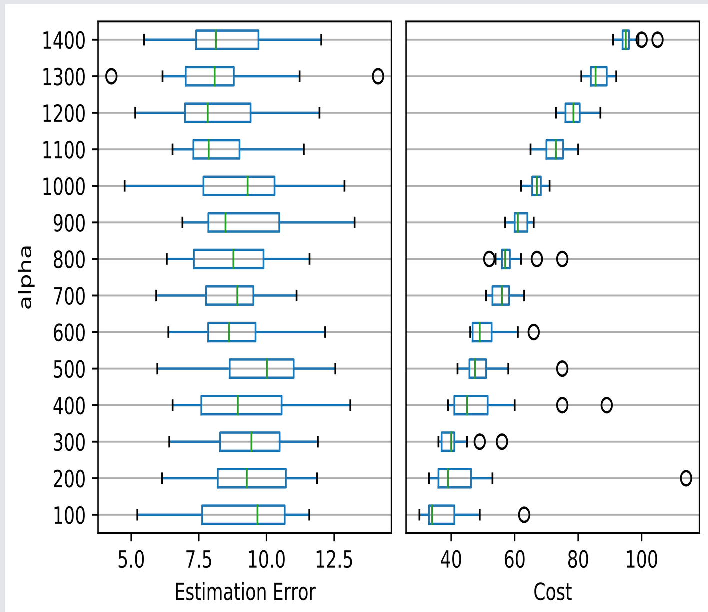
Target : d-dimensional gaussian $\mathcal{N}(50,5)^{\otimes d}$

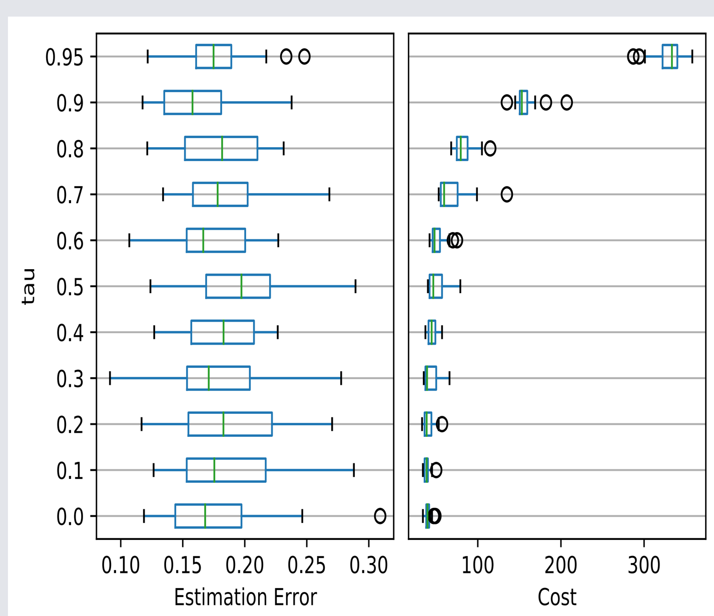
Proposal : 5-Gaussian Mixture initialized with

$$\Sigma_{0,k} = 200 \times I_{50}, \quad \mu_{0,k} \sim \mathcal{U}([-4,4]) \text{ for } k = 1,\ldots,5$$

On parameter tuning

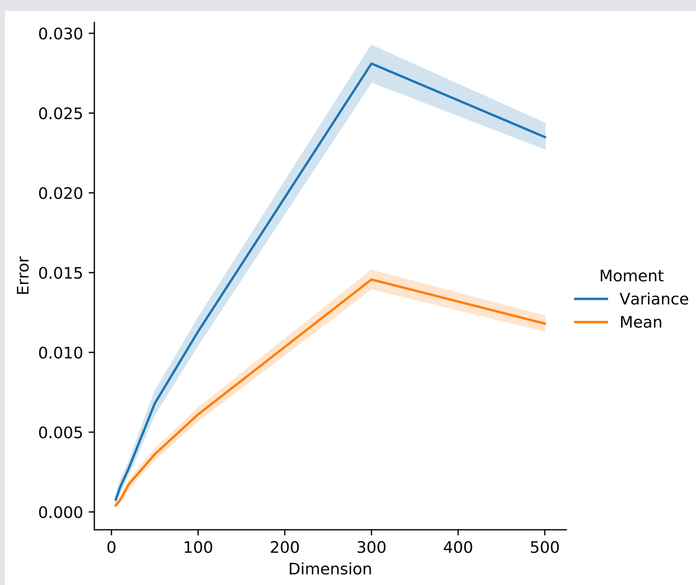
either α or τ is fixed

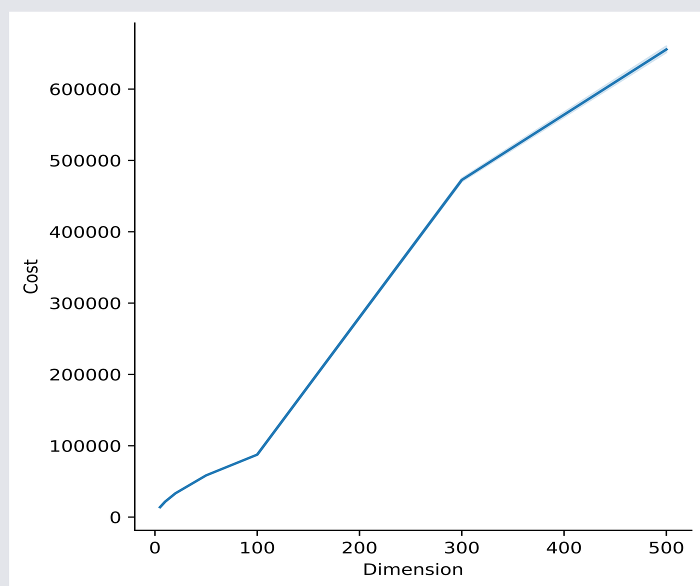




On dimensionality

$\tau = 0.4, \quad \alpha = 300$ or 1000





Target : Rosenbrock distribution

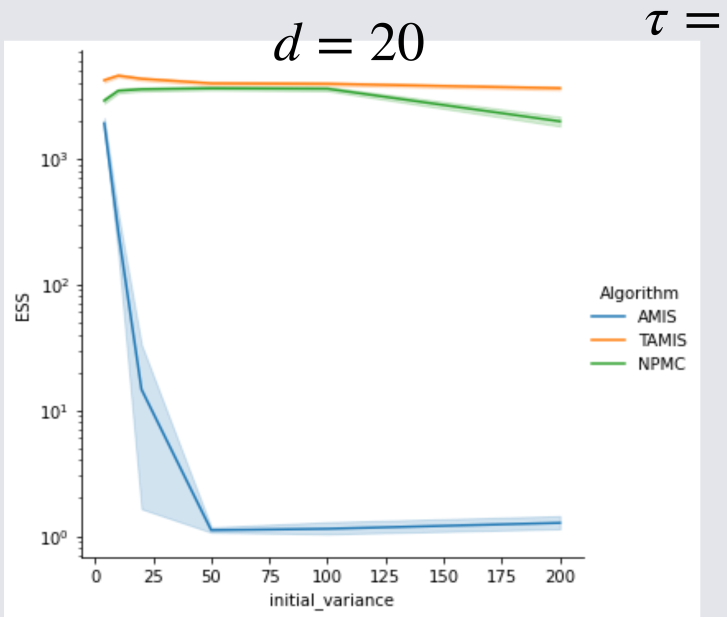
Proposal :5-Gaussian Mixture initialized with

$$\Sigma_{0,k} = \sigma^2 \times I_d, \text{ for } k = 1,\ldots,5$$

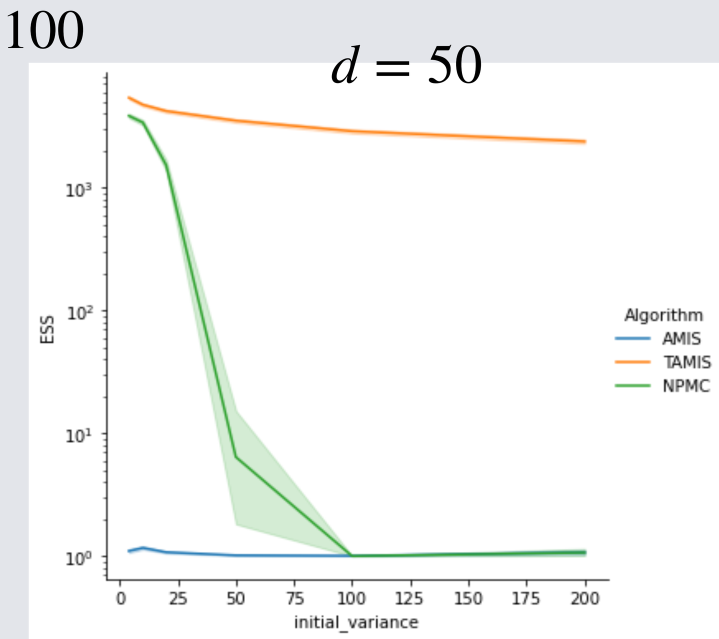
$\tau = 0.4, \alpha = 100$

On the initialization

$d = 20$

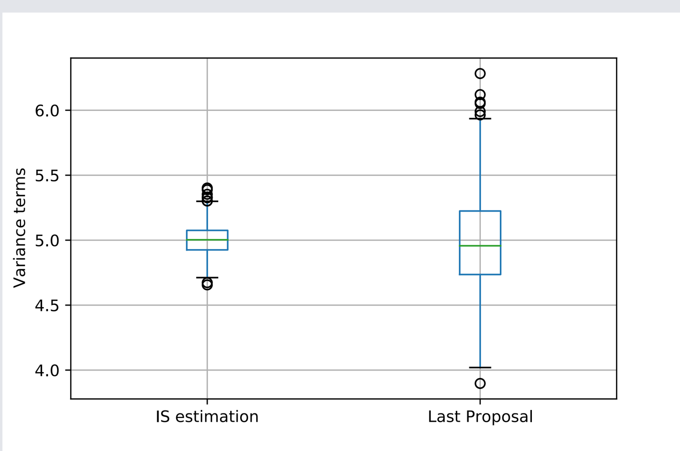


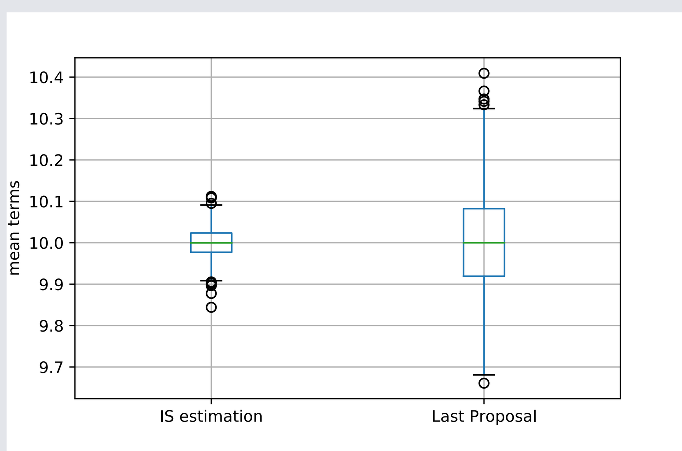
$d = 50$

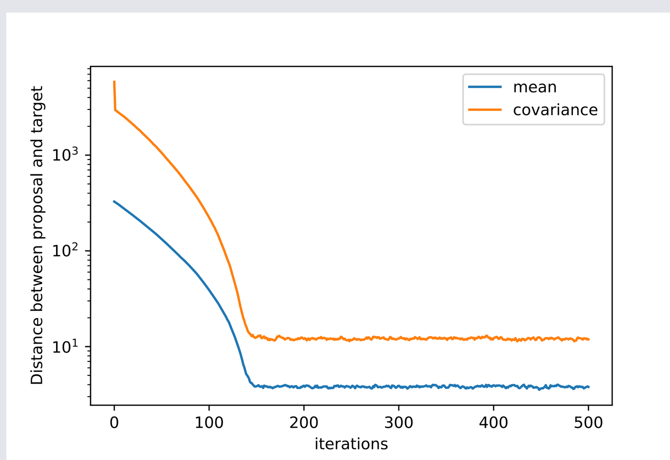


Target : 1000-dimensional Gaussian $\mathcal{N}(10,5)^{\otimes 1000}$

Proposal : Gaussian initialized with $\Sigma_0 = 100 \times I_{1000}, \quad \mu_0 \sim \mathcal{U}([-4,4])$







References

1. Beskos, A., Jasra, A., Kantas, N., and Thiery, A. (2016). On the convergence of adaptive sequentialMonte Carlo methods. The Annals of Applied Probability, 26(2):1111 – 1146
2. Koblents, E., and Mıguez, J. (2015). A population monte carlo scheme with transformed weights and its application to stochastic kinetic models.Statistics and Computing 25(2):407– 425