Adaptive Rejection Sampling with fixed number of nodes

L. Martino⋆, F. Louzada⋆

⋆ *Institute of Mathematical Sciences and Computing, Universidade de São Paulo, São Carlos (São Paulo).*

Abstract

The adaptive rejection sampling (ARS) algorithm is a universal random generator for drawing samples efficiently from a univariate log-concave target probability density function (pdf). ARS generates independent samples from the target via rejection sampling with high acceptance rates. Indeed, ARS yields a sequence of proposal functions that converge toward the target pdf, so that the probability of accepting a sample approaches one. However, sampling from the proposal pdf becomes more computational demanding each time it is updated. In this work, we propose a novel ARS scheme, called Cheap Adaptive Rejection Sampling (CARS), where the computational effort for drawing from the proposal remains constant, decided in advance by the user. For generating a large number of desired samples, CARS is faster than ARS.

Key words: Monte Carlo methods, Rejection Sampling, Adaptive Rejection Sampling

1. Introduction

Random variate generation is required in different fields and several applications [Devroye, 1986, Hörmann et al., 2003, Robert and Casella, 2004]. Rejection sampling (RS) [Robert and Casella, 2004, Chapter 2] is a universal sampling method which generates independent samples from a target probability density function (pdf). The sample is either accepted or rejected by an adequate test of the ratio of the two pdfs. However, RS needs to establish analytically a bound for the ratio of the target and proposal densities.

Given a target density, the adaptive rejection sampling (ARS) method [Gilks and Wild, 1992, Gilks, 1992] produces jointly both a suitable proposal pdf and the upper bound for the ratio of the target density over this proposal. Moreover, the main advantage of ARS is that ensures high acceptance rates, since ARS yields a sequence of proposal functions that actually converge toward the target pdf when the procedure is iterated. The construction of the proposal pdf is obtained by a non-parametric procedure using a set of support points (nodes), with increasing cardinality. When a sample is rejected in the RS test, it is added to the set of support points. One limitation of ARS is that it can be applied only with (univariate) log-concave target densities.[1] For this reason, several extensions have been proposed [Hörmann, 1995, Hirose and A.Todoroki, 2005, Evans and Swartz, 1998, Görür and Teh, 2011, Martino and Míguez, 2011], even mixing with MCMC techniques [Gilks et al., 1995, Martino et al., 2013, 2015a]. A related RS-type method, automatic but non-adaptive, that employs a piecewise constant construction of the proposal density obtained with a pruning of the initial nodes, has been suggested in [Martino et al., 2015b].

In this work, we focus on the computational cost required by ARS. The ARS algorithm obtains high acceptance rates improving the proposal function, which becomes closer and closer to target function. Hence, this enhancement of the acceptance rate is obtained building more complex proposals, which become more computational demanding. The overall time of ARS depends on both the acceptance rate and the time required for sampling from the proposal pdf. The computational cost of ARS remains bounded since the probability of updating the proposal pdf, P_t, vanishes to zero as the number of iterations t grows. However, for a finite t, there is always a positive probability $P_t > 0$ of improving the proposal function, producing an increase of the acceptance rate. This enhancement of the acceptance rate could not balance out the increase of the time required for drawing from the new updated proposal function. Namely, if the acceptance rate is enough close to 1, a further improvement of the proposal function could become prejudicial.

Thus, we propose a novel ARS scheme, called Cheap Adaptive Rejection Sampling (CARS), employing always a fixed number of nodes, i.e., the com-

[1] The possibility of applying ARS for drawing for multivariate densities depends on the ability of constructing a sequence of non-parametric proposal pdfs in higher dimensions. See, for instance, the piecewise constant construction in [Martino et al., 2015a] as a simpler alternative procedure.

putational effort required for sampling from the proposal remains constant, selected in advance by the user. The new technique is able to increase the acceptance rate on-line in the same fashion of the standard ARS method, improving adaptively the location of the support points. The configuration of the nodes converges to the best possible distribution which maximizes the acceptance rate achievable with a fixed number of support points. Clearly, the maximum obtainable acceptance rate with CARS is always smaller than 1, in general. However, for large value of required samples, the CARS algorithm is faster than ARS for generating independent samples from the target, as shown the numerical simulations.

2. Adaptive Rejection Sampling

We denote the target density as

$$\bar{\pi}(x) = \frac{1}{c_\pi}\pi(x) = \frac{1}{c_\pi}\exp\big(V(x)\big), \quad x \in \mathcal{X} \subseteq \mathbb{R},$$

with $c_\pi = \int_{\mathcal{X}} \pi(x)dx$. The adaptive proposal pdf is denoted as

$$\bar{q}_t(x) = \frac{1}{c_t}q_t(x) = \frac{1}{c_\pi}\exp\big(W_t(x)\big),$$

where $c_t = \int_{\mathcal{X}} q_t(x)dx$ and $t \in \mathbb{N}$. In order to apply rejection sampling (RS), it is necessary to build $q_t(x)$ as an envelope function of $\pi(x)$, i.e.,

$$q_t(x) \geq \pi(x), \quad \text{or} \quad W_t(x) \geq V(x), \tag{1}$$

for all $x \in \mathcal{X}$ and $t \in \mathbb{N}$. As a consequence, it is important to observe that

$$c_t \geq c_\pi, \qquad \forall t \in \mathbb{N}. \tag{2}$$

Let us assume that $V(x) = \log \pi(x)$ is concave, and we are able to evaluate the function $V(x)$ and its first derivative $V'(x)$.[2] The adaptive rejection

[2]The evaluation of $V'(x)$ is not strictly necessary, since the function $q_t(x)$ can also construct using a derivative-free procedure (e.g., see [Gilks, 1992] or the piecewise constant construction in [Martino et al., 2015a]). For the sake of simplicity, we consider the construction involving tangent lines.

sampling (ARS) technique [Gilks, 1992, Gilks and Wild, 1992] considers a set of *support points* at the t-th iteration,

$$\mathcal{S}_t = \{s_1, s_2, \ldots, s_{m_t}\} \subset \mathcal{X},$$

such that $s_1 < \ldots < s_{m_t}$ and $m_t = |\mathcal{S}_t|$, for constructing the envelope function $q_t(x)$ in a non-parametric way. We denote as $w_i(x)$ as the straight line tangent to $V(x)$ at s_i for $i = 1, \ldots, m_t$. Thus, we can build a piecewise linear function,

$$W_t(x) = \min[w_1(x), \ldots, w_{m_t}(x)], \quad x \in \mathcal{X}. \tag{3}$$

Hence, the proposal pdf defined as $\bar{q}_t(x) \propto q_t(x) = \exp(W_t(x))$, is formed by exponential pieces in such a way that $W_t(x) \geq V(x)$, so that $q_t(x) \geq \pi(x)$, when $V(x)$ is concave (i.e., $\pi(x)$ is log-concave). Figure 1 depicts an example of piecewise linear function $W_t(x)$ built with $m_t = 3$ support points.

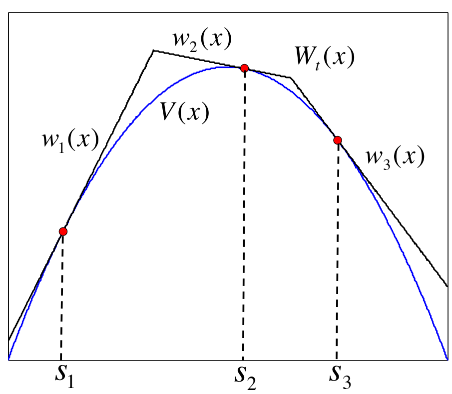


Figure 1: Example of construction of the piecewise linear function $W_t(x)$ with $m_t = 3$ support points, such that $W_t(x) \geq V(x)$.

Table 1 summarizes the ARS algorithm for drawing N independent samples from $\bar{\pi}(x)$. At each iteration t, a sample x' is drawn from $\bar{q}_t(x)$ and accepted with probability $\frac{\pi(x')}{q_t(x')}$, otherwise is rejected. Note that a new point is added to the support set $\mathcal{S}_t$ whenever it is rejected in the RS test improving the construction of $q_t(x)$. Clearly, denoting as T the total number of iterations of the algorithm, we have always $T \geq N$ since several samples are discarded.

Table 1: Adaptive Rejection Sampling (ARS) algorithm.

Initialization:

1. Set $t = 0$ and $n = 0$. Choose an initial set $\mathcal{S}_0 = \{s_1, \ldots, s_{m_0}\}$.

Iterations (while $n < N$):

2. Build the proposal $q_t(x)$, given the set of support points $\mathcal{S}_t = \{s_1, \ldots, s_{m_t}\}$, according to Eq. (3).
3. Draw $x' \sim \bar{q}_t(x) \propto q_t(x)$ and $u' \sim \mathcal{U}([0,1])$.
4. If $u' > \frac{\pi(x')}{q_t(x')}$, then reject x', update
$$\mathcal{S}_{t+1} = \mathcal{S}_t \cup \{x'\},$$
and set $t = t + 1$. Go back to step 2.
5. If $u' \leq \frac{p(x')}{\pi_t(x')}$, then accept x', setting $x_n = x'$.
6. Set $\mathcal{S}_{t+1} = \mathcal{S}_t$, $t = t + 1$, $n = n + 1$ and return to step 2.

Outputs: The N accepted samples $x_1, \ldots, x_N$.

3. Computational cost of ARS

The computational cost of an ARS-type method depends on two elements:

1. The averaged number of accepted samples, i.e., the acceptance rate.
2. The computational effort required for sampling from $q_t(x)$.

We desire that the acceptance rate is close to 1 and, simultaneously, that the spent time required for drawing from $q_t(x)$ is small. In general, there exists a trade-off since an increase of the acceptance rate requires the use of a more complicated proposal density $q_t(x)$. ARS is an automatic procedure which provides a possible compromise. Below, we analyze some important features of a standard ARS scheme.

3.1. Acceptance rate

The averaged number of accepted samples, i.e., the acceptance rate, is

$$\eta_t = \int \frac{\pi(x)}{q_t(x)} \bar{q}_t(x) dx = \frac{c_\pi}{c_t}, \tag{4}$$

that is $0 \leq \eta_t \leq 1$ since $c_t \geq c_\pi$, $\forall t \in \mathbb{N}$, by construction. Defining the L_1 distance between $\pi_t(x)$ and $p(x)$ as

$$D(q_t, \pi) = \|q_t(x) - \pi(x)\|_1 = \int_{\mathcal{X}} |q_t(x) - \pi(x)| dx, \tag{5}$$

ARS ensures that $D(q_t, \pi) \to 0$ when $t \to \infty$, and as a consequence $c_t \to c_\pi$. Thus, η_t tends to one as $t \to \infty$. Indeed, as $\eta_t \to 1$, ARS becomes virtually an exact sampler after a some iterations.

3.2. Drawing from the proposal pdf

Let us denote the exponential pieces as

$$h_i(x) = e^{w_i(x)}, \quad i = 1, \ldots, N, \tag{6}$$

so that

$$q_t(x) = h_i(x), \qquad \text{for} \qquad x \in \mathcal{I}_i = (e_{i-1}, e_i], \quad i = 1, \ldots, N,$$

where e_i is the intersection point between the straight lines $w_i(x)$ and $w_{i+1}(x)$, for $i = 2, \ldots, N-1$, and $e_0 = -\infty$ and $e_N = +\infty$ (if $\mathcal{X} = \mathbb{R}$). Thus, for drawing a sample x' from $\bar{q}_t(x) = \frac{1}{c_t} q_t(x)$, we need to:

1. Compute analytically the area A_i below each exponential piece, i.e., $A_i = \int_{\mathcal{I}_i} h_i(x) dx$ and obtain the normalized weights

$$\rho_i = \frac{A_i}{\sum_{n=1}^N A_n} = \frac{A_i}{c_t}, \tag{7}$$

   where we have observed that $c_t = \sum_{n=1}^N A_n = \int_{\mathcal{X}} q_t(x) dx$.

2. Select an index j^* (namely, one piece) according to the probability mass ρ_i, $i = 1, \ldots, N$.

3. Draw x' from $h_{j^*}(x)$ restricted within the domain $\mathcal{I}_{j^*} = (e_{j^*-1}, e_{j^*}]$, and zero outside (i.e., from a truncated exponential pdf).

Observe that, at step 2, a multinomial sampling is required. It is clear that the computational cost for drawing one sample from $q_t(x)$ increases as the number of pieces grows or, equivalently, the number of support points grows. Fortunately, the computational cost in ARS is automatically controlled by the algorithm, since the probability of adding a new support point

$$P_t = 1 - \eta_t = \frac{1}{c_t} D(q_t, \pi), \tag{8}$$

tends to zero as $t \to \infty$, since the distance in Eq. (5) vanishes to zero, i.e., $D(q_t, \pi) \to 0$.

4. ARS with fixed number of support points

We have seen that the probability of adding a new support point P_t vanishes to zero as $t \to \infty$. However, for a finite t, we have always a positive probability $P_t > 0$ of adding a new point (although small), so that a new support point could be incorporated producing an increase of the acceptance rate. After a certain iteration τ, i.e., $t > \tau$, this improvement of the acceptance rate could not balance out the increase of the time required for drawing from the proposal, due to the addition of the new point. Namely, if the acceptance rate is enough close to 1, a further addition of a support point could slow down the algorithm, becoming prejudicial.

In this work, we provide an alternative adaptive procedure for ARS, called *Cheap Adaptive Rejection Sampling* (CARS), which uses a fixed number of support points. When a sample is rejected, a test for swapping the rejected sample with the closest support point within $\mathcal{S}_t$ is performed, so that the total number of points remains constant. Unlike in the standard ARS method, in the new adaptive scheme the test is deterministic. The underlying idea is based on the following observation. The standard ARS algorithm yields a decreasing sequence of normalizing constants $\{c_t\}_{t\in\mathbb{N}}$ of the proposal pdf converging to $c_\pi = \int_{\mathcal{X}} \pi(x)dx$, i.e.,

$$c_0 \geq c_1 \ldots \geq c_t \ldots \geq c_\infty = c_\pi. \tag{9}$$

Clearly, since the acceptance rate is $\eta_t = \frac{c_\pi}{c_t}$ this means that $\eta_t \to 1$. In CARS, we provide an alternative way for producing this decreasing sequence of normalizing constants $\{c_t\}$. Indeed, an exchange between two points is accepted if it produces a reduction in the normalizing constant of the corresponding proposal pdf. More specifically, consider the set

$$\mathcal{S}_t = \{s_1, s_2, \ldots, s_M\},$$

contained M support points. When a sample x' is rejected in the RS test, the closest support point s^* in $\mathcal{S}_t$ is obtained, i.e.,

$$s^* = \arg\min_{s_i \in \mathcal{S}_t} |s_i - x'|.$$

We recall that we denote with $q_t(x)$ the proposal pdf built using $\mathcal{S}_t$ and with c_t its normalizing constant. Then, we consider a new set

$$\mathcal{G} = \mathcal{S}_t \cup \{x'\} \backslash \{s^*\}, \tag{10}$$

namely, including x' and removing s^*. We denote with $g(x)$ the proposal built using the alternative set of support points $\mathcal{G}$, and $c_g = \int_{\mathcal{X}} g(x)dx$. If

$$c_g < c_t,$$

then the swap is accepted, i.e., we set $\mathcal{S}_{t+1} = \mathcal{G}$ for the next iteration, otherwise the set remains unchanged, $\mathcal{S}_{t+1} = \mathcal{S}_t$. The complete algorithm is outlined in Table 2. Note that c_t is always computed (in any case, for both ARS and CARS) at the step 3, for sampling from $q_t(x)$. Furthermore observe that, after the first iteration, step 2 can be skipped since the new proposal pdf $q_{t+1}(x)$ has been already constructed in the previous iteration, i.e., $q_{t+1}(x) = q_t(x)$, or at step 4.3, i.e., $q_{t+1}(x) = g(x)$.

Therefore, with the CARS algorithm, we obtain again a decreasing sequence of $\{c_t\}_{t\in\mathbb{N}}$

$$c_0 \geq c_1 \ldots \geq c_t \ldots \geq c_\infty,$$

but $c_\infty \neq c_\pi$ so that $\eta_t \to \eta_\infty < 1$, in general. The value η_∞ is the highest acceptance rate that can be obtained with M support points, given the target function $\pi(x)$. Therefore, CARS yields a sequence of sets $\mathcal{S}_1, \ldots, \mathcal{S}_t, \ldots$ that converges to the stationary set $\mathcal{S}_\infty$ containing the best configuration of M support points for maximizing the acceptance rate, when the target function is $\pi(x)$ and given a specific construction procedure for the proposal $q_t(x)$.[3]

[3] The best configuration $\mathcal{S}_\infty$ depends on the specific construction procedure employed for building the sequence of proposal functions $q_1, q_2 \ldots, q_t, \ldots$

Table 2: Cheap Adaptive Rejection Sampling (CARS) algorithm.

Initialization:

1. Set $t = 0$ and $n = 0$. Choose a value M and an initial set $\mathcal{S}_0 = \{s_1, \dots, s_M\}$.

Iterations (while $n < N$):

2. Build the proposal $q_t(x)$, given the current set $\mathcal{S}_t$, according to Eq. (3) or other suitable procedures.
3. Draw $x' \sim \bar{q}_t(x) \propto q_t(x)$ and $u' \sim \mathcal{U}([0,1])$.
4. If $u' > \frac{\pi(x')}{q_t(x')}$, then reject x' and:
   4.1 Find the closest point s^* in $\mathcal{S}_t$,
   $$s^* = \arg\min_{s_i \in \mathcal{S}_t} |s_i - x'|.$$
   4.2 Build the alternative proposal $g(x)$ based on the set of points
   $$\mathcal{G} = \mathcal{S}_t \cup \{x'\} \backslash \{s^*\}$$
   and compute $c_g = \int_{\mathcal{X}} g(x)dx$.
   4.3 If $c_g < c_t$, set $\mathcal{S}_{t+1} = \mathcal{G}$, otherwise, if $c_g \geq c_t$, set $\mathcal{S}_{t+1} = \mathcal{S}_t$. Set $t = t + 1$ and go back to step 2.
5. If $u' \leq \frac{p(x')}{\pi_t(x')}$, then accept x', setting $x_n = x'$.
6. Set $\mathcal{S}_{t+1} = \mathcal{S}_t$, $t = t + 1$, $n = n + 1$ and return to step 2.

Outputs: The N accepted samples $x_1, \dots, x_N$.

5. Numerical simulations

We consider a Gaussian density as (typical) log-concave target pdf and test both ARS and CARS. Namely, we consider

$$\bar{\pi}(x) \propto \pi(x) = \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad x \in \mathbb{R},$$

with $\sigma^2 = \frac{1}{2}$. We compare ARS and CARS in terms of the time required for generating $N \in \{5000, 1000\}$ samples. In all cases and both techniques, we consider a initial set of support points $\mathcal{S}_0 = \{s_1, \ldots, s_0\}$ with cardinality $m_0 = |\mathcal{S}_0| \in \{3, 5, 10\}$ (clearly, $M = m_0$ in CARS) where the initial points are chosen uniformly in $[-2, 2]$ at each simulation, i.e., $s_i \sim \mathcal{U}([-2, 2])$.[4]

We run 500 independent simulations for each case and compute the required time for generating N samples (using a Matlab code), the averaged number of final support points (denote as $E[m_T]$) and the acceptance rate reached in the final iteration (denoted as $E[\eta_T]$; averaged over the 500 runs), for both techniques. Table 3 shows the results. The time is normalized with respect to the time spent by ARS with $N = 5000$, $|\mathcal{S}_0| = 3$. The results show that CARS is always faster than ARS. We can observe that both methods obtain acceptance rates close to 1. CARS reaches acceptance rates always greater of 0.87 using only 3 nodes. CARS obtains an more than 0.98 employing only 10 nodes and after generating $N = 5000$ independent samples. Fig. 2 depicts the wasted time, the final acceptance rate and the final number of nodes, as function of number N of generated samples. We can observe that CARS is significantly faster than ARS when N grows, owing to ARS yields a sensible increase of the number of support points that corresponds to an infinitesimal increase of the acceptance rate, whereas CARS the number of nodes remains constant. Figure 3 shows a sequence of proposal pdfs constructed by CARS, using 3 nodes and starting with $S_0 = \{-1.5, -1, 1.8\}$. The L_1 distance $D(q_t, \pi)$ is reduced progressively and the acceptance rate improved. The final set of support point is $S_t = \{-1.0261, -0.0173, 1.0305\}$, close to the optimal one $S_\infty = \{-1, 0, 1\}$.

6. Acknowledgements

This work has been supported by the Grant 2014/23160-6 of São Paulo Research Foundation (FAPESP) and by the Grant 305361/2013-3 of National Council for Scientific and Technological Development (CNPq).

References

L. Devroye. *Non-Uniform Random Variate Generation*. Springer, 1986.

[4] Clearly, the configurations of either all negative or all positive are discarded since they yield improper proposal pdf by construction.

Table 3: Results as function of the desired number of samples N and the cardinality $|\mathcal{S}_0|$ of the initial set of support points $\mathcal{S}_0$. We show the normalized spent time, the averaged final number of support points, $E[m_T]$, and the averaged final acceptance rate, $E[\eta_T]$.

Scheme	N	$\lvert\mathcal{S}_0\rvert = 3$	$\lvert\mathcal{S}_0\rvert = 5$	$\lvert\mathcal{S}_0\rvert = 10$
ARS	5000	Time=1 $E[\eta_T] = 0.9942$ $E[m_T] = 32.36$	Time=0.9709 $E[\eta_T] = 0.9945$ $E[m_T] = 32.69$	Time=0.9801 $E[\eta_T] = 0.9952$ $E[m_T] = 34.17$
CARS	5000	Time=0.9599 $E[\eta_T] = 0.8721$ $E[m_T] = M = 3$	Time=0.9477 $E[\eta_T] = 0.9224$ $E[m_T] = M = 5$	Time=0.9694 $E[\eta_T] = 0.9556$ $E[m_T] = M = 10$
ARS	10000	Time=2.2843 $E[\eta_T] = 0.9963$ $E[m_T] = 40.60$	Time=1.9862 $E[\eta_T] = 0.9964$ $E[m_T] = 41.09$	Time=1.9983 $E[\eta_T] = 0.9968$ $E[m_T] = 42.16$
CARS	10000	Time=1.9716 $E[\eta_T] = 0.8784$ $E[m_T] = M = 3$	Time=1.7311 $E[\eta_T] = 0.9350$ $E[m_T] = M = 5$	Time=1.8969 $E[\eta_T] = 0.9631$ $E[m_T] = M = 10$
ARS	50000	Time=11.2196 $E[\eta_T] = 0.9987$ $E[m_T] = 68.63$	Time=11.2887 $E[\eta_T] = 0.9987$ $E[m_T] = 69.56$	Time=11.7599 $E[\eta_T] = 0.9988$ $E[m_T] = 70.09$
CARS	50000	Time=8.7756 $E[\eta_T] = 0.8855$ $E[m_T] = M = 3$	Time=8.4322 $E[\eta_T] = 0.9540$ $E[m_T] = M = 5$	Time=9.0704 $E[\eta_T] = 0.9861$ $E[m_T] = M = 10$

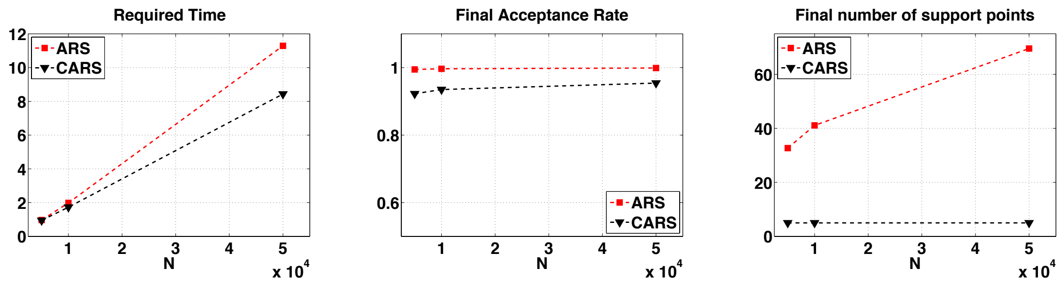
W. Hörmann, J. Leydold, and G. Derflinger. *Automatic nonuniform random variate generation*. Springer, 2003.

C. P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer, 2004.

W. R. Gilks and P. Wild. Adaptive Rejection Sampling for Gibbs Sampling. *Applied Statistics*, 41(2):337–348, 1992.

W. R. Gilks. Derivative-free Adaptive Rejection Sampling for Gibbs Sampling. *Bayesian Statistics*, 4:641–649, 1992.

L. Martino, J. Read, and D. Luengo. Independent doubly adaptive rejection Metropolis sampling within Gibbs sampling. *IEEE Transactions on Signal Processing*, 63(12):3123–3138, 2015a.



(a) (b) (c)

Figure 2: **(a)** Spent time, **(b)** final acceptance rate, and **(c)** final number of support points, as function of the number N of drawn samples, for ARS (squares) and CARS (triangles).

W. Hörmann. A rejection technique for sampling from T-concave distributions. *ACM Transactions on Mathematical Software*, 21(2):182–193, 1995.

H. Hirose and A.Todoroki. Random number generation for the generalized normal distribution using the modified adaptive rejection method. *International Information Institute*, 8(6):829–836, March 2005.

M. Evans and T. Swartz. Random variate generation using concavity properties of transformed densities. *Journal of Computational and Graphical Statistics*, 7(4):514–528, 1998.

Dilan Görür and Yee Whye Teh. Concave convex adaptive rejection sampling. *Journal of Computational and Graphical Statistics*, 20(3):670–691, September 2011.

L. Martino and J. Míguez. A generalization of the adaptive rejection sampling algorithm. *Statistics and Computing*, 21(4):633–647, October 2011.

W. R. Gilks, N. G. Best, and K. K. C. Tan. Adaptive Rejection Metropolis Sampling within Gibbs Sampling. *Applied Statistics*, 44(4):455–472, 1995.

L. Martino, R. Casarin, F. Leisen, and D. Luengo. Adaptive Sticky Generalized Metropolis. *arXiv:1308.3779*, 2013.

L. Martino, H. Yang, D. Luengo, J. Kanniainen, and J. Corander. A fast universal self-tuned sampler within Gibbs sampling. *(in press) Digital Signal Processing*, 2015b.

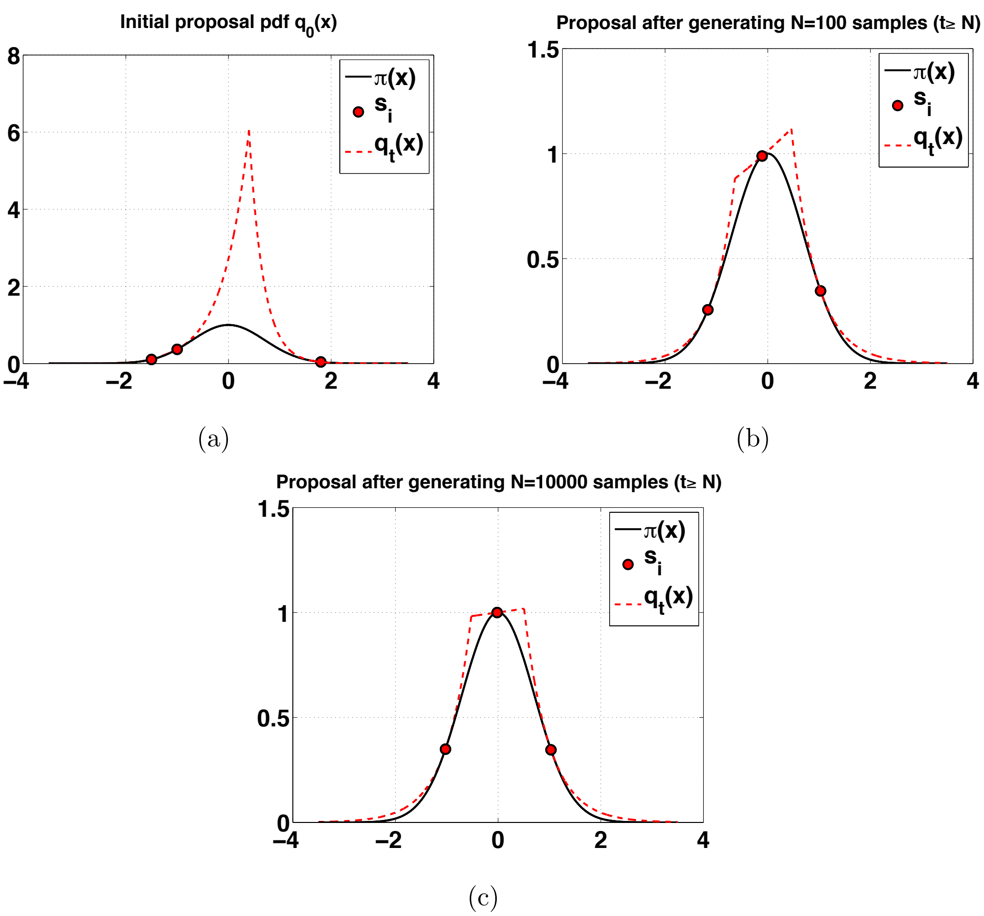


Figure 3: Example of sequence of proposal pdfs obtained by CARS, starting with $S_0 = \{-1.5, -1, 1.8\}$. We can observe that the L_1 distance $D(q_t, \pi)$ is reduced progressively. The proposal function $q_t(x)$ is depicted with dashed line, the target function $\pi(x)$ with solid line and the support points with circles. The configuration of the nodes in figure **(c)** is $S_t = \{-1.0261, -0.0173, 1.0305\}$ with $t \geq N = 10^4$. The optimal configuration with 3 nodes and $\pi(x) = \exp\left(-x^2\right)$ is $S_\infty = \{-1, 0, 1\}$.