APPLICATION OF SEQUENTIAL QUASI-MONTE CARLO TO AUTONOMOUS POSITIONING

Nicolas Chopin *
CREST-ENSAE
92245 Malakoff
France

Mathieu Gerber †
Harvard University
Department of Statistics
Cambridge, MA

ABSTRACT

SMC (Sequential Monte Carlo) algorithms (also known as particle filters) are popular methods to approximate filtering (and related) distributions of state-space models. However, they converge at the slow $1/\sqrt{N}$ rate, which may be an issue in real-time data-intensive scenarios. We give a brief outline of SQMC (Sequential Quasi-Monte Carlo), a variant of SMC based on low-discrepancy point sets proposed by [1], which converges at a faster rate, and we illustrate the greater performance of SQMC on autonomous positioning problems.

Index Terms—Low-discrepancy point sets; Particle filtering; Quasi-Monte Carlo

1. INTRODUCTION

Many problems in signal processing (and related fields) can be formalised as the filtering of data $(y_t)$ to recover an unobserved signal $(x_t)$ that follows a state-space model. For non-linear and/or non-Gaussian state-space models, particle filtering [2, 3], also known as Sequential Monte Carlo (SMC), is now the standard approach to perform filtering; see e.g. [4]. However, a potential drawback of SMC for real time applications is its slow $1/\sqrt{N}$ convergence rate (based on $N$ simulations, or ‘particles’). In real time problems, the running time per iteration of the filtering algorithm is bounded by the time interval between successive observations and, consequently, this slow convergence rate implies that in some settings the approximation error of SMC might be non negligible.

Recently, [1] proposed and studied the sequential quasi-Monte Carlo (SQMC) algorithm, which is a quasi-Monte Carlo (QMC) version of particle filtering. Based on $N$ particles, SQMC has the advantage to converge at rate $o(1/\sqrt{N})$, i.e. at a faster rate than SMC; see Theorem 7 of [1]. On the other hand, SQMC requires $O(N \log N)$ operations and is thus slower than SMC, which has complexity $O(N)$. But [1] show that, in several scenarios, the faster convergence of SQMC does more than compensate its slower running time and, consequently, for a given computational budget, SQMC typically achieves a significantly smaller error size than SMC.

In this paper we propose to apply SQMC to the problem of autonomous positioning of a vehicle moving along a two dimensional space where, following [5], we assume that the Markov transition is non Gaussian. Our numerical study show that for this real time application SQMC provides a much more accurate estimation of the position of the vehicle than SMC.

2. SEQUENTIAL QUASI-MONTE CARLO

2.1. Background on sequential Monte Carlo

To introduce SMC we consider the following generic state-space model, described in terms of probability density functions:

\[
\begin{align*}
  y_t | x_t & \sim f^Y (y_t | x_t), \quad t \geq 0 \\
  x_t | x_{t-1} & \sim f^X (x_t | x_{t-1}), \quad t \geq 1 \\
  x_0 & \sim f^X_0 (x_0)
\end{align*}
\]

(1)

where $(x_t)_{t \geq 0}$ is the unobservable Markov process on $\mathcal{X} \subseteq \mathbb{R}^d$ and $(y_t)_{t \geq 0}$ is the observation process.

The typical quantity of interest in state-space models is the filtering distribution, that is, the distribution of $x_t$ given all the available observations at time $t$, which is given by

\[
p(x_t | y_{0:t}) = \frac{1}{Z_t} \times \int_{\mathcal{X}^t} f^X_0 (x_0) \prod_{s=1}^{t} f^X (x_s | x_{s-1}) \prod_{s=0}^{t} f^Y (y_s | x_s) dx_{0:t-1}
\]

(2)

where $Z_t$ is a normalising constant. Except in linear Gaussian models, the integrals in (2) are not tractable, but one may instead run a particle filter to sequentially approximate $p(x_t | y_{0:t})$.

The basic idea of particle filtering is to use the Markov transition $f^X (x_t | x_{t-1})$ to propagate the discrete approxima-
tion (for \( t \geq 1 \))

\[
p^N(x_{t-1}|y_{0:t-1}) = \sum_{n=1}^{N} W^n_{t-1} \delta_{x^n_{t-1}}, \quad (d x_{t-1}),
\]

with \( \sum_{n=1}^{N} W^n_{t-1} = 1, \ W^n_{t-1} \geq 0 \)

of the filtering distribution at time \( t-1 \) to the approximation

\[
p^N(x_{t-1}|y_{0:t-1}) = p^N(x_{t-1}|y_{0:t-1}) f^N(x_t|x_{t-1})
\]

(3)

of \( p(x_{t-1}|y_{0:t-1}) \). Then, the marginal distribution of \( x_t \) with respect to

\[
p^N(x_{t-1},y_{0:t-1}) \propto p^N(x_{t-1}|y_{0:t-1}) f^N(y_t|x_t)
\]

(4)

may be used as an approximation of the filtering distribution at time \( t \). Thus, one can perform an importance sampling step, with proposal distribution (3) and target distribution (4), to get a weighted particle system \( \{W^n_{t}, x^n_{t}\}_{n=1}^{N} \) which is approximately distributed from \( p(x_t|y_{0:t}) \); see Algorithm 1 for a more precise description of particle filtering.

Algorithm 1 SMC Algorithm (Bootstrap filter)

Operations must be performed for all \( n \in 1 : N \)

Sample \( x^n_0 \) from \( f^N_0(x_0) \) and compute \( W^0_0 = f^N(y_0|x^n_0)/\sum_{n=1}^{N} f^N(y_0|x^n_0) \)

for \( t = 1, \ldots, T \) do

Sample \( a^n_{t-1} \) from \( \mathcal{M}(W^{1:N}_{t-1}) \), the multinomial distribution that produces outcome \( m \) with probability \( W^{m}_{t-1} \).

Sample \( x^n_t \) from \( f^N(x_t|x^n_{t-1}) \) and compute \( W^n_t = f^N(y_t|x^n_t)/\sum_{n=1}^{N} f^N(y_t|x^n_t) \)

end for

2.2. Background on quasi-Monte Carlo

Loosely speaking, a QMC point set \( u^{1:N} \) in \([0,1]^d \) is a set of (deterministic) points which are “more uniformly” distributed than uniform random samples. The most classical measure of uniformity in the QMC literature is the so called star discrepancy, defined by

\[
D^*(u^{1:N}) = \sup_{b \in (0,1)^d} \left| \frac{1}{N} \sum_{n=1}^{N} \mathbb{I}(u^n \in [0,b]) - \frac{d}{b} \right|,
\]

where \( b = (b_1, \ldots, b_d) \). We say that \( u^{1:N} \) is a QMC point set if \( D^*(u^{1:N}) = O(N^{-1}(\log N)^d) \).

The main motivation for using low discrepancy point sets in numerical integration is the Koksma–Hlawka inequality:

\[
\left| \frac{1}{N} \sum_{n=1}^{N} \varphi(u^n) - \int_{[0,1]^d} \varphi(u) \, du \right| \leq V(\varphi) D^*(u^{1:N})
\]

which explicitly links the integration error and the equidistribution property of the point set at hand, because the quantity \( V(\varphi) \) only depends on the integrand \( \varphi \); see e.g. Chap. 5 of [6] for a definition of \( V(\varphi) \).

A useful variant to QMC is randomised QMC (RQMC), which combines the advantages of random sampling and of QMC strategies. A RQMC point set \( u^{1:N} \) is such that \( u^n \sim \mathcal{U}([0,1]^d) \) for all \( n \in 1 : N \) and \( D^*(u^{1:N}) = O((N^{-1} \log N)^d) \) with probability one. A particularly interesting construction of RQMC point sets is the nested scrambled method for \((t,m,s)\)-nets (see e.g. [7], Chap. 4, for a definition) proposed by [8], which allows to approximate the integral of smooth functions with an error of size \( O(N^{-1.5 + \varepsilon}) \) for any \( \varepsilon > 0 \) [9]. In addition, and contrary to plain QMC, no smoothness assumptions on the integrand \( \varphi \) are needed for scrambled net quadrature rules to outperform Monte Carlo integration [9]. This last point is particularly important in the context of SMC because the resampling step (Step 4 of Algorithm 1) introduces discontinuities which cannot be efficiently handled by deterministic QMC strategies.

2.3. Sequential quasi-Monte Carlo

The basic idea of SQMC is to replace the sampling step from the proposal distribution (3) by a low discrepancy point set with respect to the same distribution.

The classical way to transform a low discrepancy point set with respect to the uniform distribution (i.e. a QMC point set) into a low discrepancy point set with respect to a non-uniform distribution \( \pi(x) \) on \( X \subset \mathbb{R}^d \) is to use the inverse of the Rosenblatt transformation of \( \pi \), defined by

\[
F_{\pi} (x) = (u_1, \ldots, u_d)^T, \quad x = (x_1, \ldots, x_d)^T \in X,
\]

where, \( u_1 = F_{\pi,1}(x_1) \), \( F_{\pi,1} \) being the CDF of the marginal distribution of the first component (relative to \( \pi \)), and for \( i \geq 2, u_i = F_{\pi,i}(x_i|x_{1:i-1}) \), \( F_{\pi,i} \) being the CDF of component \( x_i \), conditional on \( (x_1, \ldots, x_{i-1}) \), relative to \( \pi \).

Following this idea, and assuming for the moment that the state variable \( x_t \) is univariate, one can generate a low discrepancy point set \( (x^n_{t-1}, x^n_t) \) from (3) as follows: let \( u^{1:N}_t \) be a \((R)QMC \) point set in \([0,1]^2 \), with \( u^n = (u^n_1, u^n_2) \), and compute

\[
\hat{x}^{n}_{t-1} = F_{\pi}^{-1}(p^N(x_{t-1}|y_{0:t-1})(u^n_1), \quad x^n_t = F_{\pi}^{-1}(f^N(X_{t-1})(u^n_2)).
\]

However, the extension of this approach to \( d > 1 \) is not trivial because the distribution \( p^N(x_t|y_{0:t-1}) \) is then a (weighted) sum of Dirac measures over \( \mathbb{R}^d \).

To overcome this difficulty, [1] proposes to transform the multivariate (discrete) distribution \( p^N(x_{t-1}|y_{0:t-1})dx_t \) into a univariate (discrete) distribution \( p^N_{\mathcal{H}}(h_t|y_{0:t})dx_t \) on \([0,1]\) using the following change of variable

\[
x \in X \mapsto h \circ \psi(x) \in [0,1],
\]
The complexity of Algorithm 2 is $O(N \log N)$, because it performs two sorting steps at each iteration. Regarding the practical implementation of Algorithm 2, note that: (a) QMC generation (Steps 2 and 5) routines are available in most software (e.g. package randtoolbox in R, or class grandset in the Statistics toolbox of Matlab); to compute the $a^{N}_{t-1}$'s (Step 8), one may use the standard approach based on sorted uniforms for resampling; and (c) in order to compute $h$, see e.g. [10], and Chris Hamilton’s C++ program available at https://web.cs.dal.ca/~chamilton/hilbert/index.html.

Our SQMC implementation is available at https://bitbucket.org/mgerber/sqmc. We shall use RQMC (randomised QMC) point sets in our simulations (more precisely scrambled Sobol’ sequences; see [9, 11, 12] for more details on scrambling), as this makes it possible to evaluate the numerical error through repeated runs.

Finally, while we presented SQMC in this specific case where particles are mutated according to $f^N(x_i|x_{t-1})$, the Markov transition of the considered model, it of course extends directly to situations where particles are mutated according to some other kernel $q_t(x_i|x_{t-1})$ (assuming that the particles are reweighted accordingly, as in standard SMC).

### 3. APPLICATION: AUTONOMOUS POSITIONING

#### 3.1. Model description

We consider the problem of autonomous positioning of a vehicle moving in a two dimensional space. To determine its position, the vehicle estimates its speed every $T_s$ > 0 seconds and measures the power of $d_y$ ≥ 1 radio signals. We suppose that the radio signals are emitted from known locations $r_i \in \mathbb{R}^2$, $i = 1, \ldots, d_y$, and that the corresponding attenuation coefficients $a_i$ are known as well. This positioning problem admits the following state space representation (see [13] and [5])

$$y_t = 10 \log_{10} \left( \frac{P_0}{|r_t - x|^{2}} \right) + \nu_t, \quad t \geq 0$$

$$x_t = x_{t-1} + T_s \nu_t + T_s \epsilon_t, \quad t \geq 1$$

where $i \in 1 : d_y$, $x_i \in \mathbb{R}^2$ is the position of the vehicle at time $t$, $v_t$ is a measure of its speed, which is assumed to be constant over successive time intervals of $T_s$ seconds, $\epsilon_t$ and $\nu_t = (\nu_{t1}, \ldots, \nu_{td_y})$ represent measurement errors while $y_{it}$ is the power received at time $t$ by emitter $i$. In the sequel, $P_0$
and where the time varying constants have that, with high probability, $x_i$ starting at a location close to $(0,0)$. The dots show the locations of the 5 emitters.

is the initial signal from emitter $i$ and, following [5], we suppose that all the error terms are independent and distributed according to a Laplace distribution with parameter 0.5.

3.2. Simulation set-up

To compare the performance of SQMC and SMC for this tracking problem we simulate the trajectory of a vehicle evolving during 15 minutes according to (6). We assume that the sample period is $T_s = 1$ second, that $d_y = 5$ (5 emitters) and that $a_{ti} = 0.95$ for all $i = 1, \ldots, d_y$. The resulting trajectory and the locations of the emitters are shown in Figure 2.

The SMC algorithm is implemented using systematic re-sampling [14], which is usually recognised as being the most efficient resampling strategy.

SQMC is implemented using nested scrambled Sobol’ sequences for the point sets $u_i^1, N$. As described above, we need to use a mapping $\psi_i$ to map the particles generated at iteration $t$ of SQMC into $[0,1]^d$ before performing the Hilbert sort. We chose for $\psi_i$ a component-wise (re-scaled) logistic transform; that is, $\psi_i(x) = (\psi_{i1}(x_1), \psi_{i2}(x_2))$ with

$$\psi_{i1}(x_i) = \left[1 + \exp\left(-\frac{x_i - \bar{x}_i}{\bar{x}_i - \bar{x}_i}\right)\right]^{-1}, \quad i = 1, 2,$$

and where the time varying constants $\bar{x}_i$ and $\bar{x}_i$ are used to have that, with high probability, $x_i \in [\bar{x}_i, \bar{x}_i]$. A reasonable choice for $\bar{x}_i$ is thus

$$\bar{x}_i, \bar{x}_i = \sum_{s=0}^{t} v_s \pm 2\sqrt{\text{Var}(\bar{x}_0) + t T_s^2 \text{Var}(\epsilon_t)}.$$

More generally, we may use the same approach whenever the marginal distribution of $x_i$ may be computed, or alternatively one may simply rescale the particle cloud to zero mean and unit variance (component wise) before applying the logistic transformation.

Simulation results are presented for $N \in \{2^8, \ldots, 2^{16}\}$, where 2 is the base of the Sobol’ sequence. Taking a power of 2 for the number of simulations is the standard approach in QMC integration based on Sobol’ sequence because both good theoretical and empirical results are obtained for this choice of $N$. However, this restriction is non necessary for QMC to outperform Monte Carlo methods and little gain may be expected in the context of SQMC, see [15] for more details on this point.

3.3. Results

In Figure 3 we compare the mean square error (MSE) of the filtering expectation estimate obtained from SQMC and SMC, as a function of $t$, for $N \in \{2^5, 2^{10}, 2^{16}\}$. To save space, only the results for the first component of $x_i$ are presented; the results for the second component are essentially the same. One observes that the performance gain of SQMC (relative to standard particle filtering) increases quickly with $N$.

We now study the amount of CPU time required to have a “reasonable” Monte Carlo error using both SMC and SQMC. Letting $\hat{x}_i$ be an estimate of the filtering expectation $E[x_i | y_0:t]$, we consider the Monte Carlo error to be reasonable if it is small compared to the posterior variance, that is, if $\text{MSE}(\hat{x}_i) \leq \delta^2 \text{Var}(x_i | y_0:t)$ for $i = 1, 2$ and where we set $\delta = 0.01$.

Figure 4 shows the number of time steps $t \in \{0, \ldots, 899\}$ for which this criterion is not met, as a function of the CPU budget (i.e. CPU time per iteration). To increase the CPU budget, we simply increase $N$. We observe that much better results are achieved using SQMC. Indeed, when the CPU budget is 0.05s per iteration, the SMC error is too large for more than 600 time steps, while a CPU budget of 0.07s is enough to estimate both coordinates of $x_i$ for all iterations with SQMC.

4. CONCLUSION

In this paper we have illustrated the potential of sequential quasi-Monte Carlo for real time signal processing problems with a non-linear and non-Gaussian state-space model for autonomous positionning. Compared to Monte Carlo particle filtering, dramatic variance reductions are observed when SQMC is used, both as a function of the number of particles and of CPU time. In real time application, the running time of the filtering algorithm is a crucial element and, concerning this point, we believe that significant improvement can be achieved for SQMC, notably concerning the Hilbert sort step. For instance, the computations of the Hilbert indices involve only bits operations and therefore GPU computing may allow for dramatic cost reductions.
Fig. 3. Filtering of the state-space model (6): The plot gives the gain factor, defined as the MSE(SMC) over MSE(SQMC), as a function of \( t \) for the estimation of \( \mathbb{E}[x_{1t}|y_{0:t}] \). The results are obtained from 100 independent runs of SMC and SQMC.

Fig. 4. Filtering of the state-space model (6): The plot gives the number of time steps \( t \in \{0, \ldots, 899\} \) such that \( \text{MSE}(\hat{x}_{it}) \geq 0.01^2 \text{Var}(x_{it}|y_{0:t}) \) as a function of the CPU budget (average CPU time per iteration), where \( \hat{x}_{it} \) is either the SQMC (solid lines) or the SMC (dashed lines) estimate of \( \mathbb{E}[x_{it}|y_{0:t}] \), \( i = 1, 2 \). The results are obtained from 100 independent runs of SMC and SQMC.

REFERENCES