Markov chain Monte Carlo (MCMC) is a preferred sampling method due to its wide applicability. However, it suffers from a slow scaling of $n^{-3}$, analogously to standard Monte Carlo. For the latter, scaling close to $n^{-1}$ can be achieved by using more homogeneously distributed points in space than the usual pseudo-random numbers. With the same philosophy in mind we apply quasi-Monte Carlo to an MCMC method that uses multiple proposals. We show consistency and achieve an improved scaling of close to $n^{-1}$ numerically.

Using multiple proposals in MCMC

Metropolis-Hastings (M-H) was generalised in (Cal14) such that it allows for parallelisation of a single chain by proposing multiple points via a kernel $k$ in parallel. In every MCMC iteration, a finite state Markov chain on the proposed points, i.e. $\tilde{x}_1, ..., \tilde{x}_{N_d}$, is constructed, as displayed in the figure on the right, from which samples are drawn.

In practice, generating proposals $x_i = x_\lfloor i/N_d \rfloor$ from $k(x_i, \cdot)$ is typically performed via applying its generalised inverse $\Psi_{k_i}$ to a set of pseudo-random numbers $u_1, ..., u_{N_d} \in [0, 1]$ such that $x_i = \Psi_{k_i}(u_1, ..., u_{N_d})$. The overall procedure is described in Algorithm 1.

**Algorithm 1: Multiple proposal Metropolis-Hastings**

**Input**: Initialise starting point $x_0 \in \Omega$ and number of proposals $N_d$, auxiliary variable $i = 1$ and counter $n = 0$, generate an random sequence $u_1, u_2, ..., u_{N_d} \in (0, 1]$;

1. for each MCMC iteration $i = 1, 2, ...$ do

2. Set $u_i = (u_{i,0}, u_{i,1}, ..., u_{i,N_d}) \in (0, 1)^{N_d}$ and sample $\tilde{x}_i$ conditioned on $i$, i.e., draw $N_d$ new points from $k(\tilde{x}_i, \cdot) = \tilde{p}(\tilde{x}_i|\tilde{x}_0)\tilde{p}(\tilde{x}_0)$ by the inverse $\Psi_{k_i}(u_i)$;

3. Compute the stationary distribution of $f(\tilde{x}_{i+1}, \cdot)$, i.e. $p(\tilde{f} = j|\tilde{x}_{i+1}) = \tilde{p}(\tilde{x}_{i+1}|\tilde{x}_0)\tilde{p}(\tilde{x}_0)$, which can be done in parallel;

4. for $m = 1, ..., N_d$ do

5. Sample new $\tilde{x}$ via $p(\tilde{x}|\tilde{x}_{i+1})$;

6. Set new sample $x_{i+n} = \tilde{x}$;

7. end

8. Update counter $n = n + N_d$;

9. end

Properties of multiple proposal M-H

- Satisfies detailed balance, thus updates leave the stationary distribution invariant.
- The proposal kernel $k$ can be chosen freely, including formulations based on Langevin diffusion and Hamiltonian dynamics.
- The likelihood of the proposed states can be computed in parallel.
- Due to the multiple alternatives to accept the next state, the acceptance rate is increased.

Completely uniformly distributed numbers

The star discrepancy $D^d_n$ for a set of points $P = \{x_1, ..., x_n\} \subset [0, 1]^d$ is defined by

$$D^d_n = \sup_{b \in [0,1]^d} \left| \frac{\# \{x \in \{0,1\}^d \cap P \cap (0,b]\}}{n} - \text{vol}([0,b]) \right|.$$ 

A sequence $(u_i)_{i \geq 1} \subset [0, 1)$ is called c.u.d. if for any $d \geq 1$ the points $x_i = (u_{i,1}, ..., u_{i,d}) \in [0,1]^d$ fulfill $D^d_n(x_1, ..., x_n) \to 0$ as $n \to \infty$.

References
