Low discrepancy sequences

Low-discrepancy sequences (often called *quasi-Monte Carlo (QMC) sequences*) are deterministic sequences of points in $[0,1]^d$ designed to approximate the uniform distribution as evenly as possible. Their defining property is a small *discrepancy*, which measures the maximal deviation between the empirical distribution of the first N points and the Lebesgue measure, typically tested on axis-aligned boxes.

Standard (star) discrepancy of $X_N = \{x_1, \dots, x_N\} \subset S = \prod_{k=1}^{d} [m_k, M_k]$ is defined as

$D_N^* = \sup_{B \in \mathcal{B}_{\text{axis}}} \left| \frac{\#(X_N \cap B)}{N} - \lambda_d(B) \right|,$

where $\mathcal{B}_{\text{axis}}$ is the class of axis-aligned boxes $\prod_{k=1}^{d} [m_k, t_k),\ t_k \in [m_k, M_k]$

For this star discrepancy D_N^* , low-discrepancy sequences achieve rates of order $O(\log^d N / N))$, which is asymptotically much better than the $O(N^{-1/2})$ behaviour of independent Monte Carlo sampling. This improvement underlies their effectiveness for numerical integration of functions of bounded variation, formalised by the Koksma–Hlawka inequality, which bounds the integration error by the product of the discrepancy of the point set and the variation of the integrand.

Mathematically, most widely used low-discrepancy sequences are *digital sequences* constructed in a fixed base (usually base 2) via arithmetic over finite fields. Examples include Halton sequences (based on radical inverse functions in prime bases), Sobol' sequences (using carefully chosen direction numbers and digital nets), and Faure or Niederreiter–Xing sequences. Their construction ensures that, for every N, the initial prefix of the sequence fills dyadic (or, more generally, digital) partitions of the unit cube in a highly balanced way. A key practical feature is *extensibility*: the sequence can be generated incrementally, and every prefix remains a good low-discrepancy point set, without having to recompute earlier points.

Low-discrepancy sequences are particularly useful for high-accuracy numerical integration and sampling when the integrand is reasonably smooth and depends on many variables. In such settings, quasi-Monte Carlo methods often outperform standard Monte Carlo by orders of magnitude, especially in moderate dimensions, while remaining simple to implement and fully deterministic. Their limitations are equally important: the discrepancy notion is inherently anisotropic (based on axis-aligned boxes), they are not adaptive to local features of a specific integrand, and they are not rotationally invariant. Consequently, in practice they are often combined with randomisation (scrambling, random shifts) or with adaptive or problem-specific refinements, yielding hybrid methods that retain the global uniformity of low-discrepancy sampling while better exploiting the local structure of the integrand.

Low discrepancy vs. coordinate energy minimisation

It is interesting to compare the low-discrepancy sequences with the point sequences produced by minimising the energy based on a specific coordinate product form of the pair potential:

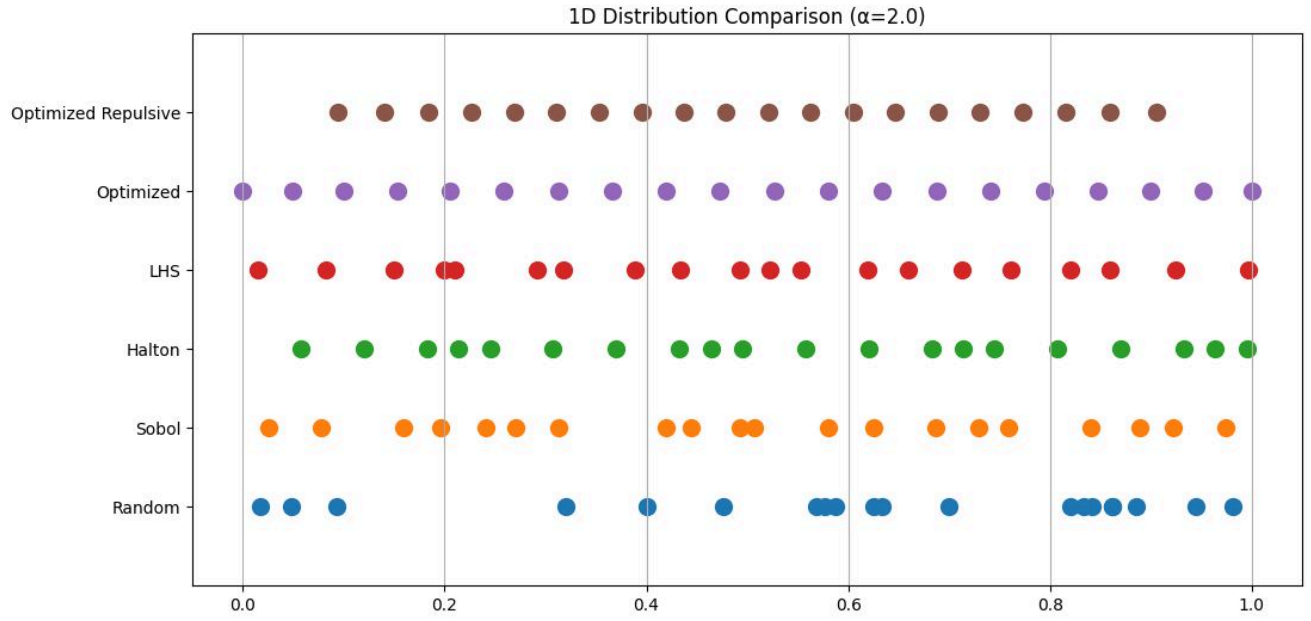
$E = \sum_{i \neq j} \prod_{k=1}^{d} |x_i(k) - x_j(k)|^{-\alpha}$, where $x_i = (x_i(1), \ldots, x_i(d))$.

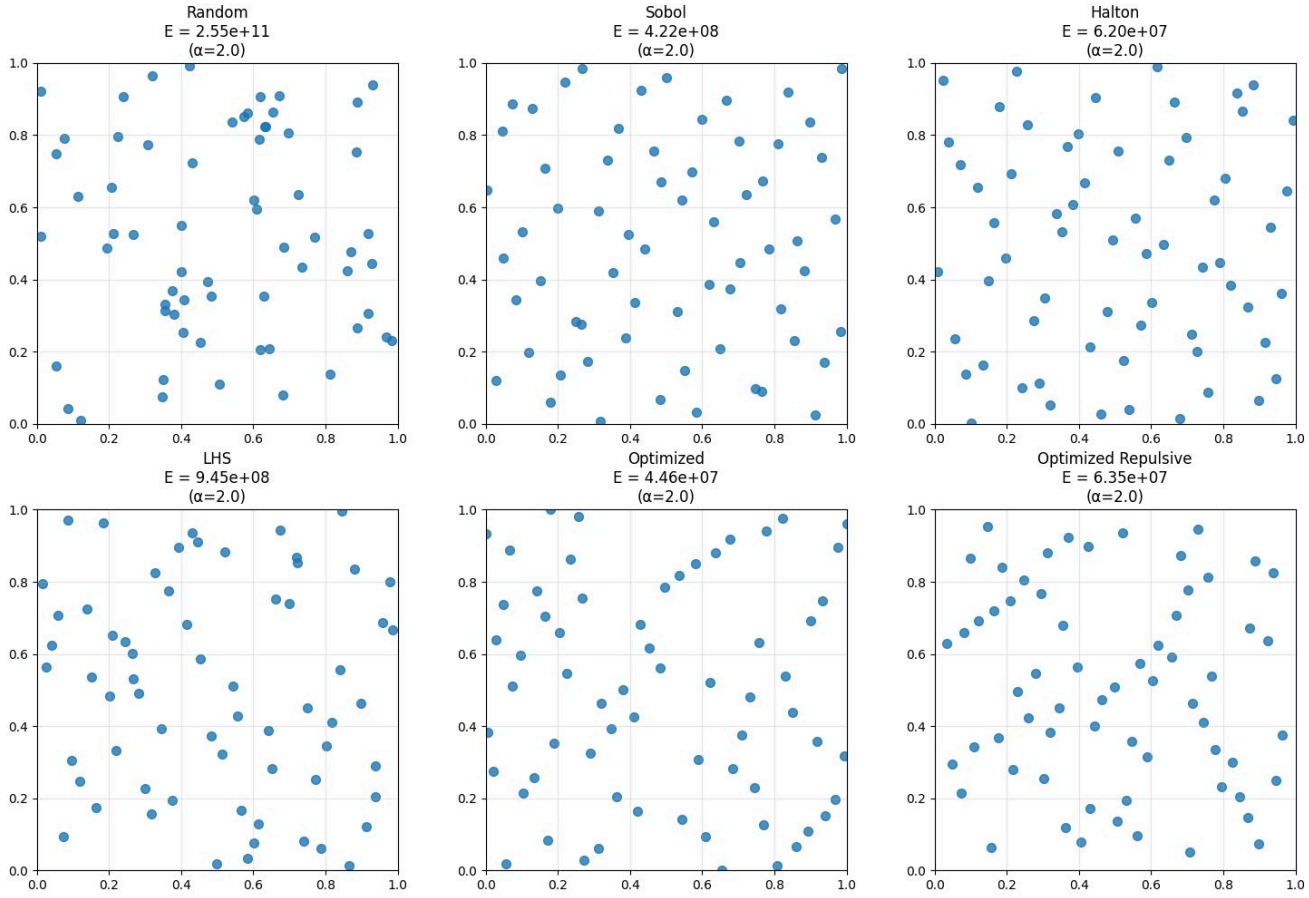
One can also add an additional potential term to repulse the points from the boundary:

$N \sum_i \prod_k (x_i(k) - m_k)^{-\alpha} (M_k - x_i(k))^{-\alpha}$,

where the domain is a box $S = \prod_k [m_k, M_k]$. The factor N is chosen to balance the pressure from the pair potential that behaves like $O(N^2)$.

The following two plots are the energy-minimising configurations for E as above with and without the boundary repulsion term in 1D and in 2D cases.





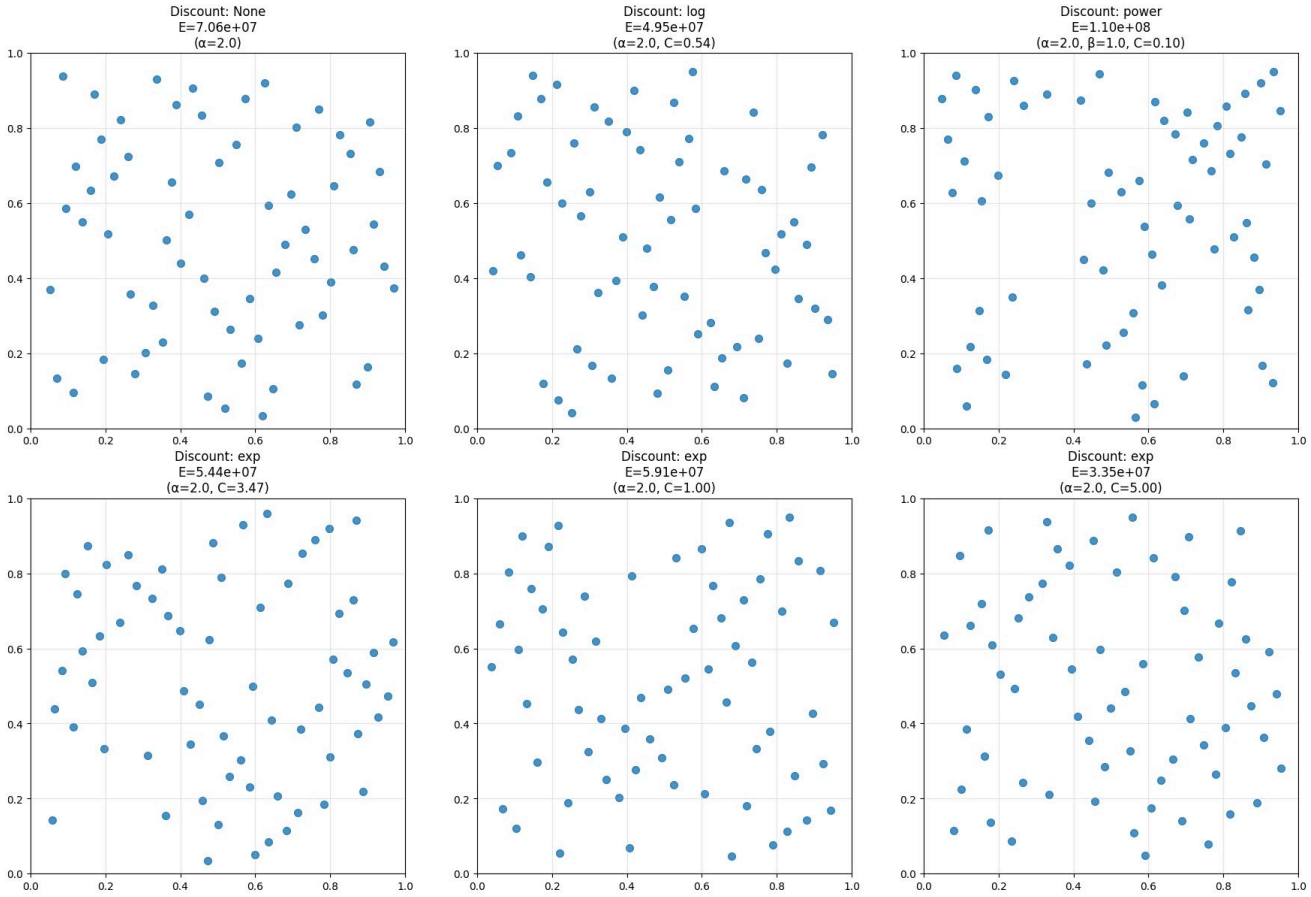
Visibly, a problem with the potential E is that it favours the configurations of points lying along the diagonals. Once such a diagonal pattern appears, it prevents the points to appear in a cross-shaped region: below or above and to the left or to the right the pattern. Another problem is an unbounded interaction: no matter how far two points, their interaction is large if at least just one their coordinates is close. Therefore it has a sense to down-weight the pair potential with a discount term that decays with the distance. For compatibility, an L_1-distance scaled by the phase space box sizes is suggested:
$||z|| = \sum_{k=1}^{d} |z(k)|/(M_k - m_k)$.

Discounted Potentials

Comparison of Optimised configurations using distance-discounted potentials:

1. **Exponential**: $e^{-C||z||}$
2. **Power Law**: $C||z||^{-\beta}$
3. **Logarithmic**: $C/\ln(e + ||z||)$

where $||z||$ is the sum of coordinate distances normalised by box size.



There does not seem a notable improvement with a distance discounted kernel. Perhaps, fast decaying exponential decay produces more promising configurations, but as to me, Sobol or Halton QMC wins.