

Halftoning and Quasi-Monte Carlo

Kenneth M. Hanson

Los Alamos National Laboratory

MS D413, CCS-2, Los Alamos, NM 87545

Phone: 505-667-1402 Fax: 505-665-4972

Email: kmh@lanl.gov

In performing statistical studies in simulation science, it is typically required that one estimate the values of an integral of a function that is based on a simulation calculation. Oftentimes simulations tend to be costly in terms of computer time and usage. Therefore, the ability to accurately estimate integrals with a minimum number of function evaluations is a critical issue in evaluating simulation codes.

The goal of the standard Monte Carlo (MC) technique is to estimate the integral of a function over a specified M-dimensional domain from evaluations of the function at points that are randomly chosen within that domain. The objective in Quasi-Monte Carlo (QMC) is to improve those estimates through a suitable specification of the sample point set. It has been shown that the rms integration errors from N samples for a fixed number of dimensions typically fall off as N^{-1} with QMC, much more quickly than with MC, namely, $N^{-1/2}$.

Digital halftoning is the process of creating a pattern of black dots on a white background to create the illusion of a gray-scale image. One of the principal goals in halftoning is to avoid introducing undesirable texture into the rendered image, which is typically caused by clumping of the dots, or uneven dot placement that accompanies random dot distributions. In a sense, QMC has the same goal, whether it is implicitly or explicitly stated. The clumpiness in random point distributions also exists in standard Monte Carlo, and leads to lower sampling efficiency than more uniformly distributed point distributions.

I will outline the parallels between QMC and halftoning, and describe a halftoning-inspired algorithm for generating a sample set with uniform density, which yields smaller integration errors in two dimensions than standard QMC algorithms. It must be kept in mind that the same algorithms that work for 2D may not work for higher dimensions. I will discuss the implications for higher dimensions and other potential approaches to enhanced QMC methods.

In standard Monte Carlo techniques, one evaluates integrals on the basis of a set of point samples. The estimate of the integral of a function $f(x)$ of the parameter vector x is proportional to a sum over the values of the function evaluated at N sample points x_i , which are randomly drawn from a uniform probability density function defined over the integration region. The objective of the quasi-Monte Carlo technique is to reduce the number of function evaluations needed to obtain a given accuracy in Monte Carlo-type integration, and to accelerate its convergence as N increases. This goal is typically achieved. One useful feature of QMC is that any number of samples can be generated.

Furthermore, an arbitrary number of additional samples can be added to an existing set of samples.

The direct-binary-search (DBS) technique represents one of the best approaches to producing high-quality digital halftoned images. It is based on a simple model for the human visual system in which the image seen by a human eye is simply a blurred version of the actual external scene. The DBS algorithm attempts to minimize the difference between the original gray-scale image to be rendered and the dot image, as perceived by the human. To quantify the perceived discrepancy between the halftone image and the actual gray-scale image, the most-often-used cost function is the total power in the error image. The goal of the DBS algorithm is then to minimize the mean-squared value of the difference between the two blurred images.

Taking a cue from the DBS algorithm, the proposed algorithm for generating quasi-MC points attempts to minimize the visual discrepancy (MVD) between a set of points and an image with uniform density. Starting with some arbitrary point pattern, the MVD algorithm considers each point in the set in a randomly permuted order. Each point is perturbed in eight different directions, parallel to the axes and along the diagonals, and the point is placed in the position with the lowest value of the objective function. This procedure is repeated until all points are visited a number of times. This algorithm produces point set patterns with very desirable halftoning properties, which are essentially indistinguishable from DBS patterns. More importantly, when used to evaluate a variety of integrals in the Monte Carlo manner, estimates of the integrals are more accurate than those obtained with typical quasi-MC point sets. For example, one of the test functions used is $\exp(-2|x - x_0|) \exp(-2|y - y_0|)$, where the accuracy of its integral over the unit square is averaged over positions (x_0, y_0) distributed uniformly over the unit square. For $N = 100$ sample points, the rms relative accuracy of the integrals of this function are 8%, 1.8%, and 0.8%, for a random points, a Halton sequence (a typical QMC point set), and the proposed MVD point set, respectively. The same advantage of the MVD point sets is observed for other values of N . Its accuracy relative to standard randomly-chosen points steadily improves as N increases; at $N = 1000$, it is over 20 times more accurate. Translated into the number of function evaluations needed, the use of random points would require over 400 times more function evaluations than MVD to achieve comparable integration accuracy.

The algorithm used for the preceding evaluations in two dimensions is not appropriate for even moderately high dimensions. The reason is that it is based on a pixilated version of the point image. This implies the necessity for storing a discretized image in M dimensions, which may be infeasible when M gets larger than four or five, even when coarsely discretized in each dimension.

Another approach to generating a suitable point set, which can be shown to be equivalent to MVD, is to draw an analogy between the point set and a collection of particles, which interact by means of a potential field. The potential field can be chosen so that the particles repel each other at close distances, but are less repulsive when they are sufficiently far apart. This type of action occurs in the MVD approach, although it is not

explicit. Appropriate conditions need to be specified at the boundary of the region. The advantage of this potential-field approach is that an integral over the M-dimensional domain is not required to evaluate the cost function.

Initial results indicate that this potential-field approach is very promising. In two dimensions, point sets generated with this approach produce integration accuracies that are comparable to those obtained with MVD. Similar advantages are observed in higher dimensions.