Speedier Simulations with Quasi-Monte Carlo Methods

Luke Bielawski[1], Brooke Feinberg[2], Aiwen Li[3], Richard Varela[4]. Advisors: Dr. Fred Hickernell[5], Dr. Yuhan Ding[5]

[1]University of Illinois Urbana-Champaign [2]Scripps College [3]University of Pennsylvania [4]California State University, Sacramento
[5]Illinois Institute of Technology

Background and Motivation

Various problems can be written in the form of a multidimensional integral taken over the d-dimensional unit cube, which can be approximated by point evaluations

$$\int_{[0,1]^d} f(\mathbf{x})\,d\mathbf{x} \approx \frac{1}{n}\sum_{i=1}^{n} f(\mathbf{x}_i) \tag{1}$$

Commonly, the point set $P_n = \{\mathbf{x}_i\}_{i=1}^n \subset [0,1]^d$ is made of independent elements distributed uniformly over the unit cube, which forms the basis of Monte Carlo simulation. Specially constructed point sets which cover the unit cube more evenly, known as low discrepancy sequences, can be proven to exhibit faster rates of convergence for (1), in particular Sobol' sequences [1]. We have primarily focused on applications to studying the expectation and density of random variables of the form

$$Y = f(X), \qquad X \sim \mathscr{U}([0,1]^d) \tag{2}$$

In this project, we explored different uses of low discrepancy sequences and other QMC methods in various areas of mathematics, finance, and statistics.

Error Analysis With QMC

Goal: Minimize error when estimating the PDF $\hat{\rho}(y)$ of a random variable.
Method: Kernel Density Estimation (KDE), a non-parametric method for estimating PDFs [7]. We tested several parameters to find ones with the smallest error:

- Kernel = weighting function. We tested the Epanechnikov and Gaussian kernels [6].
- Bandwidth = smoothing parameter, denoted h. We tested 100 different bandwidths $\{h_i\} = 2^{\{a_i\}}$ where $\{a_i\}$ is the set of 100 evenly-spaced points in [-10, 0].
- Sample size = number of points n. We tested values of $n \in \{2^{10}, 2^{11}, 2^{12}, 2^{13}, 2^{14}\}$.
- IID vs. LD: We compared the accuracy of KDEs when using Sobol' points, a type of low discrepancy (LD) sequence, to KDEs using IID points.
- Distributions: We tested the accuracies of KDEs for various distributions in both 1 dimension and higher dimensions.

To measure error, we used the root mean squared error (RMSE), defined as

$$\text{RMSE} := \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{\rho}(y_i) - \rho(y_i))^2} \tag{3}$$

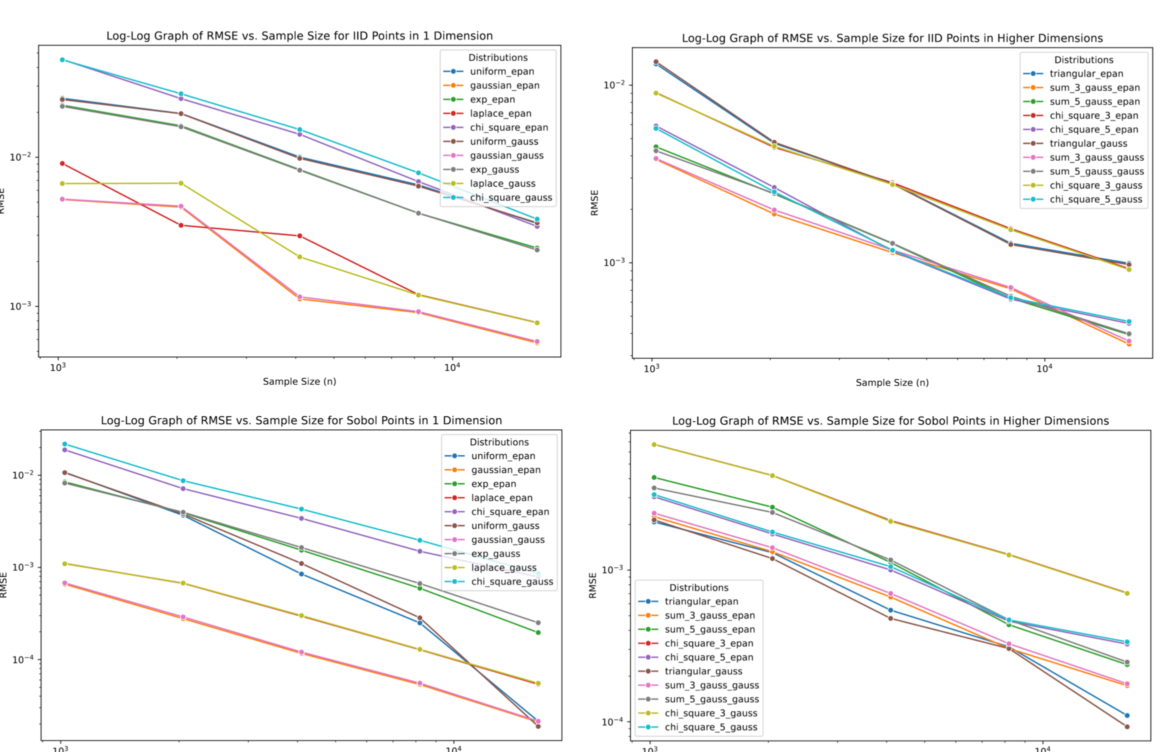


Figure 1. Log-log graphs of RMSE vs. sample size (n). We graphed linear regression models of the form $-\log_{10}(RMSE) = b_0 \cdot \log_{10}(n) + \sum_{j=1}^{N} b_j \cdot x_j + \varepsilon_j$ where b_0 is the slope of $-\log_{10}(RMSE)$ vs. $\log_{10}(n)$, and x_j are indicator variables to separate the N different distributions we test. For IID points, slope in 1 dim. is 0.8379 (top left) and in higher dims. is 0.8811 (top right). For Sobol' points, slope in 1 dim. is 1.0750 (bottom left) and in higher dims. is 0.9465 (bottom right).

Conditional Quasi-Monte Carlo

Conditional density estimation (CDE) gives an alternative method for density estimation. For $Y = f(X_1, \ldots, X_d)$, if we can calculate the conditional density for $k \in \{1:d\}$

$$\rho(y \,|\, (X_1, \ldots, \hat{X}_k, \ldots, X_d)) =: \rho(y \,|\, \mathscr{G}_{-k}) \tag{4}$$

then it can be shown [3], under certain technical assumptions, that

$$E[\rho(y \,|\, \mathscr{G}_{-k})]_{\mathscr{G}_{-k}} = \rho(y) \qquad \forall y \in [a,b] \approx \operatorname{supp}(Y) \tag{5}$$

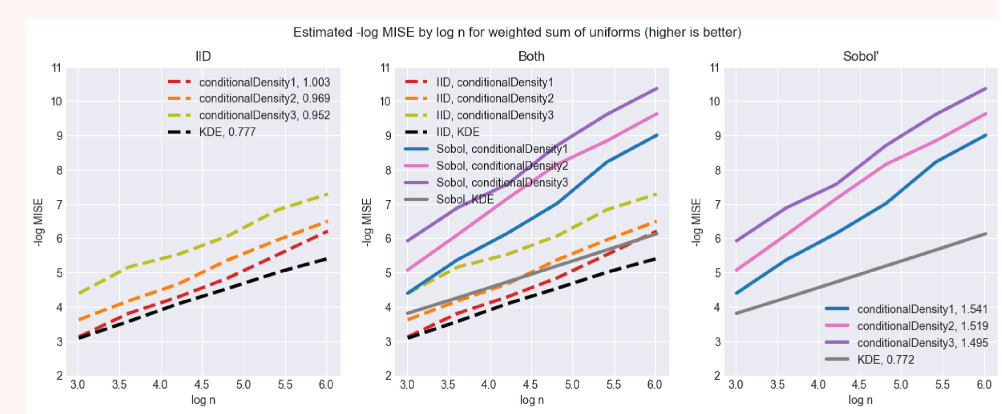
with finite variance. We can utilize RQMC to estimate this density by transforming

$$\rho(y \,|\, \mathscr{G}_{-k}) = \tilde{\rho}(y, |U_{-k}), \qquad U_{-k} \in [0,1]^{d-1} \tag{6}$$

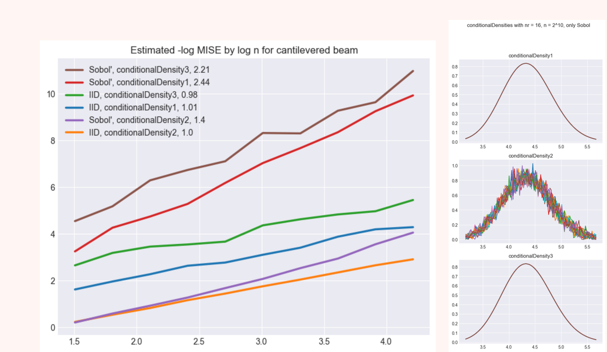
If we have the true PDF of Y, we estimate the MISE by rectangle rule. If the true PDF is not known, we estimate the MISE by: splitting the domain into n_e equal segments, realizing the estimated density $\hat{\rho}(y)$ n_r times, and getting the unbiased estimator

$$\widehat{\text{MISE}} := \frac{b-a}{n_e}\sum_{j=1}^{n_e} \widehat{\operatorname{Var}(\hat{\rho}_n(y_j))} \tag{7}$$

Example 1
$Y = X_1 + X_2 + X_3$
$X_i \sim \mathscr{U}([0, 2^{i-1}])$
$\rho(y \,|\, \mathscr{G}_{-k}) = 2^{1-k}\mathbb{1}_{y - X_1 - X_2 \in [0,2]}$
Sobol' points much better, hiding the higher variance variables is best



Example 2 (borrowed from [3])
$Y = \frac{500000}{X_1}\sqrt{\frac{X_2^2}{256} + \frac{X_3^2}{16}}$, $X_j \sim \mathscr{N}(\mu_j, \sigma_j^2)$
$\mu = (2.9 \cdot 10^7, 500, 1000), \sigma = (1.45 \cdot 10^6, 100, 100)$
No true PDF, hiding 1 or 3 much better than 2



Quasi-Monte Carlo Graph Random Features

Motivating Question: Can we identify an alternative kernel function or graph type that achieves a similarly low variance estimation for the q-QRF procedure as the previously introduced 2-regularized Laplacian kernel?

- The general graph random features (g-GRFs) algorithm estimates graph kernels by employing an ensemble of random walkers that deposit a "load" at every vertex they pass through that depends on: i) the product of weights of edges traversed by the walker, and ii) the marginal probability of the subwalk [4].
- The Quasi-Monte Carlo graph random features algorithm (q-GRFs) introduces correlated ensembles, or antithetic walkers, to correlate the lengths of random walks with some stopping criterion –improving the variance of estimators of the 2-regularised Laplacian kernel [5].

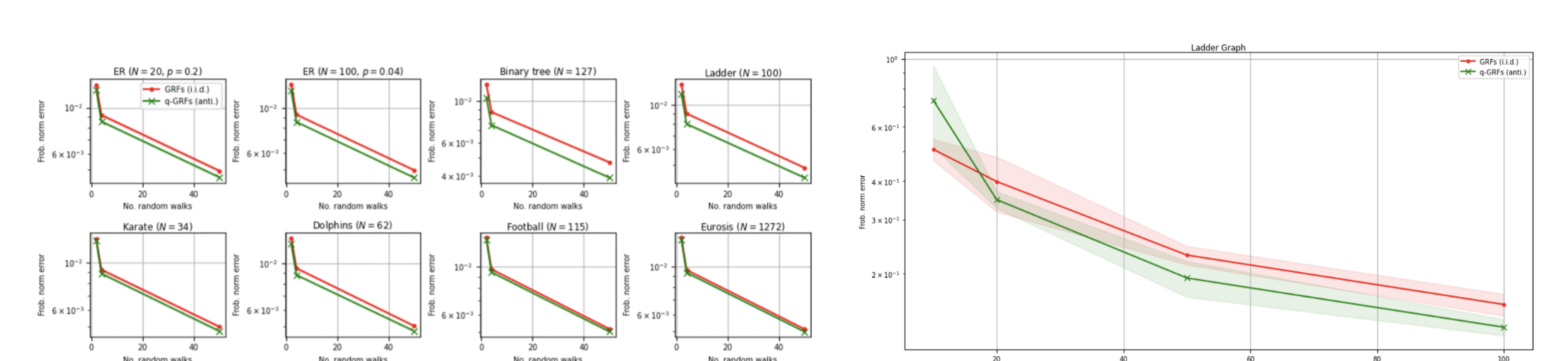


Figure 2. LEFT: Q-GRFs (green line) yield lower variance estimators of the diffusion kernel than g-GRFs (red line) on the eight previously experimented graphs [5]. RIGHT: Q-GRFs yield lower variance estimators of the diffusion kernel on ladder graphs with nine rungs.

Results: While the diffusion kernel yields comparable results to the 2-regularized Laplacian kernel, the antithetic procedure's inherent randomness results in inconsistent outcomes. Therefore, we assert there is strong evidence to suggests the q-GRFs algorithm can yield lower variance estimators of the diffusion kernel. Moreover, we found that the number of rungs on a ladder graph influences whether q-GRFs can achieve a lower variance estimator of the diffusion kernel–though theoretical results explaining these phenomena are forthcoming.

Application of QMC for Finance Options

QMC can be used to model and simulate financial options, via density estimation, to determine the best time to exercise the option. To create and compute these financial options, the software package QMCPy, a Python library which is maintained by Dr. Hickernell and his team, was used to perform the density estimation of the PDF of these different financial payoffs.

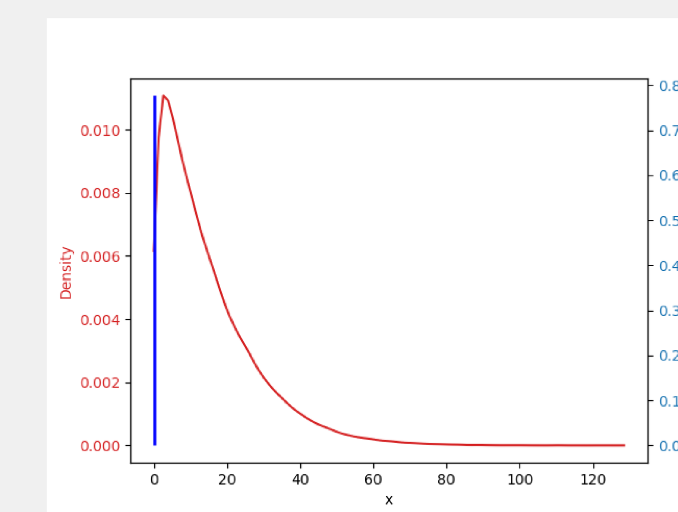


Figure 3. Here is the density estimation of the European for a starting price \$100 and a strike price of \$120, where red curve represents the estimated probability density function, while the blue bar represents the discrete probability of when the payoff of the option is zero.

Conclusion and Future Work

In each case, low discrepancy points can be seen to allow much lower error than IID uniform points can provide. The exact benefit varies and ongoing theoretical work aims to recognize its utility in a variety of situations, including density estimation and analyzing combinatorial objects, and how to transform problems to be more QMC-compatible. Beyond theory, implementing QMC methods into software has been an ongoing goal: building robust, clean, and fast code, such as through QMCPy [2]. Future work expanding on theory, applications, and implementation is anticipated.

References

[1] Amal Ben Abdellah, Pierre L'Ecuyer, Art B. Owen, and Florian Puchhammer. Density estimation by randomized quasi-monte carlo. SIAM/ASA Journal on Uncertainty Quantification, 9(1):280–301, 2021.

[2] S.-C. T. Choi, F. J. Hickernell, M. McCourt, and A. Sorokin. QMCPy: A quasi-Monte Carlo Python library, 2020+.

[3] Pierre L'Ecuyer, Florian Puchhammer, and Amal Ben Abdellah. Monte carlo and quasi–monte carlo density estimation via conditioning. INFORMS Journal on Computing, 34(3):1729–1748, 2022. Publisher: INFORMS.

[4] Isaac Reid, Krzysztof Choromanski, Eli Berger, and Adrian Weller. General graph random features. (arXiv:2310.04859).

[5] Isaac Reid, Krzysztof Choromanski, and Adrian Weller. Quasi-monte carlo graph random features. (arXiv:2305.12470).

[6] Youngsung Soh, Yongsuk Hae, Aamer Mehmood, Raja Hadi Ashraf, and Intaek Kim. Performance evaluation of various functions for kernel density estimation. 03(1):58–64.

[7] Walter Zucchini. APPLIED SMOOTHING TECHNIQUES part 1: Kernel density estimation. 2003.

Acknowledgement

We acknowledge the support and funding provided by NSF-DMS-2244553 through Illinois Tech's (SURE): Summer Undergraduate Research Experience (REU) program