Statistics and Probability Letters 132 (2018) 55–61

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Optimal Jittered Sampling for two points in the unit square

Florian Pausinger [a], Manas Rachh [b], Stefan Steinerberger [c,*]

[a] Zentrum Mathematik (M10), TU Munich, Germany
[b] Applied Mathematics Program, Yale University, New Haven, CT 06510, USA
[c] Department of Mathematics, Yale University, New Haven, CT 06510, USA

ARTICLE INFO

Article history:
Received 20 April 2017
Received in revised form 11 September 2017
Accepted 18 September 2017
Available online 27 September 2017

MSC:
primary 65C05
secondary 93E20

Keywords:
Jittered Sampling
Calculus of variations
Numerical integration
Quasirandom point sets

ABSTRACT

Jittered Sampling is a refinement of the classical Monte Carlo sampling method. Instead of picking n points randomly from $[0, 1]^2$, one partitions the unit square into n regions of equal measure and then chooses a point randomly from each partition. Currently, no good rules for how to partition the space are available. In this paper, we present a solution for the special case of subdividing the unit square by a decreasing function into two regions so as to minimize the expected squared $\mathcal{L}_2$-discrepancy. The optimal partitions are given by a *highly* nonlinear integral equation for which we determine an approximate solution. In particular, there is a break of symmetry and the optimal partition is *not* into two sets of equal measure. We hope this stimulates further interest in the construction of good partitions.

1. Introduction and statement of result

1.1. Jittered Sampling

Jittered Sampling is a mixture of classical Monte Carlo and sampling along grid-type structures: a standard approach is to partition $[0, 1]^2$ into m^2 axis aligned cubes of equal measure and placing a random point inside each of the $N = m^2$ cubes, see Fig. 1. This idea seems to date back to a paper of Bellhouse (1981) from 1981 and makes a reappearance in computer graphics in a 1984 paper of Cook et al. (1984) by the name of *Jittered Sampling* (see also Dobkin et al., 1996). Bounds on the discrepancy are due to Beck (1987); we also refer to the book Beck and Chen (1987), the exposition in Chazelle (2000), and the recent quantified version of the first and third authors (Pausinger and Steinerberger, 2016). Deterministic lower bounds are derived in Chen and Travaglini (2009).

> **Main Problem.** For any $N \in \mathbb{N}$, $d \geq 2$, which partition of the unit cube $[0, 1]^d$ into N sets gives, in expectation, the best result for the Jittered Sampling construction? What partition of space should be used?

We will study this question restricted to $d = 2$. Clearly, 'best' requires a quantitative measure of equidistribution: we will work with the squared $\mathcal{L}_2$-discrepancy, which, for a given set $P = \{p_1, \ldots, p_N\}$ of N points in $[0, 1]^2$ is given by

$$\mathcal{L}_2^2(P) := \int_{[0,1]^2} \left| \frac{\#P \cap ([0, x] \times [0, y])}{\#P} - xy \right|^2 dxdy.$$

* Corresponding author.
E-mail addresses: florian.pausinger@ma.tum.de (F. Pausinger), manas.rachh@yale.edu (M. Rachh), stefan.steinerberger@yale.edu (S. Steinerberger).

https://doi.org/10.1016/j.spl.2017.09.010
0167-7152/

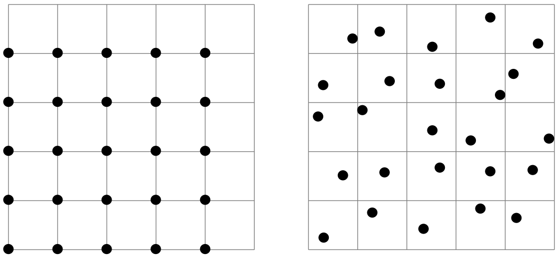

Fig. 1. The regular grid and a point set obtained by Jittered Sampling.

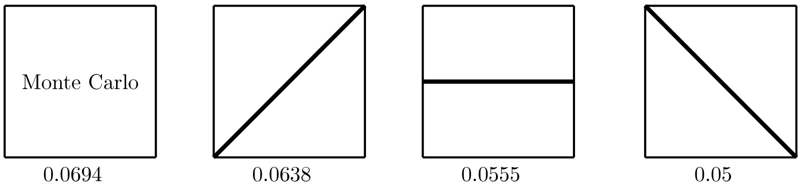

Fig. 2. Different subdivisions and their expected squared $\mathcal{L}^2$-discrepancy.

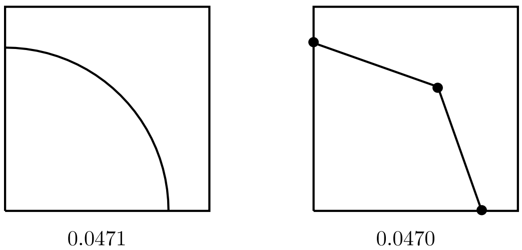

Fig. 3. A quarter disk and lines connecting $(0, 0.792)$, $(0.63, 0.63)$ and $(0.792, 0)$.

A result of the first and third authors (Pausinger and Steinerberger, 2016) shows that *any* decomposition into N sets of equal measure *always* yields a smaller expected squared $\mathcal{L}_2$-discrepancy than N completely randomly chosen points: even the most primitive Jittered Sampling construction is better than Monte Carlo. However, currently known quantitative bounds (Pausinger and Steinerberger, 2016) do not imply any effective improvement for $N \lesssim (2d)^{2d}$ points. The motivation of our paper is to gain a better understanding of Jittered Sampling. A better quantitative control is desirable since it could provide a possible way towards improving bounds on the inverse of the star-discrepancy (see Aistleitner, 2011; Dick and Pillichshammer, 2014; Heinrich et al., 2001).

1.2. Result

The purpose of this short note is to initiate the study of explicit effective Jittered Sampling constructions by obtaining a complete solution for the $d = 2, n = 2$ case within a natural family of domain partitions. We first consider some natural examples (see Fig. 2) of partitions into sets of equal measure and compute their expected squared $\mathcal{L}_2$-discrepancy (for details on the computation, see the proof).

These examples suggest that a certain type of symmetry along the $y = x$ diagonal seems to be helpful. Of course, this still leaves a very large number of shapes that could potentially be tested. Two natural examples are a quarter disk and a quadrilateral domain; see Fig. 3.

We will now restrict ourselves to the study of partitions of $[0, 1]^2 = \Omega \cup \left([0, 1]^2 \setminus \Omega\right)$, where

$$\Omega = \left\{(x, y) \in [0, 1]^2 : y \leq g(x)\right\}$$

for some monotonically decreasing function g whose graph $\{(x, g(x)) : 0 \leq x \leq 1\}$ is assumed to be symmetric around the line $y = x$ and splits the unit square into two regions with areas p and $1 - p$. Somewhat to our surprise, it is actually possible to determine a highly nonlinear integral equation that any optimal function $g(x)$ has to satisfy. The equation is so nonlinear that we would not know of any way to show existence of solution, except that it arises as the minimum of a variational problem for which compactness methods can be used.