A Large Scale Extended Algorithm for 2D Halton Points with Low-Discrepancy Sequences

Wenxing Chen*
School of Mathematics and Statistics
Wuhan University
Wuhan, 430072 China
E-mail: wenxingchen@whu.edu.cn

Shuyang Dai
School of Mathematics and Statistics
Wuhan University
Wuhan, 430072 China
E-mail: shuyang_dai@whu.edu.cn

Baojuan Zheng
Department of Algorithm Research and Development
Commercial Vehicle Technology Center of SAIC Motor Group Co., Ltd
Shanghai, 200433 China
E-mail: Zhengbaojun2020@163

***Abstract*–Random discrete points have important application value in meshless PDE equation discretization, molecular dynamics simulation, point cloud imaging and so on. There are many common methods to generate random points, such as Monte Carlo, Gibbs Sampling, and Hammersley series and etc. But Halton random point algorithm has a defect that it only generates discrete points in $[0,1]^2$ region. However, in practical applications, it is necessary to be able to generate discrete points on any area. This paper proposed a new Halton points extension algorithm to solve this defects. We defined a linear operator which can transform discrete points from $[0,1]^2$ region into any plane region. Two examples are given, the extension algorithm respectively includes square, rectangular and polar coordinates region. The numerical results show that our method is accurate, effective and more general, it also enhanced the application range by our method.**

***Keywords*–sampling strategy, halton points, extended linear operators, linear transformation, low-discrepancy sequences.**

I. INTRODUCTION

The Halton random sampling problem has important application value in scientific calculation [1,2], image processing [3], and engineering measurement [4]. This Halton extension algorithm can also be applied to SPH to solve fluid and dynamic problems by a meshless method. The use of coordinate measuring machine (CMM) for size measurement has been applied to advanced manufacturing environment to ensure the high quality and reliability of manufacturing products [5]. In addition, the use of meta-heuristic method to solve nonlinear engineering and optimization problems is also one of the important topics [6]. A new hybrid Halton-PSO algorithm without changing the original particle swarm algorithm and its structure. The results show that the Halton-PSO algorithm improves the performance and efficiency [7]. In general, given a probability distribution $\rho(x)$, how to let the computer generate samples satisfying this probability distribution. This problem is an important in statistical simulation-sampling. Sampling forms are also diverse, such as Monte Carlo (MCMC) [8], Gibbs Sampling [9], Halton, etc. It is well known that Halton sequences show a poor uniformity in high dimensions. Literature [10] combines the potential accuracy advantage of Halton sequence in multi-dimensional integration with the practical error estimation advantage of Monte Carlo methods.

In view of the limitations of general random sampling, the principle and method of sampling based on Halton point extension algorithm are proposed in this paper, the sampling strategy is given, including the division of geometric region, and the calculation principle of sampling point coordinates. We design extension algorithm needs to satisfy three conditions.

(1) The spacing between the sampling points uniform. (2) The properties of the expanded discrete point and the principle discrete point remain unchanged. (3) The generation speed is fast and the method is simple. When the number of sampling points increases on the plane, the Halton sampling method has the minimum rate value and covariance cov value, which reflects that the Halton sampling method has the best uniformity at the same sampling point.

II. HALTON POINTS EXTENSION ALGORITHM

In the PDE calculation, the size of solving domain is various, but traditional Halton random points can only generation the range of $[0,1]^2$. Therefore, this article will give a specific extension algorithm to make the discrete domian not limited. Halton algorithm can generate a set of points based on the Vander Corput sequence which are uniform distributed in high-dimensional space [10]. For all algorithms that need to be sampled, uniformly distributed random numbers lead to a better sample distribution. The uniform distribution in high-dimensional space can be judged according to the discrepancy criteria, Halton sequences are generated from radical inversion calculations.

As for the following formula Eq.(1), b is a positive integer set, as for any positive integer n can be converted to a prime b-scale system. Namely, the decomposition the integer n into sum of n power functions. Such as Eq.(2), the coefficients $a_i(j,n)$, $(0 \le a_i \le b_j)$ includes the summation of the finite term function coming from the digit expansion of the integer n in base b_j , $\Phi_{b,c(j)}$ is an element in the n-dimensional Halton coordinates matrix and the specific expression is given as follows:

$$n = \sum_{i=0}^{M-1} a_i(j,n) b_j^i \tag{1}$$

$$\Phi_{b,c(n)} = (b^{-1},...,b^{-M})[C(a_0(j,n),a_1(j,n),...,a_{M-1}(j,n))^{\mathrm{T}}] \tag{2}$$

978-1-6654-1732-7/21/$31.00 ©2021 IEEE
DOI 10.1109/AIAM54119.2021.00043

2021 3rd International Conference on Artificial Intelligence and Advanced Manufacture (AIAM) | 978-1-6654-1732-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/AIAM54119.2021.00043

If C is the identity matrix that can directly put the number with base b into the right of the decimal point, and the Vander Corput Sequence is defined as:

$$\Phi_{b_j(n)} = (b^{-1},...,b^{-M})(a_0(j,n),a_1(j,n) \\ ,...,a_{M-1}(j,n))^{\mathrm{T}} = \sum_{i=0}^{M-1}\sigma(a_i(j,n)b^{-i-1}) \tag{3}$$

The Halton sequence in relatively prime basis $b_1,b_2,...,b_s$ which defined as follows:

$$H_{b,s} = (\Phi_{b1}(n),...,\Phi_{bj}(n),...,\Phi_{bs}(n)) \tag{4}$$

Each dimension is based on Vander Corput Sequences with different bases b_n, which are prime numbers $(b_1, b_2,...,b_n)$. Finally, the Halton sequence is expressed as $\boldsymbol{H}_{b,s}$, but the defect of this sequence is that only generated some discrete points on $[0,1]^s$ domain. This paper proposed a new extension algorithm can generate arbitrary points in 2D plane by scaling transformation and translation transformation of coordinate points.

A. The Square Extension Algorithm

Recently, matlab can only generate Halton points in the region of $[0,1]^n$, but it doesn't satisfy the actual need for PDE discretization. Therefore, this algorithm needs to be improved. The coordinates of any point in the 2D domain are $(x_i^{h_0}, y_i^{h_0})$, the side length of the original square region Ω_0 can be denoted as L_0. Meanwhile, L is the length of extension square region Ω_s, the coordinates (x_i^s, y_i^s) of any point on the extension region Ω_s. The coordinates transformation formula from Ω_0 to Ω_s which can be written as Eq.(5). Among them, α is a Linear transformation factor which value is $\alpha = L / L_0$.

$$\begin{pmatrix} x_i^s \\ y_i^s \end{pmatrix} = \alpha \begin{pmatrix} x_i^{h_0} \\ y_i^{h_0} \end{pmatrix} \tag{5}$$

B. The Rectangular Region Extension Algorithm

In the first quadrant of the Cartesian Coordinate System, a and b are respectively the length and the width of the extended rectangular region Ω_I. Meanwhile, two sides a and b of a rectangle is very close to the x-axis and the y-axis. According to the above description, the following relationships formula (6) exist between regions:

$$\Omega_I = \Omega_{A_i} \bigcup \Omega_{A_r}, i = 1,2,...,k. \tag{6}$$

Ω_{A_i} is obtained by horizontal translation of square A_1. The discrete points of the remainder region Ω_{A_r} can be obtained by translation of A_0. However, the discrete points of the first square A_1 is obtained by the linear transformation of Ω_0, expansion coefficient is $\alpha = min\{a,b\}$, Ω_{A_i} is composed of k equal-sized squares, and satisfy the relationship $\Omega_{A_i} = \bigcup_{i=1}^{k} A_i$, $\bigcap_{i=1}^{k} A_i = 0$. Besides, the size $S_{A_1} = S_{A_2} =,...,= S_{A_k}$. The Halton discrete points in the sub-region A_1 is denoted as $(x_i^{A_1}, y_j^{A_1})$. The extension formula of Halton points in region Ω_{A_i}, the specific translation formula is shown in Eq.(7).

$$\begin{pmatrix} x_i^{A_i} \\ y_j^{A_i} \end{pmatrix} = \alpha \begin{pmatrix} x_i^{A_1} + (j-1)\alpha \\ y_j^{A_1} \end{pmatrix}, i = 1,2,...,n_0, j = 1,2,...,k. \tag{7}$$

As for the Ω_d, let $m = b\%a$, if $m = 0$, then the region Ω_d is not exist. If $m \neq 0$, then $\Omega_d = [b - ka, a]^2$. The Halton discrete point $(x_i^{A_r}, y_j^{A_r})$ in Ω_{A_r}, which can be obtained by translating the coordinate point $(x_i^{A_0}, y_j^{A_0})$ in A_0, $A_0 \subset A_1$, the discrete point coordinates $(x_i^{A_r}, y_j^{A_r})$ formula of Ω_{A_r} can be written as :

$$\begin{pmatrix} x_i^{A_r} \\ y_j^{A_r} \end{pmatrix} = \alpha \begin{pmatrix} x_i^{A_0} + k\,\alpha \\ y_j^{A_0} \end{pmatrix} \quad i, j = 1,2,...,m_0 \tag{8}$$

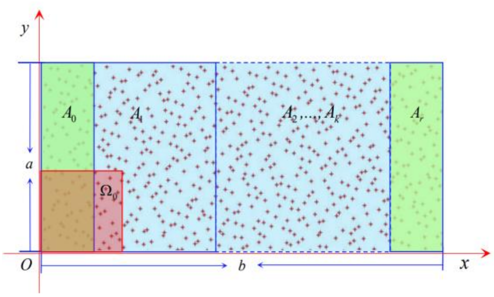
Then, all coordinates (x'_i, y'_i) of extension Halton points are formed by combining $(x_i^{A_i}, y_j^{A_i})$ and $(x_i^{A_r}, y_j^{A_r})$. In order to get an extended algorithm for a more general position. Firstly, we need to obtain the discrete points coordinates (x'_i, y'_i) of the extended region in the first quadrant. Then using the translation formula to make the Ω_I restoration the really position Ω, and the coordinates translation formula is as follows.

$$\begin{pmatrix} x_i \\ y_i \end{pmatrix} = \begin{pmatrix} x'_i + m \\ x'_j + n \end{pmatrix} \quad i, j = 1,2,...,kn_0,...,kn_0 + m_0 \tag{9}$$

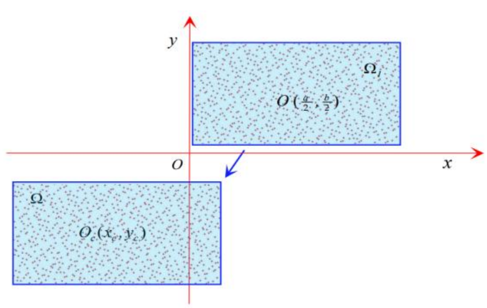
In Eq(9), the translation parameters m and n corresponding to the x and y directions. For the known region Ω_I and Ω_s center point coordinates are respectively $(\frac{a}{2}, \frac{b}{2})$ and (x_c, y_c). Actually, the extension rectangular region can be exist any quadrants in four quadrants. The translation parameters m, n can be written in the form of piecewise function. Different quadrants correspond to different shifts, and the four cases are summarized as follows:

$$(m,n)=\begin{cases}(\frac{-a+2x_c}{2},\frac{-b+2y_c}{2}) & x_c\ge\frac{a}{2},y_c\ge\frac{b}{2}\\ (\frac{a-2x_c}{2},\frac{-b+2y_c}{2}) & x_c\le\frac{a}{2},y_c\ge\frac{b}{2}\\ (\frac{a-2x_c}{2},\frac{b-2y_c}{2}) & x_c\le\frac{a}{2},y_c\le\frac{b}{2}\\ (\frac{-a+2x_c}{2},\frac{b-2y_c}{2}) & x_c\ge\frac{a}{2},y_c\le\frac{b}{2}\end{cases} \quad (10)$$

Finally, the construction process and solution steps of the Halton point extended algorithm are shown in Fig1. The Halton discrete points are extanded from $[0,1]^2$ to another rectangular area. Namely, from the red area Ω_0 to the light blue area Ω_{A_r}, the entire area Ω_I is equal to k blue square areas plus the remaining green rectangles Ω_{A_r}. Actually, the position of PDE solution region is not limited in the first quadrant, they may be exist any position. So we need to expand Ω_I first and then to move Ω_I to actual location Ω.



(a)Rectangular extension algorithm result graph.



(b)Rectangular region's position restoration.

Fig.1. llustrative diagram of Halton point extension algorithm.

III. Error analysis

The discrepancy of high dimensional Halton sequence, and its lower bound criterion satisfies the following relation Theorem1. It mainly describes the generalized Halton sequence which gives an estimate of discrepancy of the generalized Halton sequence in bounded Cantor bases. The details of Theorem1 and Theorem2 are as follows:

Theorem 1: Let $(H_s(n))_{n\ge1}$ be an s-dimensional generalized Halton's sequence. Let D_N^* be the discrepancy of the sequence $(H_s(n))_{n=1}^N$. There exists a constant $C(H_s)>0$, such that.

$$\max_{1\le M\le N} MD_M^* \ge C(H_s)\log_2^s N \qquad N=2,3,... \quad (11)$$

Theorem 2: Let $b_1=(b_{1,j})_{j=1}^\infty,...,b_s=(b_{s,j})_{j=1}^\infty$, and s is an bounded sequences of natural numbers greater than 1. such that, for all $1\le i_1\le i_2\le s$ and all $j_1,j_2\in N$, b_{i_1,j_1} and b_{i_2,j_2} are coprime. For each $1\le i\le s$, denote $\sum_i=(\sigma_{i,j})_{j=1}^\infty$. Suppose that w is the generalized Halton sequence in Cantor bases $b_1,b_2,...,b_s$ with respect to $\sum_1,...,\sum_S$. Then, for any $N\in\boldsymbol{N}$, we have

$$ND_N^*(w)\le\sum_{l=0}^{s}\frac{M_{l+1}}{l!}\prod_{i=1}^{l}(\frac{\lfloor M_i/2\rfloor \log N}{\log m_i}+l) \quad (12)$$

Where $M_i=\max(b_{i,j})_{j=1}^\infty$ and $m_i=\min(b_{i,j})_{j=1}^\infty$.

$(1\le i\le s)$ for any $N\in\boldsymbol{N}$, we obtain:

$$D_N^*(w)\le c\frac{(\log N)^s}{N}+O(\frac{(\log N)^{s-1}}{N}) \quad (13)$$

$$c=c(b_1,...,b_s)=\frac{1}{s!}\prod_{i=1}^{s}\frac{\lfloor M_i/2\rfloor}{\log m_i} \quad (14)$$

IV. Numerical Example

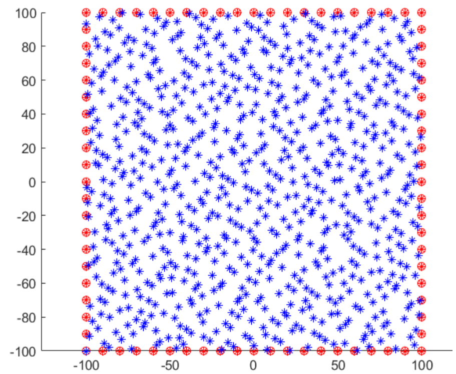
This Section will give two numerical examples. The first example is mainly divided into the Halton point expansion algorithm for square and rectangular areas. The second example mainly introduces the Halton point extension algorithm with two polar coordinate areas. The extension of discrete points can be completed efficiently, and the density and uniformity between the extended points can be kept unchanged.

A. Example 1

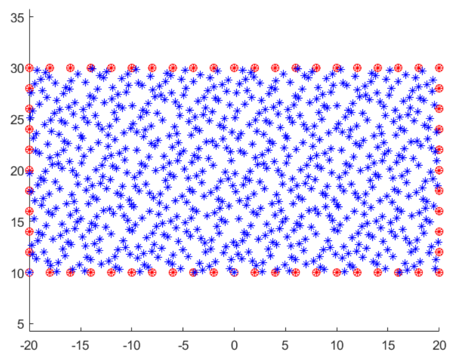
Example 1 belongs to the extended algorithm of the square region and rectangular area. The discrete points are generated by the Vander Corput Sequence. As for the square extended algorithm, the results are shown in Fig2(a). a is the coefficient of scaling transformation from the original region $\Omega_0=[0,1]^2$ to extension square Ω. The number of discrete points still unchanged, but the density of discrete points is reduced to the original $1/a^2$. Fig2(a) contain 2000 Halton points, we extended the Halton points from $[0,1]^2$ region to the $[-100,100]^2$.

As for rectangular region extension algorithm. Firstly, the Halton region $\Omega_0=[0,1]^2$ is extended to the rectangular discrete region Ω_{A_1} in the first quadrant, and then the Ω_{A_1} is moved from the first quadrant to the final actual position Ω. If the length-width ratio of the rectangular region is not integer times, the discrete point of the remaining parts needs to be considered filling problem. This example extended region is $\Omega = [-20,22] \times [10,30]$.

Implementation steps: Firstly, we need to obtain the extended ragion $\Omega_I = \Omega_{A_i} \bigcup \Omega_{A_r}$, $i = 1,2,...k$. Among them, the complementary regions are represented as: $\Omega_{A_1} = [-20,0] \times [10,30]$, $\Omega_{A_2} = [0,20] \times [10,30]$, the remaining discrete points $\Omega_{A_r} = [20,22] \times [10,30]$, the results are shown in Fig2.(b):



(a)Square extended algorithm.



(b) Rectangular extended algorithm.

Fig2. A Halton extended algorithm for regular regions.

B. Example 2

The extension algorithm of polar coordinates region is based on the rectangular extension algorithm. Our proposed method is very flexible which also includes generation of the complex multi-connected regions. this method advantage is that liberalize the generated Halton points in 2D. Marking solving domain of PDE is not restricted on $[0,1]^2$. This example given an polar coordinates region and multi-connected region extended results with Halton points as shown in Fig3.(a) and Fig3.(b). Fig3.(a) is a petal-shaped polar region with Halton points.

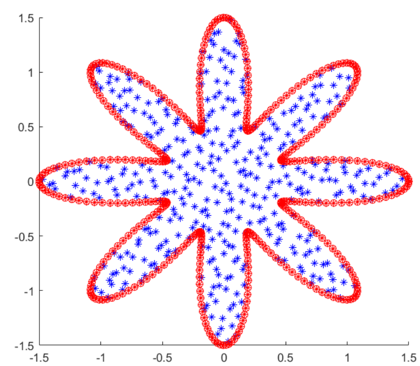
The Fig3 implementation process is as follows. Firstly, drawing the polar coordinate region Γ, and then finding the largest rectangle Ω contained with curve Γ. The coordinates of the two vertices on the diagonal of the rectangle are respectively, $A(min(x_i^b), min(y_i^b))$, $C(max(x_i^b), max(y_i^b))$. Secondly, Filling the Halton point into the rectangle Ω using our proposed rectangle extension algorithm. Thirdly, using the inpolygon () function of MATLAB to judgment whether the coordinates (x_i, y_i) are within the boundary Γ or not. The boundary points plus the internal points are all discrete points on the polar coordinate region. The red pionts are located at the boundary, blue pionts are the internal points. Fig 3.(b) is a multi-connected diagram formed by Halton points, consisting of eight small circles and a centre large circle. The Halton point can be uniformly distributed on this region. It can also be applied to analyze the stress of the bearing section and the production of fiber pipelines in deep study. For Eq. (15) is the mathematical expression of the polar coordinate function of Figure 3.(a), where N is the total number of discrete points in the boundary. The red points represent polar boundary points, and the blue points are internal points of region, the corresponding polar coordinate equation is as follows:

$$\begin{cases} x_b = r\cos\theta \\ y_b = r\sin\theta \\ r = \dfrac{1}{2} + \cos(4(\theta))^2 \\ \theta = \dfrac{2\pi i}{N} \qquad i = 1,2,...N \end{cases} \tag{15}$$

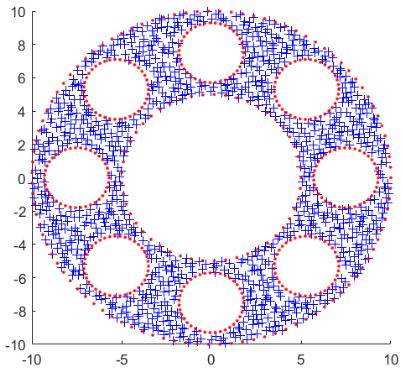
For the multi-connection area in Fig3(b), the polar coordinate formula is:

$$\begin{cases} x_b = r_i\cos\theta \\ y_b = r_i\sin\theta \\ x_c = R\cos\theta \qquad y_c = R\sin\theta \\ \theta = \dfrac{2\pi i}{N} \qquad i = 1,2,...N \end{cases} \tag{16}$$

An example diagram of the polar coordinate halton extended algorithm is shown in Fig3.(a) and Fig3.(b).



(a)The petal-shaped region with Halton points.



(b) Multi-connected region with Halton points.

Fig3. Halton points on polar coordinates region.

V. CONCLUSIONS

In this paper, the main contribution of this paper is to give the extension Halton point algorithm, which can extend the discrete points from $[0,1]^2$ to any two-dimensional plane region. It is no longer affected by the discrete region, so that the Halton point has a broader application prospect. In our work, three examples are given. The first example and the second example belong to the extension algorithm of the square and regular region. It combined with the extension transformation and the translation transformation, the Halton points of $[0,1]^2$ can be extended to any rectangular region. The third example is the extension Halton discrete points in the polar coordinate region, which is also realized on the basis of the extension algorithm of regular rectangular region. This algorithm proposed in this paper which can also be extended to 3D Halton discrete points from $[0,1]^3$ space to any 3D space. Our work provides a good foundation and application value for meshless computing and random point generation, it extended the scope of application of Halton discrete points. This article will promote effect on particle filtering, image rendering, point cloud processing, and the improvement of the initial population of genetic algorithms and etc.

ACKNOWLEDGMENTS

This work is supported by the Youth Science Fund Project (No.11701433).

REFERENCES

[1] M. Koushki, E. Jabbari, “Evaluating RBF methods for solving PDEs using Padua points distribution”. *Alexandria Engineering Journal*, vol. 59, no. 5, pp. 2999-3018, 2020.

[2] Q. Du, M. Gunzburger, L. L. Ju, “Meshfree, probabilistic determination of point sets and support regions for meshless computing”. *Computer Methods in Applied Mechanics and Engineering*, vol. 191, no. 13, pp. 1349-1366, 2020.

[3] S. Y. Wang, G. R. Cai, M. Cheng, and et al. “Robust 3D reconstruction of building surfaces from point clouds based on structural and closed constraints”. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 170, pp. 29-44, 2020.

[4] E. Rafajłowicz, R. Schwabe, “Halton and Hammersley sequences in multivariate nonparametric regression”. *Statistics & Probability Letters*, vol. 76, no. 8, pp. 803-812, 2006.

[5] G. Lee, J. MOU, Y. Shen, “Sampling strategy design for dimensional measurement of geometric features using coordinate measuring machine”. *International Journal of Machine Tools and Manufacture*, vol. 37, no. 7, pp. 917-934, 1997.

[6] C. W. Yang, W Gao, N. G. Liu, C. M. Song, “Low-discrepancy sequence initialized particle swarm optimization algorithm with high-order nonlinear time-varying inertia weight”. *Applied Soft Computing*, vol. 29, pp. 386-394, 2015.

[7] G. Weerasinghe, H. M. Chi, Y. Z. Cao “Particle Swarm Optimization Simulation via Optimal Halton Sequences”. *Procedia Computer Science*, vol. 80, pp. 772-781, 2016.

[8] G. Ökten, Y. N. Liu “Randomized quasi-Monte Carlo methods in global sensitivity analysis”. *Reliability Engineering & System Safety*, vol. 210, pp. 107520, 2021.

[9] P. Krömer, J. Platoš, V. Snášel, “Differential evolution for the optimization of low-discrepancy generalized Halton sequences”. *Swarm and Evolutionary Computation*, vol. 54, pp. 100649, 2020.

[10] D. Bednarik, P. Lertchoosakul, D. Marques, et al. “The generalized and modified Halton sequences in Cantor bases”. *Monatsh Math*, vol. 188, pp. 1-29.