"Il est facile d'apprendre.
Comprendre nécessite un effort parfois long,
rebutant mais permet de réellement goûter aux plats et de se réjouir de saveurs nouvelles."

Albert Jacquard

A la mémoire de ma grand mère
A ma mère Hayat
A mon père Abdellah
A mes enfants Yacine et Malak
A mes frères et ma soeur
A mon mari...

Remerciements

La réalisation de cette thèse a été possible grâce au concours de plusieurs personnes à qui je voudrais témoigner toute ma reconnaissance. Je remercie spécialement mon encadreur, Monsieur Bachir Djebbar, Professeur à l'USTO-Oran. Sans lui cette thèse n'aurait pas vu le jour. Je le remercie vivement pour toutes les qualités dont il a fait preuve. Ses précieux conseils et sa constante disponibilité m'ont été d'un grand secours.

Une attention particulière est à mettre en exergue pour Monsieur Christian Lécot, Professeur à l'Université de Savoie au Bourget du Lac, Co-Directeur de cette thèse. Je lui exprime ma gratitude pour l'attention, le soutien et la disponibilité dont il a fait preuve pour la concrétisation de ce mémoire.

Je remercie Monsieur Hacen Dib, Professeur à l'Université de Tlemcen, d'avoir accepté la présidence du jury de cette thèse. Je lui suis reconnaissante et le remercie vivement pour l'appréciation qu'il portera à ce document.

Je suis très sensible à l'honneur que me font Madame Malika Dali Youcef, Maître de conférences à l'Université de Tlemcen, Monsieur Khemisti Belaib, Maître de conférences à l'Université d'Oran 1, Monsieur Mounir Tlemcani, Professeur à l'USTO-Oran et Monsieur Sidi Mohammed Bouguima, Professeur à l'Université de Tlemcen, en acceptant d'examiner ce travail et de faire partie du jury.

Je voudrais adresser ma gratitude à ma collègue Madame Naima Tebbal et au chef de département de mathématiques, Monsieur Benmiloud Mebkhout pour leur support inestimable et leurs judicieux conseils.

Un grand merci à ma collègue, Madame Lila Khitri et à mon époux Monsieur Abdellatif Benchaib pour leur aide dans le domaine de la programmation.

J'exprime une reconnaissance sans fin pour ma famille qui a toujours cru en moi et m'a toujours soutenue. Je remercie spécifiquement mes parents qui n'ont jamais cessé de m'encourager.

Mes remerciements vont également à Mme Colette Delhaye de Chambéry, qui m'a aidée et hébergée lors de mes déplacements au laboratoire LAMA au Bourget du Lac (France).

Je tiens à remercier tous ceux qui de près ou de loin, m'ont apporté leurs concours à la réalisation de ce travail.

Table des matières

Introduction

Le domaine de l'analyse numérique offre plusieurs méthodes d'approximation. Certaines de ces méthodes sont plus adaptées à quelques types de problèmes que d'autres. Dans ce travail, nous allons nous intéresser aux méthodes quasi-Monte Carlo. Les méthodes de Monte Carlo sont des techniques de simulation utilisant des nombres aléatoires, ce qui a donné leur nom. Ce dernier fait allusion aux jeux de hasard pratiqués à Monte-Carlo. L'appellation "Monte Carlo" est due à N. Metropolis, inspiré de l'intérêt de S. Ulam pour le pocker, car Monte-Carlo est un grand centre de casinos, et a pour origine la similarité avec les jeux de hasard. Les méthodes de Monte Carlo ont été portées au nues à leurs débuts et c'est après la deuxième guerre mondiale qu'ils ont acquis une véritable reconnaissance. S. Ulam, N. Metropolis et notamment Von Neumann, ont utilisé les méthodes de Monte Carlo à Los Alamos pendant la préparation de la bombe atomique avec la collaboration de nombreux scientifiques. Comme les recherches à Los Alamos étaient secrètes, la première publication dans le domaine n'est parue qu'en 1949 voir [38]. Par la suite, le développement de ces méthodes a accompagné les développements de

l'informatique. En effet, dès 1945 J. Von Neumann conjecturait le potentiel des ordinateurs pour la simulation stochastique [60]. Ces méthodes ont jouit d'une bonne réputation dans de nombreux domaines où les méthodes d'analyse numérique étaient inapplicables ou très coûteuses. Par exemple, on a recours aux méthodes de Monte Carlo pour résoudre des problèmes d'équations aux dérivées partielles où les données sont peu régulières, la dimension est élevée et les frontières sont compliquées, théorie du transfert de rayonnement, phénomène d'attente, formules de cubature, équations intégrales et équations de Boltzmann non linéaires. L'inconvénient de ces méthodes est leur faible vitesse de convergence. L'erreur est en $\mathcal{O}(\frac{1}{\sqrt{N}})$ si l'on simule N états, par contre elle est indépendante de la dimension du domaine et c'est pour cette raison qu'elle reste viable pour les problèmes de dimension élevée.

Dès les années 50, les expérimentateurs ont essayé de renoncer au caractère aléatoire des points, en substituant aux suites aléatoires, ou plutôt pseudo-aléatoires (car calculées par des algorithmes déterministes) des suites quasi-aléatoires, dites aussi à discrépance faible. Ces suites sont construites de façon à être réparties le plus uniformément possible dans le domaine considéré. Les méthodes quasi-Monte Carlo contrairement aux méthodes de Monte Carlo fournissent des bornes d'erreur déterministes. Le nom de quasi-Monte Carlo a été employé pour la première fois dans un rapport de recherche de R. D. Richtmyer en 1951. K. F. Roth, médaillé Fields en 1958, a déterminé en 1954 une vitesse de convergence optimale pour l'approximation des intégrales, ainsi qu'une suite utilisant l'idée de J.G. Van der

Corput permettant une convergence rapide. Au cours des années, ont été proposées plusieurs suites à discrépance faible et des théorèmes de bornes que nous préciserons par la suite.

Le plan du document est le suivant: Dans le chapitre 1, nous présentons les méthodes Monte Carlo dans le contexte de l'intégration numérique. Nous rappelons les principaux résultats de convergence, avec la notion d'intervalle de confiance et nous présentons différents générateurs de nombres pseudo-alétoires.

Les méthodes quasi-Monte Carlo sont présentées au chapitre 2. Nous commençons par introduire la discrépance, qui est un outil essentiel d'analyse de ces méthodes. Puis nous définissons les suites à discrépance faible. Ensuite, nous donnons les bornes pour la discrépance de ces suites et nous indiquons un procédé de construction dû à H. Niederreiter. Nous rappelons aussi quelques majorations d'erreur des méthodes de quasi-Monte Carlo et enfin nous prsentons un exemple où on compare les erreurs dans le calcul d'une intégrale par les quadratures Monte Carlo et quasi-Monte Carlo.

Dans le chapitre 3, nous proposons une méthode de Runge-Kutta quasi-Monte Carlo d'ordre 3 pour résoudre un système différentiel ordinaire. Ces méthodes consistent à formuler le problème avec un terme intégral pour ensuite effectuer une quadrature quasi-Monte Carlo. Enfin, nous concluons sur les résultats obtenus et les travaux futurs.

Chapitre 1

Méthodes de Monte Carlo

Nous présentons les outils des méthodes de Monte Carlo (MC). Ce sont des méthodes probabilistes qui permettent d'évaluer certaines quantités en utilisant des nombres (pseudo-)aléatoires.

1.1 Méthodes de Monte Carlo

Les méthodes de Monte Carlo sont introduites dans le contexte de l'intégration numérique.

1.1.1 Intégration numérique multidimensionnelle

Dans de nombreux problèmes scientifiques, on rencontre souvent une intégrale de la forme

$$I(f) = \int_{\bar{\mathcal{I}}^s} f(x)\mathrm{d}x,$$

où $\mathcal{I} = [0, 1),\ s \geq 1$. Le calcul exact de ces intégrales étant souvent impossible, on a recours à des méthodes de quadrature numérique.

En dimension $s = 1$, les formules de quadrature classiques permettent d'approcher l'intégrale d'une fonction par une somme pondérée de ses valeurs en différents points. Des exemples de telles méthodes sont: La formule des trapèzes, de Simpson, des rectangles et des points-milieux.

Pour calculer l'intégrale d'une fonction f sur l'intervalle $\bar{\mathcal{I}}$, où $\mathcal{I} = [0, 1)$, considérons la formule composite des trapèzes. Elle consiste à diviser $\bar{\mathcal{I}}$ en m parties égales pour obtenir l'approximation suivante:

$$\int_{\bar{\mathcal{I}}} f(x)\mathrm{d}x \approx \sum_{k=0}^{m} \omega_k f\Big(\frac{k}{m}\Big),$$

où les ω_k sont des poids définis par

$$\omega_0 = \omega_m = \frac{1}{2m} \text{ et } \omega_k = \frac{1}{m},\ 1 \leq k \leq m-1.$$

Ainsi, il faut évaluer f en $N = m + 1$ points; dans le cas où $f \in \mathcal{C}^2(\bar{\mathcal{I}})$, l'erreur d'approximation est d'ordre

$$\mathcal{O}\Big(\frac{1}{m^2}\Big) = \mathcal{O}\Big(\frac{1}{N^2}\Big).$$

Pour une dimension $s \geq 2$, la généralisation de la méthode des trapèzes consiste à faire s quadratures unidimensionnelles, en considérant successivement les s variables. Dans ce cas, on évalue f en $N = (m + 1)^s$ noeuds de $\bar{\mathcal{I}}^s = [0, 1]^s$. Si f est deux fois continûement dérivable alors l'erreur est d'ordre

$$\mathcal{O}\Big(\frac{1}{m^2}\Big) = \mathcal{O}\Big(\frac{1}{N^{\frac{2}{s}}}\Big).$$

Pour garantir une erreur d'ordre de grandeur inférieur à 10^{-2}, il faudrait utiliser un ordre de 10^s nœuds. Ce nombre croit donc exponentiellement avec s; ce qui rend cette approche inutilisable en dimension élevée.

Pour remédier à cette situation, on a développé les méthodes de Monte Carlo (MC) dont la vitesse de convergence est indépendante de la dimension du problème. L'idée principale de ces méthodes est d'exprimer une intégrale comme l'espérance d'une variable aléatoire.

1.1.2 Quadrature de Monte Carlo

Soit $f\colon \bar{\mathcal{I}}^s \to \mathbb{R}$ une fonction intégrable et $\mathcal{I} = [0,1)$, $s \geq 1$; on veut calculer l'intégrale.

$$I(f) = \int_{\bar{\mathcal{I}}^s} f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}.$$

Les méthodes MC sont basées sur l'idée suivante: si X est une variable aléatoire uniformément distribuée sur $\bar{\mathcal{I}}^s$, et que l'on note $X \sim \mathcal{U}(\bar{\mathcal{I}}^s)$, alors l'espérance de la variable aléatoire $f \circ X$ est

$$\mathbb{E}[f \circ X] = \int_{\bar{\mathcal{I}}^s} f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = I(f).$$

Le problème revient donc à approcher l'espérance de $f \circ X$. Pour cela on se donne une suite $X_1, X_2, \ldots$ de variables aléatoires indépendantes de distribution uniforme sur $\bar{\mathcal{I}}^s$ et on définit, pour tout $N \in \mathbb{N}^*$, la moyenne empirique de la suite

$(f \circ X_n)_{1 \leq n \leq N}$ par

$$M_N := \frac{1}{N} \sum_{n=1}^{N} f \circ X_n.$$

La loi forte des grands nombres [59] assure la convergence presque sûre de la suite $(M_N)_{n \geq 1}$ vers $\mathbb{E}[f \circ X] = I(f)$ quand $N \to \infty$, c'est-à dire :

$$\mathbb{P}\Big[\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f \circ X_n = I(f) \Big] = 1.$$

D'où l'approximation de Monte Carlo.

$$\int_{\bar{\mathcal{I}}^s} f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \approx \frac{1}{N} \sum_{n=1}^{N} f \circ X_n. \tag{1.1}$$

1.1.3 Analyse de l'erreur

Pour analyser l'erreur due à l'approximation (1.1), on suppose que f est de carré intégrable $f \in \mathrm{L}^2(\bar{\mathcal{I}}^s)$. On note la variance $\sigma^2(f)$ de la fonction f par.

$$\sigma^2(f) = \int_{\bar{\mathcal{I}}^s} (f(\boldsymbol{x}) - I(f))^2 \mathrm{d}\boldsymbol{x} = \int_{\bar{\mathcal{I}}^s} (f(\boldsymbol{x}))^2 \mathrm{d}\boldsymbol{x} - I(f)^2,$$

qui est finie puisque $f \in \mathrm{L}^2(\bar{\mathcal{I}}^s)$. Si $X \sim \mathcal{U}(\bar{\mathcal{I}}^s)$, alors la variance de la variable aléatoire $f \circ X$ est

$$Var(f \circ X) = \int_{\bar{\mathcal{I}}^s} (f(\boldsymbol{x}))^2 \mathrm{d}\boldsymbol{x} - (I(f))^2 = \sigma^2(f).$$

Un résultat sur l'intégrale du carré de l'erreur est donné dans [46]

Proposition 1.1.1. *Si $f \in \mathrm{L}^2(\bar{\mathcal{I}}^s)$, on a*

$$\int_{\bar{\mathcal{I}}^s} \dots \int_{\bar{\mathcal{I}}^s} \Big(\frac{1}{N} \sum_{n=1}^{N} f(x_n) - I(f) \Big)^2 \mathrm{d}x_1 \dots \mathrm{d}x_N = \frac{\sigma^2(f)}{N}.$$

D'après cette proposition, l'erreur moyenne d'une quadrature MC est d'ordre

$$\mathcal{O}\Big(\frac{1}{\sqrt{N}}\Big),$$

qui est indépendante de la dimension s. Une estimation probabiliste de l'erreur est obtenue en utilisant le théorème de limite centrale, où $\sigma(f) := \sqrt{\sigma^2(f)}$.

Théorème 1.1.2. *Si $f \in \mathrm{L}^2(\bar{\mathcal{I}}^s)$, on a pour toute constante $c \in \mathbb{R}_+$:*

$$\lim_{N\to+\infty} \mathbb{P}\Big[\Big|\frac{1}{N}\sum_{n=1}^{N} f\circ X_n - I(f)\Big| \leq \frac{c\sigma(f)}{\sqrt{N}}\Big] = \frac{1}{\sqrt{2\pi}}\int_{-c}^{+c} e^{-t^2/2}\mathrm{d}t.$$

C'est à dire que l'erreur

$$\frac{\sqrt{N}}{\sigma(f)}\Big(\frac{1}{N}\sum_{n=1}^{N} f\circ X_n - I(f)\Big)$$

converge en lois vers une loi normale centrée réduite $\mathcal{N}(0,1)$. Nous pouvons donc construire un intervalle de confiance pour $I(f)$ à un niveau de confiance α de la forme.

$$\Big[\frac{1}{N}\sum_{n=1}^{N} f\circ X_n - \frac{c_\alpha\sigma(f)}{\sqrt{N}}, \frac{1}{N}\sum_{n=1}^{N} f\circ X_n + \frac{c_\alpha\sigma(f)}{\sqrt{N}}\Big],$$

où

$$\alpha := 2\phi^*(c_\alpha) - 1,$$

ϕ^* étant la fonction de répartition normale définie par [59]

$$\phi^*(x) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}\mathrm{d}t.$$

Exemple 1.1.3. *En utilisant la table ci-dessous où l'erreur est donné par*

$$Err := \Big|\frac{1}{N}\sum_{n=1}^{N} f\circ X_n - I(f)\Big|,$$

c_α	α
0.84	*0.6*
1.04	*0.7*
1.16	*0.75*
1.64	*0.9*
3.80	*0.99*

et pour N grand, on déduit du théorème 1.1.2 les résultats suivants:

- *avec une probabilité de 60%,* $Err \leq \frac{0.84\sigma(f)}{\sqrt{N}}$;
- *avec une probabilité de 90%,* $Err \leq \frac{1.64\sigma(f)}{\sqrt{N}}$;
- *avec une probabilité de 99%,* $Err \leq \frac{3.80\sigma(f)}{\sqrt{N}}$.

1.1.4 Réduction de la variance

La majoration d'erreur dans les quadratures de MC est de la forme

$$\frac{c_\alpha \sigma(f)}{\sqrt{N}}.$$

Pour un niveau de confiance α fixé, il existe deux manières de réduire cette borne: augmenter le nombre N de nœuds ou réduire la variance $\sigma^2(f)$.

La première méthode est coûteuse puisqu'une augmentation de N d'un facteur 100 entraîne une réduction de l'erreur d'un facteur 10 seulement.

On explore alors la deuxième technique et il existe plusieurs méthodes de réduction de la variance: Echantillonnage préférentiel, conditionnement, stratification, voir [17]. Nous allons présenter l'une de ces méthodes, la technique des variables antithétiques.

On note

$$\boldsymbol{x}_0 := \left(\frac{1}{2}, \ldots, \frac{1}{2}\right) \in \mathbb{R}^s.$$

La densité $\mathbb{1}_{\bar{\mathcal{I}}^s}$ de la loi uniforme est symétrique par rapport à $\boldsymbol{x}_0$:

$$\forall \boldsymbol{x} \in \mathrm{R}^s \ \mathbb{1}_{\bar{\mathcal{I}}^s}(\boldsymbol{x}) = \mathbb{1}_{\bar{\mathcal{I}}^s}(2\boldsymbol{x}_0 - \boldsymbol{x}).$$

Soit X une variable aléatoire, de loi uniforme sur $\bar{\mathcal{I}}^s$. On note

$$V := f(X) \text{ et } V_a := \frac{1}{2}(f(X) + f(2\boldsymbol{x}_0 - X)).$$

On a vu que.

$$\mathbb{E}[V] = I(f) \text{ et } Var(V) = \sigma^2(f).$$

Or

$$\mathbb{E}[f(2\boldsymbol{x}_0 - X)] = \mathbb{E}[f(X)],$$

d'où

$$\mathbb{E}[V_a] = \mathbb{E}[V].$$

On a aussi

$$\mathbb{E}[(f(2\boldsymbol{x}_0 - X))^2] = \mathbb{E}[(f(X))^2],$$

donc

$$Var(f(2\boldsymbol{x}_0 - X)) = Var(f(X)),$$

et par conséquent

$$Var(V_a) = \frac{1}{2}Var\big(f(X)\big) + \frac{1}{2}Cov\Big(f(X), f(2\boldsymbol{x}_0 - X)\Big).$$

D'après l'inégalité de Cauchy-Schwarz,

$$Cov\Big(f(X), f(2\boldsymbol{x}_0 - X)\Big) \leq \sqrt{Var\big(f(X)\big)}\sqrt{Var\big(f(2\boldsymbol{x}_0 - X)\big)} = Var\big(f(X)\big),$$

d'où

$$Var(V_a) \leq Var\big(f(X)\big) = Var(V) = \sigma^2(f).$$

L'utilisation de V_a à la place de V réduit la variance, mais demande deux fois plus d'évaluations de la fonction f. On peut obtenir un résultat plus intéressant [34].

Proposition 1.1.4. *Soit $A_1, A_2, \ldots, A_s$ des sous-ensembles de $\mathbb{R}$, $X_1, X_2, \ldots, X_s$ des variables aléatoires réelles indépendantes, la v.a.r. X_i étant à valeurs dans A_i. Soit $B := A_1 \times A_2 \times \cdots \times A_s$ et deux fonctions $h\colon B \to \mathbb{R}$ et $k\colon B \to \mathbb{R}$. On suppose qu'il existe $\mathrm{R} \subset \{1, 2, \ldots, s\}$ tel que*

- *h et k sont croissantes par rapport à chacune des variables x_i, pour $i \in \mathrm{R}$,*
- *h et k sont décroissantes par rapport à chacune des variables x_i, pour $i \in \mathrm{R}^c$.*

Alors

$$Cov\Big(h(X_1, X_2, \ldots, X_s), k(X_1, X_2, \ldots, X_s)\Big) \geq 0.$$

On déduit une majoration plus précise de la variance de V_a.

Proposition 1.1.5. *Soit une fonction $f : \bar{\mathcal{I}}^s \to \mathbb{R}$, monotone par rapport à chacune de ses variables et soit X une variable aléatoire de loi uniforme sur $\bar{\mathcal{I}}^s$. Si*

$$V := f(X) \text{ et } W := f(2\boldsymbol{x}_0 - X),$$

alors

$$Cov(V, W) \leq 0.$$

Ainsi, si

$$V_a := \frac{1}{2}(f(X) + f(2x_0 - X)),$$

alors

$$Var(V_a) \leq \frac{1}{2} Var(V).$$

1.2 Nombres aléatoires et pseudo-alétoires

Dans tous les calculs de type Monte Carlo, on doit substituer à une variable aléatoire un ensemble de valeurs réelles ayant des propriétés statistiques de la variable aléatoire. Ces valeurs sont appelées les **nombres aléatoires** et doivent être produites "au hasard" à l'aide d'un processus adéquat. Comme l'ordinateur n'est pas capable d'engendrer de telles suites, on a recours à des suites nommées **pseudo-aléatoires** engendrées par des algorithmes utilisant un petit nombre de paramètres.

Il existe plusieurs types de nombres pseudo-aléatoires. Nous citerons ceux qui simulent une loi uniforme $\mathcal{U}(\bar{\mathcal{I}}^s)$. Les générateurs de suites pseudo-aléatoires sont l'objet de nombreuses études théoriques. Ils doivent passer un certain nombre de

tests statistiques afin que les suites engendrées aient certaines propriétés de suites aléatoires [43, 46]. Nous présentons brièvement dans ce paragraphe les générateurs les plus utilisés.

1.2.1 Générateur à congruence linéaire simple

C'est un générateur classique qui a été introduit par D.H. Lehmer [31]. On choisit, trois paramètres M, a et c qui sont des entiers positifs appelés respectivement module, multiplicateur et incrément, et vérifient

$$1 \leq a < M,\ pgcd(a, M) = 1 \ et\ c \in \mathcal{Z}_M = \{0, 1, ..., M-1\}.$$

En partant d'une valeur initiale $y_0 \in \mathcal{Z}_M$ dite germe, telle que $pgcd(y_0, M) = 1$, on construit une suite $(y_n)_{n\geq 0}$ d'éléments de $\mathcal{Z}_M$ par la relation de congruence:

$$y_{n+1} \equiv ay_n + c \pmod{M}. \tag{1.2}$$

Pour éliminer les cas triviaux, on évite

$$a \equiv 1 \quad (M) \text{ et } (a-1)y_0 + c \equiv 0 \pmod{M}.$$

La suite $(x_n)_{n\geq 0}$ de nombres pseudo-aléatoires est alors obtenue en posant:

$$x_n := g(y_n) = \frac{y_n}{M} \in \mathcal{I} \text{ pour tout } n \geq 0, \tag{1.3}$$

g est appelée fonction de sortie. Il est clair qu'un tel générateur est périodique de période maximale M. Il s'agit alors de déterminer les paramètres qui fournissent

la suite ayant la plus grande période possible. On a un résultat classique de D.E. Knuth [25].

Proposition 1.2.1. *La période du générateur à congruence linéaire simple:*

$$y_0 \in \mathcal{Z}_M$$

$$y_{n+1} \equiv a y_n + c$$

$$x_n = \frac{y_n}{M}$$

est égale à M si et seulement si

1. *c et M sont premiers entre eux,*

2. *tout facteur premier de M divise $a - 1$,*

3. *si 4 divise M, alors 4 divise $a - 1$.*

Les trois cas standards rappelés par H. Niedderreiter dans [46] sont les suivants.

- Si M est premier et a est une racine primitive modulo M[1], $c = 0$ et $y_0 \neq 0$, alors la période de la suite $\{x_n\}_{n \geq 0}$ est $M - 1$.

- Si M est une puissance de 2, $a \equiv 5 \pmod 8$ et c est impair, alors la période de la suite $\{x_n\}_{n \geq 0}$ est M.

- Si M est une puissance de 2, $a \equiv 5 \pmod 8$, $c = 0$ et y_0 est impair, alors la période de la suite $\{x_n\}_{n \geq 0}$ est $\frac{M}{4}$.

[1] a est une racine primitive modulo M si le plus petit entier positif p tel que $a^p \equiv 1 \pmod M$ est égal à $M - 1$.

1.2.2 Générateur à congruence linéaire multiple

C'est une généralisation du générateur précédent. On choisit un module M premier, un entier k appelé ordre de la congruence et des multiplicateurs $a_0, a_1, ..., a_{k-1} \in \mathcal{Z}_M$ avec $a_0 \neq 0$. En partant de k valeurs initiales (non toutes nulles) $y_0, y_1, ..., y_{k-1} \in \mathcal{Z}_M$, on construit une suite $(y_n)_{n\geq 0}$ par la formule de récurrence suivante:

$$y_{n+k} \equiv \sum_{l=0}^{k-1} a_l y_{n+l} \pmod{M}.$$

La suite à congruence linéaire multiple $(x_n)_{n\geq 0}$ est obtenue par la relation (1.3).

Ce générateur est périodique de période maximale $M^k - 1$. Cette valeur peut être atteinte pour certains choix des paramètres, voir [46].

1.2.3 Générateur à congruence non linéaire

On part d'un module M et d'un germe $y_0 \in \mathcal{Z}_M$. La suite $(y_n)_{n\geq 0}$ est engendrée par:

$$y_{n+1} \equiv f(y_n) \pmod{M},$$

où f est une fonction à valeurs entières dans $\mathcal{Z}_M$.

Comme dans les cas précédents, la suite $(x_n)_{n\geq 0}$ est obtenue par la formule (1.3). La période maximale de ce générateur est M.

Pour les générateurs à congruence non linéaire quadratique, un résultat analogue à la proposition 1.2.1 est donné par D.E. Knuth [26]

Proposition 1.2.2. *Soit $a, b, c \in \mathcal{Z}_M$. La période du générateur à congruence quadratique:*

$$y_0, y_1 \in \mathcal{Z}_M$$

$$y_{n+1} \equiv a y_n^2 + b y_n + c \pmod{M}$$

$$x_n = \frac{y_n}{M}$$

est égale à M si et seulement si

1. *c et M sont premiers entre eux,*
2. *tout entier premier impair qui divise M divise aussi a et $b-1$,*
3. *si 4 divise M, alors a est pair et 4 divise $a-b+1$; si 2 divise M, alors 2 divise $a-b+1$,*
4. *si 9 divise M, alors ou bien 9 divise a, ou bien 9 divise $b-1$ et 9 divise $ac-6$.*

1.2.4 Générateur à congruence inverse

C'est un générateur à congruence non linéaire pour un choix particulier de la fonction f. Si $c \in \mathcal{Z}_M$, on note $\bar{c}$ l'unique élément de $\mathcal{Z}_M$ où M est premier tel que

$$\begin{cases} c\bar{c} \equiv 1 \pmod{M} & \text{si } c \neq 0, \\ \bar{c} = 0 & \text{si } c = 0. \end{cases}$$

Soit $a \neq 0$ et b deux éléments de $\mathcal{Z}_M$. On part d'un germe $y_0 \in \mathcal{Z}_M$; on construit une suite $(y_n)_{n \geq 0}$ d'éléments de $\mathcal{Z}_M$ par la relation de congruence :

$$y_{n+1} \equiv a\bar{y}_n + b \pmod{M}. \tag{1.4}$$

La suite $(x_n)_{n\geq 0}$ est obtenue par la relation (1.3). Dans ce cas encore, la période est inférieure ou égale à M. On a deux résultats rappelés dans le livre de H.Niederreiter [46].

- Si $M \geq 5$ est un nombre premier, on identifie $\mathcal{Z}_M$ au corps fini $\mathbb{F}_M$ de cardinal M.

  Si $a, b \in \mathbb{F}_M$ sont tels que le polynôme $x^2 - bx - a$ est primitif sur $\mathbb{F}_M$ [2], alors la suite $(x_n)_{n\geq 1}$ est de période M.

- Si M est de la forme $M = 2^\alpha$ avec $\alpha \geq 3$, on choisit y_0 impair; alors la période de la suite est inférieure ou égale à $M/2$. Elle est égale à $M/2$ si et seulement si 4 divise $a - 1$ et $b - 2$.

1.2.5 Générateur MRG32k3a

Ce générateur a été proposé par P. L'Ecuyer [30]; il combine deux générateurs d'ordre 3. Son initialisation nécessite deux vecteurs $\mathbf{s}_{1,0}, \mathbf{s}_{2,0} \in \mathbb{N}^3$.

A l'étape i, deux vecteurs

$$\mathbf{s}_{1,i} = (x_{1,i}, x_{1,i+1}, x_{1,i+2}) \; et \; \mathbf{s}_{2,i} = (x_{2,i}, x_{2,i+1}, x_{2,i+2})$$

[2] Le polynôme $x^2 - bx - a$ est primitif sur $\mathbb{F}_M$ s'il a une racine dans $\mathbb{F}_{M^2}$ qui engendre le groupe cyclique $\mathbb{F}^*_{M^2}$.

sont engendrés à partir des précédents par des congruences linéaires multiples:

$$x_{1,i} = 1\ 403\ 580\ x_{1,i-2} - 810\ 728\ x_{1,i-3} \pmod{M_1},$$

$$x_{2,i} = 527\ 612\ x_{2,i-1} - 1\ 370\ 589\ x_{2,i-3} \pmod{M_2},$$

où

$$M_1 := 2^{32} - 209 = 4\ 294\ 967\ 087 \text{ et } M_2 := 2^{32} - 22\ 853 = 4\ 294\ 944\ 443.$$

Puis, la suite $(x_n)_{n\geq 0}$ de nombres pseudo-aléatoires est définie par :

$$x_n = \begin{cases} \frac{z_n}{M_1+1} & \text{si } z_n > 0, \\ \\ \frac{M_1}{M_1+1} & \text{si } z_n = 0, \end{cases}$$

où l'on a posé

$$z_n = x_{1,n} - x_{2,n} \pmod{M_1}.$$

Il s'agit d'un générateur de période longue, égale à:

$$(M_1^3 - 1)(M_2^3 - 1) \approx 2^{191} \approx 3.1 \times 10^{57}.$$

Ce générateur s'est avéré performant même pour des problèmes en dimension élevée, plus précisément jusqu'à la dimension 45.

1.2.6 Générateur Mersenne Twister (MT19937)

Ce générateur a été proposé par Matsumoto et Nishimura [37], leur idée était de définir la récurrence du générateur, non pas à partir des opérations arithmétiques

classiques sur les entiers, mais à partir des opérations d'arithmétique matricielle dans le corps fini $\mathcal{Z}_2 = \{0, 1\}$.

Nous présentons quelques notions pour comprendre l'algorithme de MT:

1. L'ensemble des entiers représentables en machine est de la forme $\mathcal{Z}_{2^\omega}$, où ω désigne le nombre de bits de l'ordinateur. Tout entier $Y \in \mathcal{Z}_{2^\omega}$, de décomposition binaire $\sum_{i=0}^{\omega-1} y_i 2^i$, est stocké sous la forme d'un vecteur de bits; $Y = (y_{\omega-1}, \ldots, y_0)$.

2. Décalage de bits:

   - décalage de v bits vers la droite: $Y \gg v = (0, \ldots, 0, y_{\omega-1}, y_{v+1})$
   - décalage de v bits vers la gauche: $Y \ll v = (y_{\omega-v-1}, y_0, 0, \ldots, 0)$.

3. Opération bit à bit: Soit $X = \sum_{i<\omega} x_i 2^i$ et $Y = \sum_{\imath<\omega} y_i 2^i$. Les opérateurs bit à bit sont définis par:

$$X \oplus Y = \sum_{i<\omega} (x_i \oplus y_i) 2^i \text{ et } X \otimes Y = \sum_{i<\omega} (x_i \otimes y_i) 2^i$$

   où $x_i \oplus y_i \equiv (x_i + y_i) \pmod 2$ et $x_i \otimes y_i \equiv (x_i \times y_i) \pmod 2$.

4. On pose

$$A(x) = (x \gg 1) \oplus 0 \text{ si } x \equiv 0 \pmod 2 \ (x \text{ est pair}),$$

$$A(x) = (x \gg 1) \oplus a \text{ si } x \equiv 1 \pmod 2 \ (x \text{ est impair}).$$

5. On pose

$\underline{M}_r = (1, \ldots, 1, 0, \ldots, 0)$ où les r premiers bits valent 1, les autres 0,

$\overline{M}_r = (0, \ldots, 0, 1, \ldots, 1)$ où les r derniers bits valent 1, les autres 0.

- La dynamique de Mersenne Twister est définie par 2 étapes:

1. Opération de récurrence:

   $X_{k+n} = X_{k+m} \oplus A\Big((X_{k+1} \otimes \underline{M}_r) \oplus (X_k \otimes \overline{M}_r)\Big)$

   La dynamique de MT est donc basée sur un shéma récurrent d'ordre n dans l'ensemble des entiers machines. Pour $k \geq 0$ le terme X_{k+n} est construit à partir de X_k, X_{k+1} et X_{k+m} $(0 \leq m \leq n)$.

2. Opération de tempring:

   Cette opération consiste à mélanger les bits de X_{k+n}, afin d'augmenter encore l'imprédictibilité des valeurs générées. X_{k+n} est transformé de la manière suivante:

$$Y \leftarrow X_{k+n}$$

$$Y \leftarrow Y \oplus (Y \gg u)$$

$$Y \leftarrow Y \oplus ((Y \ll s) \otimes b)$$

$$Y \leftarrow Y \oplus ((Y \ll t) \otimes c)$$

$$Y \leftarrow Y \oplus (Y \gg l)$$

- La fonction de sortie utilisée est donnée par $U_k = g(Y_k) = \frac{Y_k+0.5}{2^\omega}$. U_k est bien à valeurs dans $[0, 1]$.

- Le générateur MT19937 génère les nombres U_k, avec les paramètres:

  - Paramètre de récurrence: $\omega = 32, n = 624, r = 31, m = 397, a = 2567483615$.

  - Paramètre de tempering: $u = 11, s = 7, t = 15, l = 18, b = 2636928640, c = 4022730752$.

Ce choix permet de maximiser la période, égale à $T_{MT} = 2^{\omega n-r} - 1 = 2^{19937} - 1$. La période étant de la forme $M_n = 2^n - 1$, il s'agit d'un Mersenne number plus précisément, comme $2^{19937} - 1$ est un nombre premier, la période est un Mersenne prime.

Ainsi Mersenne Twister est un générateur pseudo-aléatoire qui possède une construction similaire aux générateurs congruentiels, car on retrouve la récurrence du générateur basée sur une période. Néanmoins, la différence est que la dynamique de MT est construite non pas à partir d'opérations arithmétiques sur les entiers, mais à partir d'opérations sur les bits.

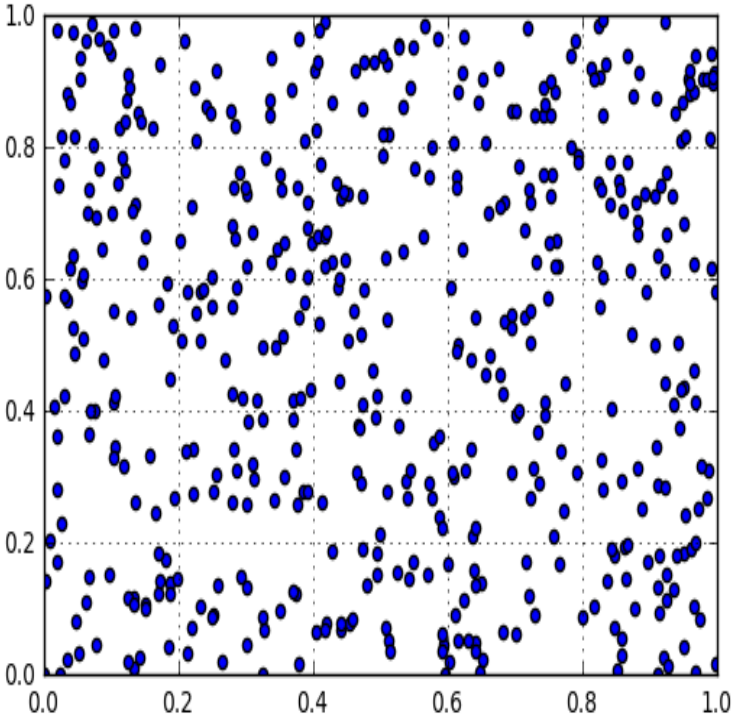


Figure 1.1: Ensemble de 500 points pseudo-aléatoires

Chapitre 2

Méthodes quasi-Monte Carlo

Dans ce chapitre nous présentons les méthodes quasi-Monte Carlo (QMC) qui sont les analogues déterministes des méthodes MC.

Elles sont basées sur l'utilisation d'ensembles de points déterministes, les points quasi-aléatoires ou à faible discrépance. Ces points sont caractérisés à l'aide d'une notion de discrépance vue en théorie des nombres, qui est essentielle dans l'analyse des méthodes QMC. Elle mesure "la qualité de la répartition uniforme" des points dans le tore s-dimensionnel $\mathcal{I}^s = [0,1)^s$ et que nous présenterons dans la première section de ce chapitre.

Dans la deuxième section, nous décrivons quelques ensembles et suites quasi-aléatoires, ou à discrépance faible, qui sont utilisées dans les chapitres suivants.

2.1 Discrépance

L'étude de l'équirépartition d'une suite de points dans le domaine d'intégration $\mathcal{I}^s$ est fondamentale pour les quadratures quasi-Monte Carlo. La discrépance est une mesure de l'uniformité de la suite, ou de sa déviation par rapport à la distribution uniforme. Les références classiques sont le livre de L.Kuipers et H.Niederreiter [27], [46]. Nous donnons en premier la définition d'une suite uniformément répartie.

Définition 2.1.1. *Une suite $\{x_n : n \geq 1\}$ de points de $\bar{\mathcal{I}}^s$ est dite uniformément répartie ou équirépartie si pour tout sous-intervalle E de $\bar{\mathcal{I}}^s$, on a:*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_E(x_n) = \lambda_s(E), \tag{2.1}$$

où $\mathbb{1}_E$ désigne la fonction indicatrice de E et λ_s la mesure de Lebesgue dans $\mathbb{R}^s$.

Pour simplifier l'écriture, nous allons utiliser la notation suivante: Si $X = \{x_n : 1 \leq n \leq N\}$ est un ensemble de N points de $\bar{\mathcal{I}}^s$ et si E est un sous-ensemble de $\bar{\mathcal{I}}^s$, on note $A(E, X)$ le nombre d'indices n (compris entre 1 et N) tels que $x_n \in E$: La relation (2.1) peut alors s'écrire:

$$\lim_{N\to\infty} \frac{A(E, X)}{N} = \lambda_s(E).$$

On peut donc définir la discrépance.

Définition 2.1.2. *La discrépance (appelée aussi discrépance extrême) d'un ensemble $X = \{x_n : 1 \leq n \leq N\}$ de points de $\bar{\mathcal{I}}^s$ est définie par:*

$$D_N(X) := \sup_{E\in\mathcal{J}} \left| \frac{A(E, X)}{N} - \lambda_s(E) \right|,$$

où $\mathcal{J}$ est l'ensemble des sous intervalles de $\mathcal{I}^s$ de la forme $\prod_{i=1}^{s}[u_i, v_i)$, avec $0 \leq u_i < v_i \leq 1$.

Il est clair que $0 < D_N(X) \leq 1$ et que plus la discrépance est faible, plus la répartition des points est uniforme. On peut définir d'autres types de discrépance en changeant la classe de sous-ensembles sur laquelle on définit le maximum.

Définition 2.1.3. *La discrépance à l'origine d'un ensemble $X = \{x_n : 1 \leq n \leq N\}$ de points de $\bar{\mathcal{I}}$ est définie par:*

$$D_N^*(X) := \sup_{E \in \mathcal{J}^*} \left| \frac{A(E, X)}{N} - \lambda_s(E) \right|,$$

où $\mathcal{J}^*$ est l'ensemble des sous-intervalles de $\mathcal{I}^s$ de la forme $\prod_{i=1}^{s}[0, u_i)$, avec $0 < u_i \leq 1$.

Définition 2.1.4. *La discrépance isotrope d'un ensemble $X = \{x_n : 1 \leq n \leq N\}$ de points de $\bar{\mathcal{I}}^s$ est définie par :*

$$J_N(X) := \sup_{C \in \mathcal{J}_C} \left| \frac{A(C, X)}{N} - \lambda_s(C) \right|,$$

où $\mathcal{J}_C$ est l'ensemble des sous-ensembles convexes de $\bar{\mathcal{I}}^s$.

Des inclusions $\mathcal{J}^* \subset \mathcal{J} \subset \mathcal{J}_C$ on déduit

$$0 < D_N^*(X) \leq D_N(X) \leq J_N(X) \leq 1.$$

On note

$$D_N(E, X) := \frac{1}{N} A(E, X) - \lambda_s(E).$$

Cette quantité est appelée la discrépance locale de X en E.

Dans le cas où X est un ensemble de points de $\mathcal{I}^s$, on a:

$$D_N(X) = \sup_{E \in \bar{\mathcal{J}}} |D_N(E, X)| \text{ et } D_N^*(X) = \sup_{E \in \bar{\mathcal{J}}^*} |D_N(E, X)|,$$

où $\bar{\mathcal{J}}$ est l'ensemble des sous-intervalles de $\mathcal{I}^s$ de la forme $\prod_{i=1}^s [u_i, v_i]$ et $\bar{\mathcal{J}}^*$ est l'ensembles des sous-intervalles de la forme $\prod_{i=1}^s [0, u_i]$.

L'évaluation des discrépances $D_N(X)$ et $D_N^*(X)$ étant en général très difficile en dimension $s \geq 2$ [40], des algorithmes de calcul effectif ou de majoration ont été mis au point par E. Thiémard en dimension $s = 1$. Nous présentons quelques résultats [41], [44], [46] sur la discrépance.

Proposition 2.1.5. *Pour tout ensemble X de N nombres de $\bar{\mathcal{I}} = [0, 1]$, on a*

$$\frac{1}{N} \leq D_N(X) \leq 1. \tag{2.2}$$

Proposition 2.1.6. *Les discrépances $D_N(X)$ et $D_N^*(X)$ sont reliées par*

$$D_N^*(X) \leq D_N(X) \leq 2^s D_N^*(X). \tag{2.3}$$

Proposition 2.1.7. *Si $X = \{x_n \ : 1 \leq n \leq N\}$ est un ensemble de points de $\bar{\mathcal{I}}$ où $0 \leq x_1 \leq x_2 \leq \cdots \leq 1$, alors*

- *pour la discrépance à l'origine*

$$D_N^*(X) = \frac{1}{2N} + \max_{1 \leq n \leq N} \left| x_n - \frac{2n-1}{2N} \right|. \tag{2.4}$$

- *pour la discrépance*

$$D_N(X) = \frac{1}{N} + \max_{1\leq n<N} \left| \frac{n}{N} - x_n \right| - \min_{1\leq n<N} \left| \frac{n}{N} - x_n \right|. \tag{2.5}$$

Il résulte de (2.4) que $D_N^*(X) \geq \frac{1}{2N}$ avec égalité pour

$$x_n = \frac{2n-1}{2N},\ 1 \leq n \leq N$$

On peut étendre la notion de discrépance à une suite infinie $S = \{x_n : n \geq 1\}$ de $\bar{\mathcal{I}}^s$ en notant $D_N(S)$ la discrépance (respectivement $D_N^*(S)$ la discrépance à l'origine) des N premiers termes de S. Cela permet de caractériser l'équirépartition d'une suite de points de $\bar{\mathcal{I}}^s$.

Définition 2.1.8. *Une suite infinie $S = \{x_n : n \geq 1\}$ de $\bar{\mathcal{I}}^s$ est uniformément répartie si*

$$\lim_{N\to+\infty} D_N(\{x_n, 0 \leq n < N\}) = 0.$$

On a alors l'équivalence entre les propositions suivantes:

1. S est uniformément répartie sur $\bar{\mathcal{I}}^s$,
2. $\lim_{N\to+\infty} D_N(S) = 0$,
3. $\lim_{N\to+\infty} D_N^*(S) = 0$.

Par définition D_N et D_N^* sont des discrépances en norme L^∞. Il existe également des discrépances en norme L^p.

Définition 2.1.9. *Soit $X = \{x_0, x_1, ..., x_{N-1}\}$ une suite finie d'éléments de $\bar{\mathcal{I}}^s$. Si $D_N(J, X)$ désigne la discrépance locale, on définit la discrépance dans L^p de X par:*

$$T_N^{(p)}(X) = \left[\int_{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{I}^{2s};x_i<y_i} |D_N(J,X)|^p \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{y} \right]^{\frac{1}{p}}, \tag{2.6}$$

où $J = \prod_{i=1}^{s}[x_i, y_i)$, *avec* $\boldsymbol{x} = (x_1, x_2, ..., x_s), \boldsymbol{y} = (y_1, y_2, ..., y_s)$. *La discrépance à l'origine dans* L^p *de* X *est donnée par:*

$$T_N^{(p)*}(X) = \left[\int_{\mathcal{I}^s} (D_N(J, X))^p \mathrm{d}\boldsymbol{b}\right]^{\frac{1}{p}}, \tag{2.7}$$

où $J = \prod_{i=1}^{s}[0, b_i)$ *et* $\boldsymbol{b} = (b_1, b_2, ..., b_s)$.

On montre dans [43] que $D_N^*(X) \leq D_N(X) \leq 2^s D_N^*(X)$, en utilisant le fait que la norme L^∞ est plus grande que la norme L^2, on a $T_N^{(2)*}(X) \leq D_N^*(X)$. On a aussi la relation suivante : $k_s(D_N(X))^{(s+2)/2} \leq T_N^{(2)*}(X)$, où k_s est une constante positive dépendant de s. Dans ce cas on a les équivalences suivantes.

1. S est uniformément répartie sur $\bar{\mathcal{I}}^s$,
2. $\lim_{N\to+\infty} T_N^{(p)}(S) = 0$,
3. $\lim_{N\to+\infty} T_N^{(p)*}(S) = 0$.

De même que pour le cas fini, un calcul exact de $D_N^*(X)$ est possible quoique très couteux en dimension $s > 1$. Il est aussi montré dans [40] que si

$$M_{m,n,i} = \max(x_{n,i}, x_{m,i})$$

et

$$m_{m,n,i} = \min(x_{n,i}, x_{m,i})$$

alors

$$\left(T_N^{(2)*}(X)\right)^2 = \frac{1}{N^2}\sum_{n=1}^{N}\sum_{m=1}^{N}\prod_{i=1}^{s}(1-M_{m,n,i}) - \frac{2^{-s+1}}{N}\sum_{n=1}^{N}\prod_{i=1}^{s}(1-(x_{n,i}^2)) + 3^{-s}, \tag{2.8}$$

où $x_{n,i}$ est la $i^{ème}$ coordonnée du $n^{ème}$ élément de la suite considérée, dans [39] que

$$\left(T_N^{(2)}(X)\right)^2 = \frac{1}{N^2}\sum_{n=1}^{N}\sum_{m=1}^{N}\prod_{i=1}^{s}(1-M_{m,n,i})m_{m,n,i} - \frac{2^{-s+1}}{N}\sum_{n=1}^{N}\prod_{i=1}^{s}(1-x_{n,i})x_{n,i} + 12^{-s}. \tag{2.9}$$

La figure 2.1 représente respectivement les 1024 points pseudo-aléatoires, obtenue par le générateur **MT19937** et à discrépance faible (ensemble de Hammersley), on constate que les points dans la deuxième figure sont équirépartis. Par contre dans les suites pseudo-aléatoires, il apparaît des rectangles ne contenant aucun point.

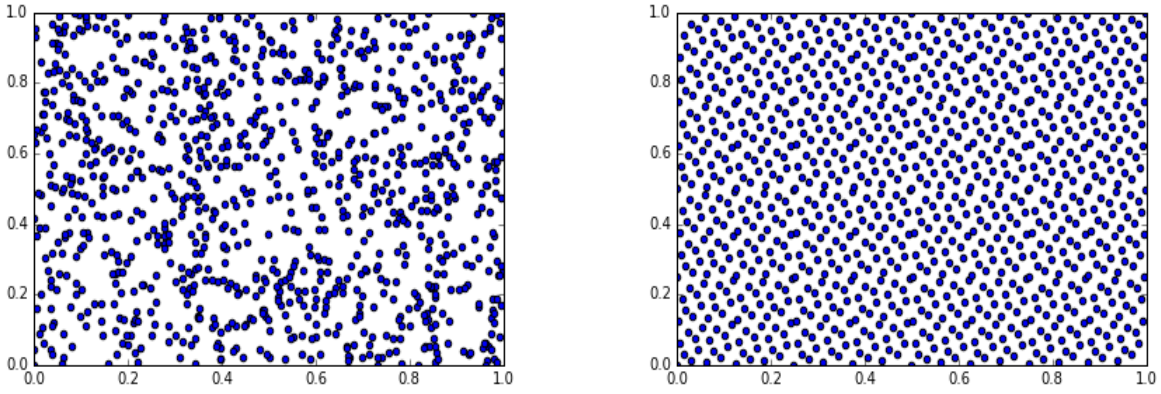


Figure 2.1: Ensembles de 1024 points pseudo-aléatoires (à gauche) et de 1024 points à discrépance faible (à droite).

2.2 Suites à discrépance faible

Dans cette partie on présente quelques suites de points de $\mathcal{I}^s$ à discrépance faible. Nous commençons par donner les bornes classiques de discrépance.

Dans [46], H. Niederreiter a indiqué une technique permettant de construire, à partir d'une suite à discrépance faible dans $\mathcal{I}^{s-1}$, un ensemble fini à discrépance faible dans $\mathcal{I}^s$ qui est basée sur le résulat suivant:

Lemme 2.2.1. *Soit $s \geq 2$ un entier et $S = \{\boldsymbol{x}_n : n \geq 0\}$ une suite dans $\mathcal{I}^{s-1}$. Pour $N \geq 1$, soit*

$$\tilde{X}_N := \{\Big(\frac{n}{N}, \boldsymbol{x}_n\Big) : 0 \leq n \leq N-1\} \subset \mathcal{I}^s.$$

Alors

$$ND_N^*(\tilde{X}_N) \leq \max_{1 \leq M \leq N} MD_M^*(S) + 1. \tag{2.10}$$

Dans le cas unidimensionnel, il est facile de déterminer le minimum de $D_N(X)$ et $D_N^*(X)$, où X est un ensemble de N points dans $\bar{\mathcal{I}} = [0,1]$. En effet, la proposition 2.1.7 montre que.

$$D_N^*(X) \geq \frac{1}{2N} \quad \text{et} \quad D_N(X) \geq \frac{1}{N}.$$

De plus ces bornes sont atteintes pour l'ensemble

$$X = \{\frac{2n-1}{2N} : 1 \leq n \leq N\}.$$

Par contre, un résultat de W.M. Schmidt [53] montre qu'il existe une constante $c > 0$ telle que pour toute suite infinie $S = \{x_n : n \geq 1\} \subset \bar{\mathcal{I}}$, on ait

$$D_N(S) \geq c\frac{\log(N)}{N},$$

pour une infinité de valeurs de N. La meilleure valeur de c connue a été donnée par R. Béjian voir [2] et elle vaut 0.12. L'inégalité $D_N(X) \leq 2D_N^*(X)$ implique une minoration pour la discrépance à l'origine analogue à la précédente avec une constante égale à 0.06.

Ainsi, en dimension 1, il n'existe aucune suite infinie dont la discrépance décroisse plus rapidement que $\mathcal{O}\Big(\frac{(\log N)}{N}\Big)$; on connaît des suites ayant exactement cette décroissance, comme les suites de Van de Corput (voir plus loin).

En dimension s quelconque, on conjecture que la discrépance à l'origine d'un ensemble X de N points vérifie:

$$D_N^*(X) \geq K_s \frac{(\log N)^{s-1}}{N}, \tag{2.11}$$

où $K_s > 0$ est une constante qui ne dépend que de s. Pour $s = 2$, cette inégalité à été établie par W.M. Schmidt voir [54]. Il n'existe pas de preuve pour $s \geq 3$. La borne connue jusqu'à présent, due à K.F. Roth [52], est la suivante.

Théorème 2.2.2. *Il existe une constante $B_s > 0$ ne dépendant que de s telle que tout ensemble fini $X \subset \bar{\mathcal{I}}^s$ vérifie:*

$$D_N^*(X) \geq B_s \frac{(\log N)^{(s-1)/2}}{N}.$$

Si l'inégalité (2.11) était vérifiée, on démontrerait à l'aide du lemme 2.2.1 l'existence d'une constante $K'_s > 0$ ne dépendant que de s telle que pour toute suite infinie S,

$$D_N^*(S) \geq K'_s \frac{(\log N)^s}{N}$$

pour une infinité de valeurs de N.

On présente un résultat analogue à celui de K.F. Roth pour les suites infinies:

Théorème 2.2.3. *Il existe une constante $B'_s > 0$ ne dépendant que de s telle que toute suite infinie $S \subset \bar{\mathcal{I}}^s$ vérifie :*

$$D_N^*(S) \geq B'_s \frac{(\log N)^{s/2}}{N},$$

pour une infinité de valeurs de N.

On ne connaît pas d'ensemble fini de points dont la disrcépance soit de l'ordre donné par le théorème 2.2.2, ni de suite infinie dont la discrépance soit de l'ordre donné par le théorème 2.2.3. J.M. Hammersley a été le premier à proposer voir [20] un ensemble fini X de points vérifiant:

$$D_N^*(X) = \mathcal{O}\Big(\frac{(\log N)^{s-1}}{N}\Big). \tag{2.12}$$

De même, J.H. Halton a été le premier à proposer voir [19] une suite infinie S vérifiant:

$$D_N^*(S) = \mathcal{O}\Big(\frac{(\log N)^{s}}{N}\Big), \tag{2.13}$$

pour tout $N \geq 2$. On appelle ensemble à discrépance faible tout ensemble fini vérifiant (2.12) et suite à discrépance faible toute suite vérifiant (2.13) pour tout $N \geq 2$. Nous présentons dans la suite des exemples d'ensembles et de suites à discrépance faible avec l'estimation de la discrépance à l'origine correspondante.

Définition 2.2.4. *Une suite $S \subset [0,1)^s$ vérifiant*

$$\forall N \geq 2;\ D_N^*(S) = \mathcal{O}\Big(\frac{(\log N)^{s}}{N}\Big)$$

est appelée suite à discrépance faible et une méthode d'approximation utilisant une telle suite est appelée une méthode quasi-Monte Carlo.

Asymptotiquement la convergence est donc plus rapide que pour la méthode de Monte Carlo standard qui, on le rappelle, est en $\mathcal{O}(1/\sqrt{N})$.

2.2.1 Suites de Van der Corput, de Halton et ensemble de Hammersley

Ce sont des suites à discrépance faible dans $\mathcal{I}$. On a besoin de définir la fonction radicale inverse. Soit $b \geq 2$ un entier. Tout entier $n \in \mathbb{N}$ a un développement unique en base b de la forme

$$n = \sum_{j=0}^{\infty} a_j(n) b^j, \tag{2.14}$$

où $a_j(n) \in \mathcal{Z}_b = \{0, \ldots, b-1\}$ pour tout $j \geq 0$ L'expression (2.14) est une somme finie car $a_j(n) = 0$ pour j suffisamment grand.

La fonction radicale inverse ϕ_b en base b est définie par :

$$\phi_b(n) := \sum_{j=0}^{\infty} a_j(n) b^{-j-1} \in \mathcal{I},\ n \in \mathbb{N},$$

où les $a_j(n)$ sont les coefficients de la représentaion (2.14).

- Suite de Van der Corput

Définition 2.2.5. *La suite de van der Corput en base b est la suite*

$$S_b = \{\phi_b(n) \colon n \geq 0\} \subset \mathcal{I}.$$

La discrépance à l'origine de la suite S_b est d'ordre $\mathcal{O}\Big((\log N)/N\Big)$ pour tout $N \geq 2$, avec une constante qui ne dépend que de b. Plus précisément, on a le résultat asymptotique suivant dû à Faure [13]:

$$\limsup_{N\to\infty} \frac{ND_N^*(S_b)}{\log N} = \limsup_{N\to\infty} \frac{ND_N(S_b)}{\log N} = \begin{cases} \frac{b^2}{4(b+1)\log b} & \text{si } b \text{ est pair,} \\ \frac{b-1}{4\log b} & \text{si } b \text{ est impair.} \end{cases}$$

Exemple 2.2.6. *En base $b = 3$, les premiers points de la suite de Van der Corput sont*

n	*0*	*1*	*2*	*3*	*4*	*5*	*6*
$\phi_3(n)$	*0*	*1/3*	*2/3*	*1/9*	*4/9*	*7/9*	*2/9*

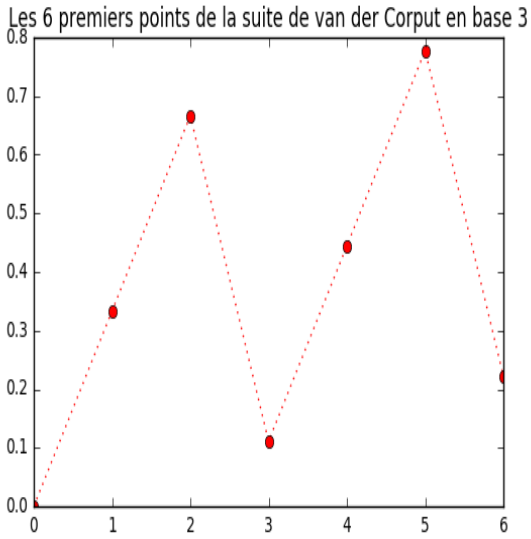


Figure 2.2: Suite de Van der Corput en base 3

Asymptotiquement, S_3 est la meilleure suite de Van der Corput. R. Béjian et H. Faure ont également obtenu une majoration pour la discrépance des n premiers points de la suite de Van der Corput en base 2 [1].

Théorème 2.2.7. *Pour la suite de Van der Corput en base* 2*, on a*

$$D_N^*(S_2) = D_N(S_2) \leq \frac{\log N}{3\log 2} + 1, \text{ pour tout, } N \geq 1$$

et

$$\limsup_{N\to\infty} \left(D_N(S_2) - \frac{\log N}{3\log 2} \right) = \frac{4}{9} + \frac{\log 3}{3\log 2}.$$

Remarque 2.2.8. *On peut améliorer ces résultats si l'on considère une suite de Van der Corput généralisée en base b dont le terme d'ordre n est défini par*

$$\boldsymbol{x}_n = \sum_{j=0}^{\infty} \sigma(a_j(n)) b^{-j-1},$$

où σ est une permutation de $\mathcal{Z}_b$.

La suite présentant la plus faible discrépance asymptotique connue en dimension 1 est de ce type. Elle est donnée dans le théorème suivant [15].

Théorème 2.2.9. *Si $\boldsymbol{S}$ est la suite de Van der Corput généralisée en base $b = 12$ engendrée par la permutation*

$$\sigma_{12} = (0\ 5\ 9\ 3\ 7\ 1\ 10\ 4\ 8\ 2\ 6\ 11).$$

Alors

$$\limsup_{N\to\infty} \frac{N D_N^*(\boldsymbol{S})}{\log N} = \frac{1919}{3454 \log 12} \approx 0.224.$$

Le fait que la constante obtenue soit deux fois plus petite que celle de la meilleure suite de Van der Corput S_3 montre le potentiel que l'on peut tirer de l'utilisation de permutations dans la construction de suites à discrépance faible.

- Suites de Halton: C'est une généralisation des suites de Van der Corput en dimension $s \geq 2$.

Définition 2.2.10. *Soit $b_1, ..., b_s$ des entiers ≥ 2. La suite de Halton associée aux bases $b_1, \ldots, b_s$ est la suite*

$$H_{b1,\ldots,b_s} = \{(\phi_{b_1}(n), \ldots, \phi_{b_s}(n)) : n \geq 0\} \subset \mathcal{I}^s,$$

où, pour $1 \leq i \leq s$, ϕ_{b_i} est la fonction radicale inverse associée à b_i

H. Niederreiter a montré dans [46] que si $b_1, \ldots, b_s$ sont premiers entre eux deux à deux, la discrépance à l'origine de la suite de Halton vérifie, pour tout $N \geq 2$:

$$D_N^*(H_{b_1,\ldots,b_s}) \leq A(b_1, \ldots, b_s)\frac{(\log N)^s}{N} + \mathcal{O}\Big(\frac{(\log N)^{s-1}}{N}\Big), \tag{2.15}$$

où

$$A(b_1, \ldots, b_s) = \prod_{i=1}^{s} \frac{(b_i - 1)}{2 \log b_i}. \tag{2.16}$$

La valeur minimale de cette constante est obtenue en prenant pour bases les s premiers nombres premiers. Dans ce cas la constante $A(b_1, \ldots, b_s)$ est notée A_s.

- Ensemble de Hammersley [21],[38]: On obtient l'ensemble de Hammersley en appliquant le lemme 2.2.1 à la suite de Halton $H_{b_1,\ldots,b_{s-1}} \subset \mathcal{I}^s$. Ses éléments sont de la forme :

$$\boldsymbol{x}_n = \big((\frac{n}{N}, \phi_{b_1}(n), \ldots, \phi_{b_{s-1}}(n)\big),$$

pour $0 \leq n \leq N-1$.

Dans le cas où $b_1, \ldots, b_{s-1}$ sont premiers entre eux deux à deux, il résulte de (2.10) et (2.15) que la discrépance à l'origine d'un ensemble de Hammersley $X = \{\boldsymbol{x}_n;\ 0 \leq n \leq N-1\}$ vérifie:

$$D_N^*(X) \leq A(b_1, \ldots, b_{s-1}) \frac{(\log N)^{s-1}}{N} + \mathcal{O}\Big(\frac{(\log N)^{s-2}}{N}\Big), \tag{2.17}$$

où $A(b_1, \ldots, b_{s-1})$ est la constante définie par (2.16). Comme dans le cas précédent, la valeur minimale de cette constante est notée A_{s-1}.

Remarque 2.2.11. *Le majorant A_s a une croissance exponentielle quand $s \to \infty$:*

$$\lim_{s\to\infty} \frac{\log A_s}{s \log s} = 1.$$

Cela fait perdre de l'intérêt aux bornes (2.15) et (2.17) en grande dimension.

Il est indispensable de choisir les bases b_i premières entre elles pour que la suite remplisse l'hypercube. Par exemple la suite bidimensionnelle définie avec les bases $b_1 = 2$ et $b_2 = 6$ ne comporte aucun point dans l'ensemble $[0, 1/2] \times [5/6, 1]$ (Figure 2.4). La suite de Van der Corput en base b se décompose en cycles

monotones de longueur b et les projections des suites de Halton sur des coordonnées associées à des grandes valeurs de la base produiront de longs ensembles parallèles à la diagonale (Figure 2.5).

Une manière de conjurer l'apparition de ce balayage du carré unité par une succession de diagonales progressivement décalées au fur et à mesure des itérations consiste à généraliser les suites de Halton. Il s'agit d'une simple adaptation de la transformation déja appliquée aux suites de Van der Corput. Plus explicitement, les composantes des suites de Halton généralisées sont des suites de Van der Corput généralisées engendrées par différentes permutations. Par exemple, l'utilisation des permutations

$$\sigma_{17} = (0\ 8\ 13\ 3\ 11\ 5\ 16\ 1\ 10\ 7\ 14\ 4\ 12\ 2\ 15\ 6\ 9)$$

et

$$\sigma_{19} = (0\ 9\ 14\ 3\ 17\ 6\ 11\ 1\ 15\ 7\ 12\ 4\ 18\ 8\ 2\ 16\ 10\ 5\ 13)$$

proposées par E. Braaten et G. Weller [3] pour les bases 17 et 19 semble conduire à des résultats plus satisfaisants au niveau de la distribution à l'intérieur du carré unité (Figure 2.6). D'autres auteurs se sont intéréssés à la question du choix des permutations [15] et [58].

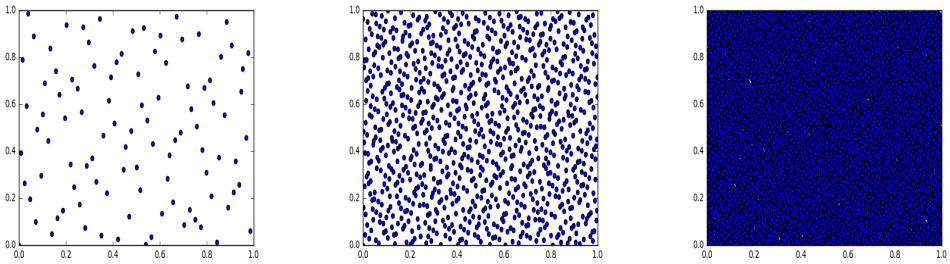


Figure 2.3: 100, 1000 et 10000 premiers points d'une suite de Halton en bases $(2, 3)$

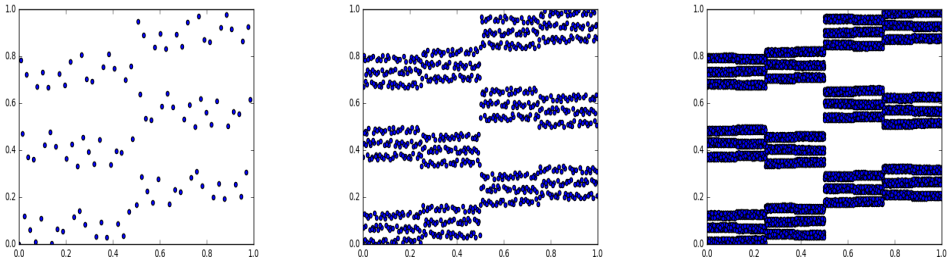


Figure 2.4: 100, 1000 et 10000 premiers points d'une suite de Halton en bases $(2, 6)$

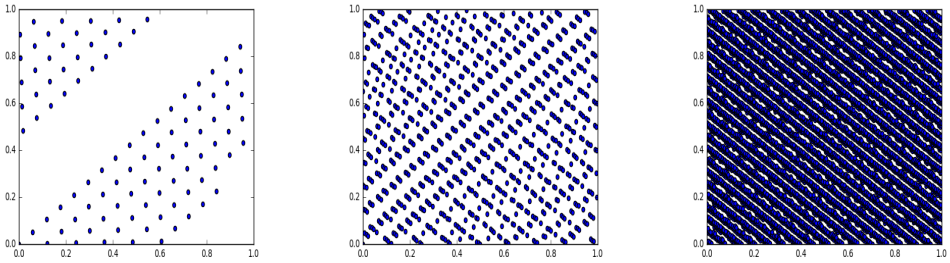


Figure 2.5: 100, 1000 et 10000 premiers points d'une suite de Halton en bases $(17, 19)$

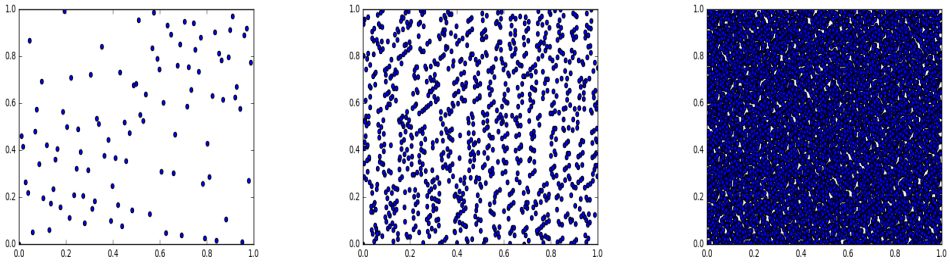


Figure 2.6: 100, 1000 et 10000 premiers points d'une suite de Halton généralisée en bases $(17, 19)$ engendrée par les permutations σ_{17} et σ_{19}

2.2.2 Suite de Faure

Dans [14] H. Faure a proposé une construction de suite utilisant une numération $b-$adique. Soit s et b deux entiers avec b premier et $1 \leq s \leq b$. On choisit s entiers distincts $r_1, ..., r_s$ compris entre 1 et b. On écrit pour tout $n \in \mathbb{N}$, son développement (2.14) en base b. On note alors

$$y_{n,j}^{(i)} := \sum_{k=j-1}^{\infty} \binom{k}{j-1} r_i^{k-j+1} a_k(n) \pmod{b}$$

et

$$x_n^{(i)} := \sum_{j=1}^{\infty} y_{n,j}^{(i)} b^{-j} \in \mathcal{I}.$$

Puisque les termes $a_k(n)$ sont nuls à partir d'un certain rang alors les sommes précédentes sont finies. La suite de Faure en base b est la suite

$$F_b := \{(x_n^{(1)}, \ldots, x_n^{(s)} : n \geq 0\} \subset \mathcal{I}^s.$$

Faure a montré que ces suites vérifient l'inégalité:

$$D_N^*(F_b) \leq C_s \frac{(\log N)^s}{N} + \mathcal{O}\Big(\frac{(\log N)^{s-1}}{N}\Big),$$

pour tout $N \geq 1$ où

$$C_s = \begin{cases} \frac{3}{16(\log 2)^2} \text{ si } b = s = 2, \\ \\ \frac{1}{s!}\Big(\frac{b-1}{2\log b}\Big)^s \text{ si } s \geq 2 \text{ et } b \geq s \text{ impair premier.} \end{cases}$$

Remarque 2.2.12. *Si $b(s)$ est le plus petit entier premier supérieur ou égal à s, alors le majorant*

$$C_s = \frac{1}{s!}\Big(\frac{b(s)-1}{2\log b(s)}\Big)^s$$

tend vers 0 *lorsque* $s \to \infty$. *La suite de Faure a donc un comportement asymptotique à priori meilleur que celui des suites de Halton pour les grandes valeurs de* s.

Une étude expérimentale des discrépances des principales suites à discrépance faible a été faite par W. J. Morokoff et R. E. Caflisch [29]

2.2.3 Réseaux-(t, m, s) et suites-(t, s)

H. Niederreiter a généralisé la construction de H.Faure [42] en introduisant une classe particulière d'ensembles finis et de suites infinies à discrépance faible, appelés respectivement réseaux-(t, m, s) et suites-(t, s). Un réseau est un ensemble fini X de points dans $\mathcal{I}^s$ dont la discrépance locale $D_N(E, X)$ est nulle pour une famille particulière (mais assez exhaustive) de sous-intervalles E de $\mathcal{I}^s$. Cela implique qu'il est à discrépance faible. Une suite-(t, s) est une suite infinie de points de $\mathcal{I}^s$ dont des segments successifs particuliers forment des réseaux-(t, m, s): cela implique de même qu'elle est à discrépance faible. On précise cela plus loin.

2.2.4 Définitions

Dans la suite du document, s désigne une dimension supérieure ou égale à 1 et λ_s est la mesure de Lebesgue sur $\mathbb{R}^s$.

Définition 2.2.13. *On appelle intervalle élémentaire de* $\mathcal{I}^s$ *en base* b *tout intervalle de la forme*

$$\prod_{i=1}^{s} \left[\frac{a_i}{b^{d_i}}, \frac{a_i + 1}{b^{d_i}} \right),$$

où d_i et a_i sont des entiers tels que $d_i \geq 0$ et $0 \leq a_i < b^{d_i}$ pour tout $1 \leq i \leq s$.

Définition 2.2.14. *Soit m et t deux entiers tels que $0 \leq t \leq m$. Un réseau-(t, m, s) en base b est un ensemble X de b^m points dans $\mathcal{I}^s$ tel que $A(E, X) = b^t$ pour tout intervalle élémentaire E en base b vérifiant $\lambda_s(E) = b^{t-m}$.*

On peut conclure de la définition 2.2.14 que la discrépance locale $D_N(E, X)$ d'un réseau-(t, m, s) en base b est nulle pour tout intervalle élémentaire E en base b tel que $\lambda_s(E) = b^{t-m}$.

Définition 2.2.15. *Soit $t \geq 0$ un entier. Une suite $S = \{\boldsymbol{x}_n \colon n \geq 0\} \subset \mathcal{I}^s$ est appelée suite-(t, s) en base b si, pour tous les entiers $k \geq 0$ et $m > t$, l'ensemble $X = \{\boldsymbol{x}_n \colon kb^m \leq n < (k+1)b^m\}$ est un réseau-(t, m, s) en base b.*

Exemple 2.2.16.

- *La suite de Van der Corput en base b est une suite-$(0, 1)$ en base b; par contre, si $s \geq 2$, les suites de Halton ne sont pas des suites-(t, s) car elles ne sont pas associées à une base unique.*

- *Si $s = 2$ et $m \in \mathbb{N}$, l'ensemble de Hammersley :*

$$X = \{\big(\frac{n}{N}, \phi_b(n)\big) \colon 0 \leq n < N\},$$

où $N = b^m$, est un réseau-$(0, m, 2)$ en base b.

On présente ci-dessous un résultat dû à H. Faure [14]

Proposition 2.2.17. *La suite de Faure F_b est une suite-$(0, s)$ en base b pour tout nombre premier $b \geq s$.*

Les figures 2.7, 2.8, 2.9 et 2.10 représentent respectivement les points d'un réseau-(0,4,2) en base 2, les 1000 et 10000 points de Hammersley et les 3^5 premiers points de la suite de Faure.

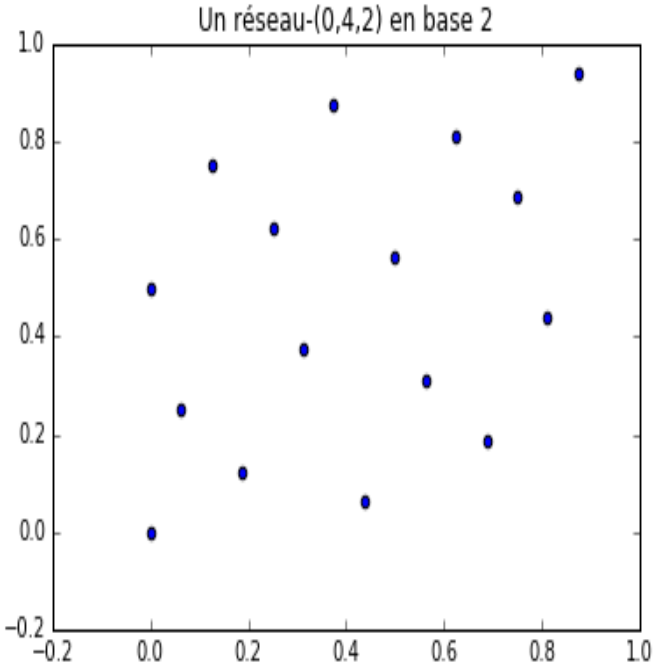


Figure 2.7: Ensemble de Hammersley où $N = 2^4$

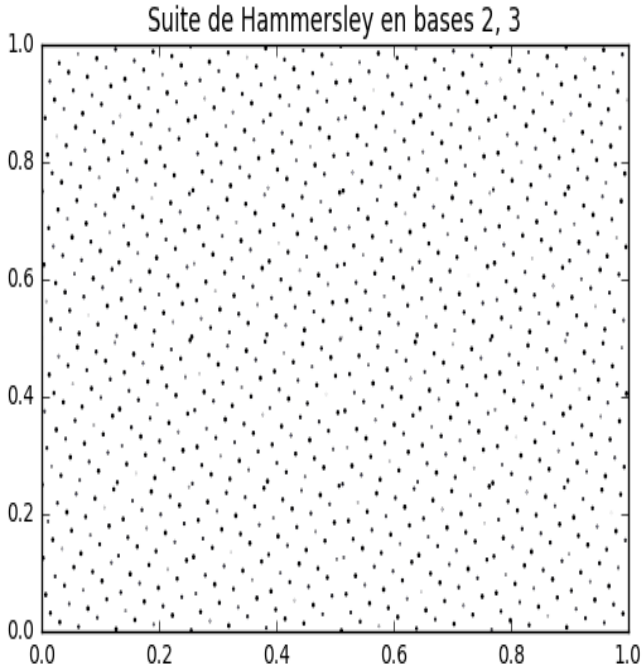


Figure 2.8: 1000 points de Hammersley

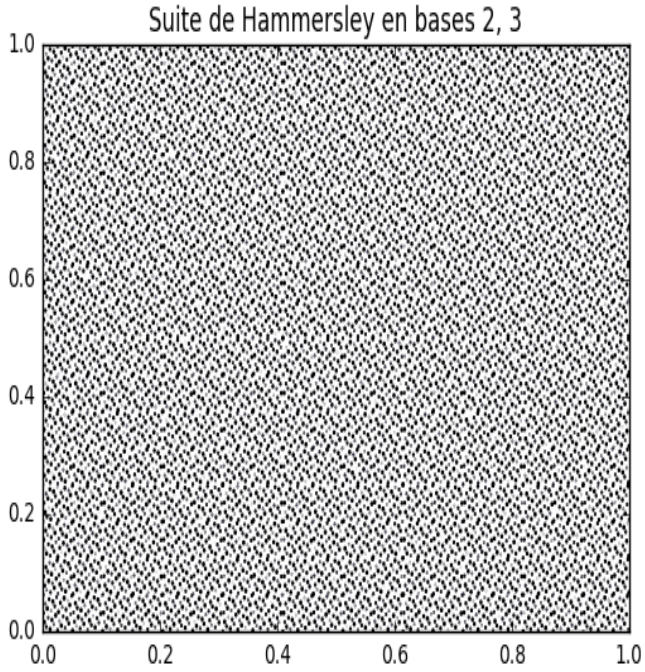


Figure 2.9: 10000 points de Hammersley

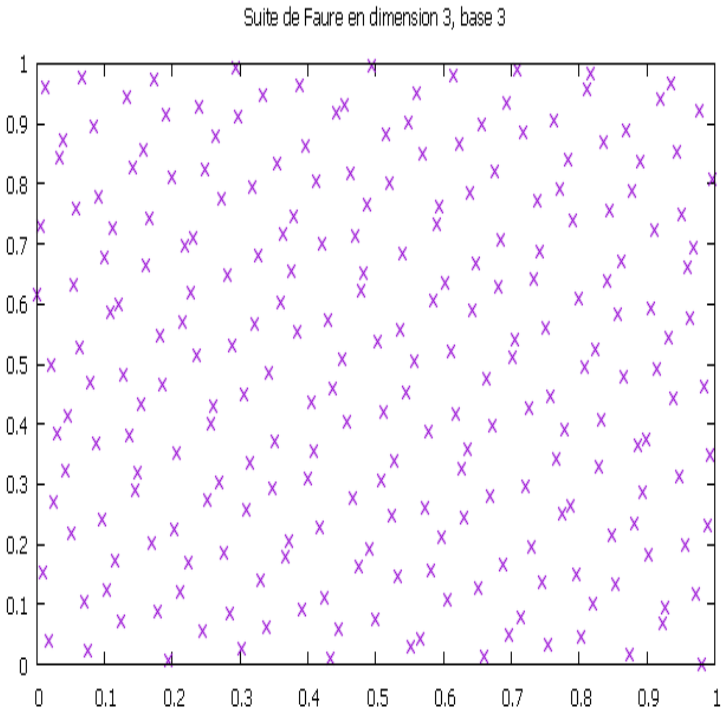


Figure 2.10: Les 3^5 premiers points de la Suite de Faure

2.2.5 Suites de Niederreiter

H.Niederreiter a proposé dans [45] un procédé général de construction de suites$-(t, s)$ en base q. On se limite ici au cas où q est une puissance première, c'est-à-dire $q = p^\alpha$ avec p premier et $\alpha \geq 1$. On note $\mathbb{F}_q$ le corps fini à q éléments. On choisit ce qui suit:

(N1) des bijections $\psi_r \colon \mathcal{Z}_b \to \mathbb{F}_q$ pour tout $r \geq 0$, telles que $\psi_r(0) = 0$ pour r suffisamment grand;

(N2) des bijections $\eta_{i,j} \colon \mathbb{F}_q \to \mathcal{Z}_b$ pour tout $1 \leq i \leq s$ et $j \geq 1$, telles que $\eta_{i,j}(0) = 0$ pour tout $0 \leq i \leq s$ et j suffisamment grand;

(N3) des éléments $c_{j,r}^{(i)} \in \mathbb{F}_q$ pour tout $1 \leq i \leq s$, tout $r \geq 0$ et j suffisamment grand.

On note $\mathbb{F}_q((x^{-1}))$ le corps des séries de Laurent formelles sur $\mathbb{F}_q$. Tout élément $L \in \mathbb{F}_q((x^{-1}))$ s'écrit

$$L = \sum_{k=\omega}^{\infty} t_k x^{-k}.$$

où $\omega \in \mathbb{Z}$ et $t_k \in \mathbb{F}_q$ pour tout $k \geq \omega$. On choisit des polynômes irréductibles [1], unitaire [2]distincts $p_1, ..., p_s \in \mathbb{F}_q[x]$. On pose $e_i := deg(p_i)$ pour tout $1 \leq i \leq s$. Pour $1 \leq i \leq s,\ j \geq 1$ et $0 \leq k \leq e_i$, on écrit le développement en série de Laurent

[1]Un polynôme est irréductible si dans toute factorisation en deux polynômes, l'un est constant.

[2]Un polynôme est unitaire si le coefficient du monôme de plus grand degré est égal à 1.

:

$$\frac{x^k}{p_i(x)^j} = \sum_{r=0}^{\infty} a^{(i)}(j,k,r)x^{-r-1}. \tag{2.18}$$

Ce développement permet de définir pour tout $1 \leq i \leq s,\ j \geq 1$ et $r \geq 0$ des éléments $c_{j,r}^{(i)}$ de $\mathbb{F}_q$:

$$c_{j,r}^{(i)} := a^{(i)}(Q(i,j)+1, k(i,j), r), \tag{2.19}$$

où

$$j-1 = Q(i,j)e_i + k(i,j) \text{ avec } 0 \leq k(i,j) < e_i.$$

On remarque que pour tout $1 \leq i \leq s$ et $r \geq 0$ les éléments $c_{j,r}^{(i)}$ sont nuls pour j suffisamment grand.

Ainsi la suite de Niederreiter est définie par $S = \{\boldsymbol{x}_n \colon n \geq 0\}$ dans $\mathcal{I}^s$. Soit $n \in \mathbb{N}$ et

$$n = \sum_{r=0}^{\infty} a_r(n)q^r$$

son développement en base q, avec $a_r(n) \in \mathcal{Z}_q$ pour $r \geq 0$. Pour tout $n \geq 0,\ 1 \leq i \leq s$, soit

$$x_n^{(i)} := \sum_{j=1}^{\infty} y_{n,j}^{(i)} b^{-j},$$

où, pour tout $j \geq 1$

$$y_{n,j}^{(i)} := n_{i,j}\Big(\sum_{r=0}^{\infty} c_{j,r}^{(i)} \psi_r(a_r(n))\Big) \in \mathcal{Z}_b.$$

Les éléments $\boldsymbol{x}_n$ sont alors définis par :

$$\boldsymbol{x}_n := \{x_n^{(1)}, ..., x_n(s)),\ n \geq 0.$$

Un résultat fondamental sur ces suites est donné par le théorème suivant.

Théorème 2.2.18. *La suite de Niederreiter est une suite*$-(t(q,s),s)$ *en base* q*, où*

$$t(q,s) := \sum_{i=1}^{s}(e_i - 1).$$

Afin d'obtenir la meilleure majoration de discrépance, il faut minimiser la valeur de $t(q,s)$ pour s et q fixés. Pour cela on range les polynômes irréductibles unitaires de $\mathbb{F}_q[x]$ par degré croissant: $p_1, p_2, ...$ et on choisit les s premiers. Des tables des polynômes irréductibles unitaires de $\mathbb{F}_q[x]$ pour q premier sont données dans [32], [33].

Cas particulier

Dans le cas où q est une puissance première et s une dimension $\leq q$, on peut choisir des polynômes de la forme $p_i(x) = x - b_i$ pour tout $1 \leq i \leq s$, où $b_1, ..., b_s$ sont des éléments distincts de $\mathbb{F}_q$. Avec ce choix, on a $t(q,s) = 0$.

La relation (2.19) s'écrit alors

$$c_{j,r}^{(i)} = a^{(i)}(j,0,r),\ 1 \leq i \leq s,\ j \geq 1,\ r \geq 0.$$

Ces éléments s'obtiennent en écrivant le développement en série de Laurent :

$$\begin{aligned}\frac{1}{p_i(x)^j} = \frac{1}{x^j(1-b_i x^{-1})^j} &= x^{-j}\sum_{r=0}^{\infty}\binom{r+j-1}{j-1}b_i^r x^{-r} \\ &= \sum_{r=j-1}^{\infty}\binom{r}{j-1}b_i^{r-j+1}x^{-r-1}.\end{aligned}$$

Pour $1 \leq i \leq s$ et $j \geq 1$, on a

$$c_{j,r}^{(i)} = \begin{cases} 0 & \text{pour } 0 \leq r < j-1, \\ \binom{r}{j-1}b_i^{r-j+1} & \text{pour } r \geq j-1, \end{cases}$$

en convenant que $0^0 = 1 \in \mathbb{F}_q$. Quand q est premier, on identifie $\mathcal{Z}_q$ à $\mathbb{F}_q$ et on choisit les bijections ψ_r et $\eta_{i,j}$ égales à l'identité : on retrouve alors la construction de H. Faure [14].

Les ensembles digitaux et les suites digitales généralisent les réseaux (t, m, s) et les suites (t, s) [10], [56]. On présente dans ce qui suit leur principe de construction.

1) Soient $s \geq 1,\ b \geq 2$ et $k \geq 1$ trois entiers. On considère:

   - Un anneau commutatif unitaire R de cardinal b;
   - des bijections $\psi_r \colon \mathcal{Z}_b \to R$, pour $0 \leq r \leq k-1$;
   - des matrices (dites génératrices) $C_1, \ldots, C_s$ de dimension $k \times k$ d'éléments de R.

   Pour $i = 0, ..., b^k - 1$, on écrit la représentation de i en base b:

$$i = \sum_{r=0}^{k-1} a_r b^r,$$

   où $a_r \in \mathcal{Z}_b$. Soit

$$y = (\psi_0(a_0), \ldots, \psi_{k-1}(a_{k-1}))^t \in R^k$$

   et

$$(b_{j,1}, b_{j,2}, \ldots, b_{j,k})^t = C_j.y,$$

   avec $b_{j,l} \in R$. Pour $j = 1, \ldots, s$ soit

$$u_{i,j} = \frac{\eta_{j,1}(b_{j,1})}{b} + \frac{\eta_{j,2}(b_{j,2})}{b^2} + \cdots + \frac{\eta_{j,l}(b_{j,k})}{b^k}.$$

L'ensemble $P = \{u_i = (u_{i,1}, \ldots, u_{i,s}),\ i = 0, \ldots, b^k - 1\}$ est appelé réseau digital sur R en base b.

2) Soient $s \geq 1$ et $b \geq 2$ deux entiers. On considère :

- Un anneau commutatif unitaire R de cardinal b;
- des bijections $\psi_r \colon \mathcal{Z}_b \to R$, pour $r \geq 0$, vérifiant $psi_r(0) = 0$, pour r suffisamment grand;
- des bijections $\eta_{j,l} \colon R \to \mathcal{Z}_b$, pour $1 \colon j \colon s$ et $l \geq 1$;
- des matrices (dites génératrices) $C_1, \ldots, C_s$ d'éléments de R, d'indices dans $\mathcal{N}^* \times N^*$.

Pour $i \geq 0$, on écrit la représentation de i en base b:

$$i = \sum_{r=0}^{\infty} a_r b^r,$$

où $a_r \in \mathbb{Z}_b$. Soit

$$y = (\psi_0(a_0), \psi_1(a_1), \ldots)^t \in R^{\mathcal{N}^*}$$

et

$$(b_{j,1}, b_{j,2}, \ldots)^t = C_j, y,$$

avec $b_{j,l} \in R$. Pour $j = 1, \ldots, s$, soit

$$u_{i,j} = \frac{\eta_{j,1}(b_{j,1})}{b} + \frac{\eta_{j,2}(b_{j,2})}{b} + \ldots.$$

La suite $P = \{u_i = (u_{i,1}, \ldots, u_{i,s}),\ i \geq 0\}$ est appelée suite digitale sur R en base b.

Exemple 2.2.19. *Soient $s \geq 1$ et $b \geq 2$, puissance première supérieure à s. Les suites de Faure généralisées [56] sont obtenues en utilisant la construction précédente dans laquelle les matrices génératrices sont de la forme:*

$$C_j = A_j P^{j-1},\ j = 1, \ldots, s,$$

où P est la transposée de la matrice de Pascal:

$$P_{i,j} := \binom{j-1}{i-1} \in \mathbb{F}_b$$

et A_j est une matrice triangulaire inférieure régulière dont les éléments sont dans $\mathbb{F}_b$. Ces suites sont des suites $(0, s)$ en base b.

La suite de Faure originale est obtenue en prenant b le plus petit entier premier $\geq s$ et les matrices A_j toutes égales à la matrice identité I.

Exemple 2.2.20. *Une construction de suites, élaborée par H. Faure et S. Tezuka [16], [57] consiste à multiplier à droite les matrices génératrices des suites -(t, s) par des matrices régulières triangulaires supérieures (NUT : non-singular upper triangular). Les nouvelles matrices génératrices sont:*

$$C'_j = C_j U_j,$$

où les U_j, $j = 1, \ldots, s$ sont des matrices NUT. Comme la multiplication par les matrices U_j ne préserve pas les propriétés de répartition dans les intervalles élémentaires, la suite ainsi obtenue n'est généralement pas de type (t, s). Une possibilité pour conserver la propriété d'uniformité des suites est de choisir U_j sous la

forme:

$$U_j = \gamma_j U,$$

où U est une matrice NUT fixée et $\gamma_j \in \mathbb{F}_b,\ \gamma_j \neq 0$. Les matrices génératrices deviennent alors:

$$C'_j = \gamma_j C_j U.$$

On a le résultat suivant [16]: Si une suite -(t, s) en base b est engendrée par les matrices génératrices $C_j,\ 1 \leq j \leq s$, alors la suite engendrée avec les matrices $C'_j = \gamma_j C_j U,\ 1 \leq j \leq s$, est également une suite -(t, s) en base b.

2.3 Hasardisation

Le but des techniques de hasardisation est de construire des suites de points à faible discrépance, vérifiants :

1. chaque suite hasardisée est uniformément distribué sur $[0, 1)^s$,

2. la régularité de la suite hasardisée est celle de la suite déterministe de départ.

2.3.1 Méthodes de décalage linéaire

Les décalages linéaires sont les procédés les plus simples de hasardisation

Décalage aléatoire modulo 1

Cette méthode de hasardisation est aussi connue sous le nom de rotation de Cranley-Patterson [8]. Soit $P_n := \{u_i; i = 0, 1, \ldots, n-1\}$ un ensemble de points de

$[0,1)^s$ et Δ un vecteur aléatoire $s-$dimensionnel uniformément distribué sur $[0,1)^s$. L'ensemble hasardisé $\tilde{P}_n := \{\tilde{u}_i; i = 0, 1, \ldots, n-1\}$ est défini par :

$$\tilde{u}_i \equiv (u_i + \Delta) \pmod 1$$

L'ensemble ainsi obtenu est uniformément distribué sur $[0,1)^s$. Par contre cette technique de hasardisation ne préserve pas les propriétés d'équirépartition de l'ensemble de départ.

Décalage digital $b-$adique

Soit $P_n := \{u_i;\ i = 0, 1, \ldots, n-1\}$ un réseau (t, m, s) en base b. Cette méthode est analogue à la précédente mais elle consiste à écrire la représentation $b-$adique du vecteur Δ et à additionner ses composantes à celles des points u_i, en utilisant les opérations sur $\mathbb{F}_b$ [35],[24]. Plus précisément, si $\Delta = (\Delta_1, \ldots, \Delta_s)$ avec

$$\Delta_j = \sum_{l=1}^{\infty} d_{j,l} b^{-l},\ u_{i,j} = \sum_{l=1}^{\infty} u_{i,j,l} b^{-l},$$

on calcule

$$u_i \oplus \Delta = (\tilde{u}_{i,1}, \ldots, \tilde{u}_{i,s}),$$

où

$$\tilde{u}_{i,j} = \sum_{l=1}^{\infty} ((u_{i,j,l} + d_{j,l}) \pmod b)^{-l}.$$

L'ensemble hasardisé est alors $\tilde{P}_n = \{\tilde{u}_i;\ i = 0, \ldots, n-1\}$. Ce type de décalage digital est le plus approprié aux réseaux (t, m, s) puisqu'il préserve l'uniformité des points et les valeurs du paramètre t.

2.3.2 Ensembles de points à faible discrépance brouillés

Cette méthode a été proposée par A. Owen dans [47], [48]. L'idée est de perturber les digits des ensembles de points à faible discrépance à l'aide de permutations aléatoires de digits, tout en préservant les propriétés d'équirépartition.

Soit $P = (u_i)_i$ un ensemble de points de $[0,1)^s$. On considère la représentation $b-$adique de chacune des composantes du $i^{ème}$ terme $u_i = (u_{i,1}, \ldots, u_{i,s})$ de P:

$$u_{i,j} = \sum_{k=1}^{\infty} u_{i,j,k} b^{-k},$$

où $0 \leq u_{i,j,k} < b, \forall i, j, k$.

La version hasardisée de P est l'ensemble $\tilde{P}$ dont le $i^{ème}$ élément $\tilde{u}_i = (\tilde{u}_{i,1}, \ldots, \tilde{u}_{i,s})$ est donné par :

$$\tilde{u}_{i,j} = \sum_{k=1}^{\infty} \tilde{u}_{i,j,k} b^{-k},$$

où les $\tilde{u}_{i,j,k}$ sont définis comme suit:

$$\begin{aligned}
\tilde{u}_{i,j,1} &= \pi_j(u_{i,j,1}) \\
\tilde{u}_{i,j,2} &= \pi_{j,u_{i,j,1}}(u_{i,j,2}) \\
&\vdots \\
\tilde{u}_{i,j,k} &= \pi_{j,u_{i,j,1},\ldots,u_{i,j,k-1}}(u_{i,j,k}),
\end{aligned}$$

les $\pi_{j,i_{i,j,1},\ldots,u_{i,j,l}} : \mathcal{Z}_b \to \mathcal{Z}_b$ étant des permutations aléatoires indépendantes uniformément distribuées sur l'ensemble des $b!$ permutations possibles de $\{0, 1, \ldots, b-$

1} et les permutations sont mutuellement indépendantes. Notons que la permutation π_j permute le premier digit en base b de $u_{i,j}$ pour tout i. Le deuxième digit est permuté par $\pi_{ju_{i,j,1}}$. La permutation appliquée au deuxième digit $u_{i,j,2}$ dépend de la valeur du premier digit $u_{i,j,1}$. De même, la permutation appliquée au $k^{ème}$ digit $u_{i,j,k}$ dépend de la valeur du premier $k-1$ digit $u_{i,j,1}$ par $u_{i,j,k-1}$, ce qui fait que l'implémentation numérique de ces méthodes demande des espaces de stockage et des temps de calcul importants.

Dans [48] A. Owen a montré que si P est un réseau -(t, m, s) (respectivement une suite -(t, s)) en base b alors $\tilde{P}$ est un réseau -(t, m, s) (respectivement une suite -(t, s)) en base b presque sûrement et que les points de l'ensemble $\tilde{P}$ sont uniformément distribués sur $[0, 1)^s$.

Des alternatives permettant la réduction de la place mémoire ont été proposées par J. Matousek [35], S. Tezuka [56], H. H. Hong et F. J. Hickernell [24]. L'idée est d'appliquer des permutations affines aux différents coefficients comme suit. Si $n = b^k$, on considère:

- s matrices régulières triangulaires inférieures $L_1, \ldots, L_s$, à indices dans $\mathbb{N}^* \times \mathbb{N}^*$, dont les éléments sont choisis aléatoirement et indépendamment dans $\mathbb{F}_b$, avec des éléments diagonaux non nuls;

- s vecteurs (à indice dans $\mathbb{N}^*$) $d_1, \ldots, d_s$ dont les composantes sont indépendantes et uniformément distribuées dans $\mathbb{F}_b$.

  Les coefficients $\tilde{u}_{i,j,1}, \tilde{u}_{i,j,2} \ldots$ de la $j^{ème}$ composante $\tilde{u}_{i,j}$ de l'élément $\tilde{u}_i$ sont

donnés par:

$$\begin{pmatrix} \tilde{u}_{i,j,1} \\ \tilde{u}_{i,j,2} \\ \vdots \end{pmatrix} = L_j \begin{pmatrix} u_{i,j,1} \\ u_{i,j,2} \\ \vdots \end{pmatrix} + d_j,$$

toutes ces opérations étant effectuées dans $\mathbb{F}_b$. Le décalage digital garantit que chaque point obtenu est uniformément distribué dans $[0,1)^s$ [49]. Cette technique préserve les propriétés d'équirépartition de l'ensemble original. En effet, dans [36] J. Matousek a montré que la valeur du paramètre t de la suite ainsi obtenue ne dépasse pas celle de la suite originale, le brouillage peut donc potentiellement améliorer la qualité de la suite.

2.4 Intégration numérique quasi-Monte Carlo

Les méthodes QMC, comme déjà vu, sont des versions déterministes des méthodes MC.

$$\int_{\bar{I}^s} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \approx \frac{1}{N}\sum_{n=1}^{N} f(\boldsymbol{x}_n),$$

où $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ est un ensemble de points à discrépance faible. Cette méthode a une meilleure vitesse de convergence que la méthode de Monte Carlo et l'estimation d'erreur est dans ce cas déterministe.

2.4.1 Estimation d'erreur

On commence par le cas unidimensionnel.

Définition 2.4.1. *Soit f une fonction réelle définie sur $\bar{\mathcal{I}} = [0,1]$. On dit que f est à variation bornée sur $\bar{\mathcal{I}}$ s'il existe une constante $M > 0$ telle que*

$$V(f,\mathcal{P}) = \sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| \leq M,$$

pour toute partition $\mathcal{P}$ de $\bar{\mathcal{I}}$ de la forme $0 = x_0 < x_1 < \cdots < x_n = 1$. On appelle alors variation de f la quantité

$$V(f) = \sup_{\mathcal{P}\in\mathfrak{B}} V(f,\mathcal{P}),$$

où $\mathfrak{B}$ est l'ensemble des partitions de $\bar{\mathcal{I}}$.

On a une inégalité fondamentale, appelée inégalité de Koksma-Hlawka[27]:

Théorème 2.4.2. *Si $f\colon \bar{\mathcal{I}} \to \mathbb{R}$ est une fonction à variation bornée, alors pour tout ensemble $X = \{x_n\colon 1 \leq n \leq N\} \subset \bar{\mathcal{I}}$, on a*

$$\left| \frac{1}{N} \sum_{n=1}^{N} f(x_n) - \int_{\bar{\mathcal{I}}^s} f(u)du \right| \leq V(f) D_N^*(X).$$

Si f est une fonction continue sur $\bar{\mathcal{I}}$, on note $\omega(f;t)$ son module de continuité, défini par

$$\omega(f;t) = \sup_{\substack{x,y\in\bar{\mathcal{I}} \\ |x-y|\leq t}} |f(x) - f(y)|\,,\ t \geq 0.$$

Pour les fonctions continues, on a la majoration d'erreur suivante, due à H. Niederreiter [29]:

Proposition 2.4.3. *Si f est une fonction continue sur $\bar{\mathcal{I}}$ alors pour tout ensemble $X = \{x_n\colon 1 \leq n \leq N\} \subset \bar{\mathcal{I}}$, on a :*

$$\left| \frac{1}{N} \sum_{n=1}^{N} f(x_n) - \int_{\bar{\mathcal{I}}} f(u)\mathrm{d}u \right| \leq \omega(f; D_N^*(X)).$$

Dans le cas multidimensionnel nous devons étendre la notion de variation d'une fonction. Nous introduisons alors quelques notations et définitions. On trouve leur origine dans le livre de E. W. Hobson [23].

Pour une fonction $\varphi\colon \bar{\mathcal{I}}^s \to \mathbb{R}^p$, deux vecteurs $\boldsymbol{w}, \boldsymbol{w'} \in \bar{\mathcal{I}}^s$ et $1 \leq i \leq s$, on note $T^i_w\varphi$ la restriction de φ à l'hyperplan $x_i = w_i$ et

$$\Delta^i_{\boldsymbol{w},\boldsymbol{w'}}\varphi := T^i_{\boldsymbol{w'}}\varphi - T^i_{\boldsymbol{w}}\varphi.$$

Si $K = \{i_1, i_2, \ldots, i_k\}$ est un ensemble d'entiers compris entre 1 et s, on note

$$T^K_{\boldsymbol{w}}\varphi := T^{i_1}_{\boldsymbol{w}} \ldots T^{i_k}_{\boldsymbol{w}}\varphi \text{ et } \Delta^K_{\boldsymbol{w},\boldsymbol{w'}}\varphi := \Delta^{i_1}_{\boldsymbol{w},\boldsymbol{w'}} \ldots \Delta^{i_k}_{\boldsymbol{w},\boldsymbol{w'}}\varphi.$$

Dans le cas où $K = \{1, \ldots, s\}$, on note

$$T_{\boldsymbol{w}}\varphi := T^{\{1,\ldots,s\}}_{\boldsymbol{w}}\varphi \text{ et } \Delta_{\boldsymbol{w},\boldsymbol{w'}}\varphi := \Delta^{\{1,\ldots,s\}}_{\boldsymbol{w},\boldsymbol{w'}}\varphi.$$

Si l'on considère une partition de $\bar{\mathcal{I}}^s$ en sous-intervalles définie par

$$0 = w_{0,i} < w_{1,i} < \cdots < w_{n_i,i} = 1,\ 1 \leq i \leq s$$

et si $\boldsymbol{a} = (a_1, \ldots, a_s)$ est un vecteur d'entiers tels que $0 \leq a_i < n_i$, on note

$$\boldsymbol{w_a} := (w_{a_1,1}, \ldots, w_{a_s,s}) \text{ et } \boldsymbol{a}+ := (a_1 + 1, \ldots, a_s + 1).$$

Définition 2.4.4. *La variation au sens de Vitali de φ sur $\bar{\mathcal{I}}^s$ est définie par*

$$V^{(s)}(\varphi) := \sup_{\mathcal{P}\in\mathfrak{B}} \sum_{\boldsymbol{a}} |\Delta_{\boldsymbol{w_a},\boldsymbol{w_{a+}}}\varphi|,$$

où $\mathfrak{B}$ est l'ensemble des partitions de $\bar{\mathcal{I}}^s$.

On dit que φ est à variation bornée au sens de Vitali si $V^{(s)}(\varphi) < +\infty$.

Remarque 2.4.5. *Si la dérivée partielle $\frac{\partial^s \varphi}{\partial x_1 \ldots \partial x_s}$ est continue sur $\bar{\mathcal{I}}^s$, la variation au sens de Vitali s'écrit :*

$$V^{(s)}(\varphi) := \int_0^1 \ldots \int_0^1 \left| \frac{\partial^s \varphi}{\partial x_1 \ldots \partial x_s}(x_1, \ldots, x_s) \right| \mathrm{d}x_1 \ldots \mathrm{d}x_s.$$

La variation au sens de Vitali mesure imparfaitement l'ampleur des fluctuations de φ, car elle s'annule dès que φ ne dépend pas d'une variable. Une mesure plus précise est la variation au sens de Hardy-Krause, qui prend également en compte les fluctuations des restrictions de φ aux faces de $\bar{\mathcal{I}}^s$.

Définition 2.4.6. *La variation au sens de Hardy-Krause de φ sur $\bar{\mathcal{I}}^s$ est définie par*

$$V(\varphi) := \sum_{k=1}^{s} \sum_{\substack{K \subset [1,s] \\ \#K = k}} V^{(k)}(T_{\mathbf{1}}^{K^c} \varphi),$$

où $\mathbf{1} = (1, ..., 1) \in \bar{\mathcal{I}}^s$, K^c est le complémentaire de K dans $\{1, \ldots, s\}$ et $V^{(k)}(T_{\mathbf{1}}^{K^c} \varphi)$ est la variation au sens de Vitali k-dimensionnelle de $T_{\mathbf{1}}^{K^c} \varphi$ sur $\bar{\mathcal{I}}^k$.

On dit que φ est à variation bornée au sens de Hardy-Krause si $V(\varphi) < +\infty$.

Le résultat fondamental sur l'erreur des méthodes QMC est l'inégalité de Koksma-Hlawka, établie dans [22] et reprise dans [62], qui généralise en dimension quelconque l'inégalité de Koksma il est donné par le théorème.

Théorème 2.4.7. *Si φ est une fonction à variation bornée $V(\varphi)$ au sens de Hardy-*

Krause sur $\bar{\mathcal{I}}^s$, alors pour tout ensemble $X = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}\} \in \mathcal{I}^s$, on a

$$\left\| \frac{1}{N} \sum_{n=0}^{N-1} \varphi(\boldsymbol{x}_n) - \int_{\bar{\mathcal{I}}^s} \varphi(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \right\| \leq V(\varphi) D_N^*(X).$$

L'ensemble de toutes les fonctions bornées au sens de Hardy-Krause sur $\bar{\mathcal{I}}^s$ sera noté par $BVHK^s$.

Une version multidimensionelle de la proposition 2.4.3 existe. La majoration suivante est due à P.D. Proinov [51].

Proposition 2.4.8. *Si f est une fonction continue sur $\bar{\mathcal{I}}^s$, alors pour tout ensemble $X = \{\boldsymbol{x}_n : 1 \leq n \leq N\} \subset \bar{\mathcal{I}}^s$, on a*

$$\left| \frac{1}{N} \sum_{n=1}^{N} f(\boldsymbol{x}_n) - \int_{\bar{\mathcal{I}}}^{s} f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \right| \leq 4\omega(f; D_N^*(X)^{1/s}).$$

La borne d'erreur donnée par l'inégalité de Koksma-Hlawka est en général la meilleure possible; dans [46] H. Niederreiter a montré le résultat suivant:

Lemme 2.4.9. *Pour tout ensemble $X = \{\boldsymbol{x}_n : 1 \leq n \leq N\} \subset \mathcal{I}^s$ et pour tout $\epsilon > 0$ il existe une fonction infiniment dérivable $f \in \mathcal{C}^\infty(\bar{\mathcal{I}}^s)$ telle que $V(f) = 1$ et qui vérifie:*

$$\left| \frac{1}{N} \sum_{n=1}^{N} f(\boldsymbol{x}_n) - \int_{\bar{\mathcal{I}}^s} f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \right| > D_N^*(X) - \epsilon.$$

L'inégalité de Koksma-Hlawka montre que l'erreur d'une quadrature QMC dépend de la variation $V(f)$ de la fonction que l'on intègre et de la discrépance $D_N^*(X)$ de l'ensemble des points de quadrature. Puisqu'on utilise des ensembles à discrépance faible, la majoration de l'erreur de quadrature est d'ordre

$$\mathcal{O}\Big(\frac{(\ln N)^{s-1}}{N}\Big).$$

Elle converge plus rapidement que la longueur de l'intervalle de confiance de la méthode de Monte Carlo qui est d'ordre

$$\mathcal{O}\Big(\frac{1}{\sqrt{N}}\Big).$$

2.4.2 Autres majorations

L'inégalité de Koksma-Hlawka ne permet pas de majorer l'erreur de quadrature QMC de certaines fonctions simples; par exemple les fonctions indicatrices d'ensembles dont la frontière n'est pas formée de faces parallèles aux faces de coordonnées.

Si $E \subset \bar{\mathcal{I}}^s$ est mesurable au sens de Lebesgue, et si $X = \{\boldsymbol{x}_n \colon 1 \leq n \leq N\}$ est un ensemble fini de $\mathcal{I}^s$, on a

$$D_N(E, X) = \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_E(\boldsymbol{x}_n) - \int_{\bar{\mathcal{I}}^s} \mathbb{1}_E \mathrm{d}\lambda_s, \tag{2.20}$$

où $\mathbb{1}_E$ est la fonction indicatrice de E. Comme cette fonction n'est pas nécessairement à variation bornée au sens de Hardy-Krause, on ne peut pas toujours estimer le membre de droite de l'égalité 2.20 en utilisant l'inégalité de Koksma-Hlawka. Des majorations utiles ont été donnée par H. Niederreiter et J. M. Wills [42].

On considère l'ensemble des parties de $\bar{\mathcal{I}}^s$ mesurables au sens de Jordan, c'est-à-dire les sous-ensembles de E de $\bar{\mathcal{I}}^s$ tels que $\mathbb{1}_E$ soit intégrable au sens de Riemann. Pour $E \subset \bar{\mathcal{I}}^s$ et $\varepsilon > 0$, on définit :

$$\begin{aligned} E_\varepsilon &:= \{\boldsymbol{x} \in \bar{\mathcal{I}}^s \colon \exists \boldsymbol{y} \in E \ \|\boldsymbol{x} - \boldsymbol{y}\|_2 < \varepsilon\}, \\ E_{-\varepsilon} &:= \{\boldsymbol{x} \in \bar{\mathcal{I}}^s \colon \forall \boldsymbol{y} \in \bar{\mathcal{I}}^s \backslash E \ \|\boldsymbol{x} - \boldsymbol{y}\|_2 \geq \varepsilon\}, \end{aligned} \tag{2.21}$$

où $\|.\|_2$ est la norme euclidienne de $\mathbb{R}^s$.

Soit σ une fonction positive et croissante, définie sur $\mathbb{R}_+^*$ vérifiant

$$\lim_{\varepsilon \to 0^+} \sigma(\varepsilon) = 0$$

On note $\mathcal{M}_\sigma$ la famille des sous-ensembles $E \subset \bar{\mathcal{I}}^s$ mesurables au sens de Lebesgue vérifiant :

$$\forall \varepsilon > 0;\ \max\Big(\lambda_s(E_\varepsilon \backslash E), \lambda_s(E \backslash E_{-\varepsilon})\Big) \leq \sigma(\varepsilon).$$

Alors $\mathcal{M}_\sigma$ est formé de sous-ensembles mesurables au sens de Jordan. On a les résultat suivant :

Théorème 2.4.10. *Soit X un ensemble de N points de $\bar{\mathcal{I}}^s$. Pour tout $E \in \mathcal{M}_\sigma$, on a :*

$$\Big|D_N(E,X)\Big| \leq 2\sigma\Big(\sqrt{s}\Big\lfloor \frac{1}{D_N(X)^1/s}\Big\rfloor^{-1}\Big) + \sigma\Big(\Big\lfloor \frac{1}{D_N(X)^{1/s}}\Big\rfloor^{-1}\Big) + D_N(X)^{1/s}.$$

Ce résultat a été repris et complété par C. Lécot [29], dont nous suivons la présentation.

En remplaçant la norme $\|.\|_2$ par la norme $\|.\|_\infty$ dans (2.21), on obtient l'inégalité suivante :

$$|D_N(E,X)| \leq 3\sigma\left(\Big\lfloor \frac{1}{D_N(X)^{1/s}}\Big\rfloor^{-1}\right) + D_N(X)^{1/s}. \tag{2.22}$$

Et si X est un réseau-(t,m,s) en base b, on a:

$$|D_N(E,X)| \leq \sigma\big(b^{-\lfloor \frac{m-t}{s} \rfloor}\big). \tag{2.23}$$

On peut s'interésser à d'autres domaines d'intégration spécifiques.

Soit $f\colon \bar{\mathcal{I}}^{s-1} \to \bar{\mathcal{I}}$ une fonction à variation bornée au sens de Hardy-Krause; on définit :

$$E_f := \{(\boldsymbol{x}', x_s) \in \mathcal{I}^s \colon x_s < f(\boldsymbol{x}')\},$$

où $\boldsymbol{x}' = (x_1, ..., x_{s-1})$.

Théorème 2.4.11. *Soit $f\colon \bar{\mathcal{I}}^{s-1} \to \bar{\mathcal{I}}$ une fonction à variation bornée $V(f)$ au sens de Hardy-Krause et soit X un ensemble de N points de $\mathcal{I}^s$. Si $D_N(X) \leq V(f)$, alors*

$$|D_N(E_f, X)| \leq sV(f)\left[\left(\frac{V(f)}{D_N(X)}\right)^{1/s}\right]^{-1}.$$

Pour un réseau-(t, m, s) en base b, on a:

Théorème 2.4.12. *Soit $f\colon \bar{\mathcal{I}}^{s-1} \to \bar{\mathcal{I}}$ une fonction à variation bornée $V(f)$ au sens de Hardy-Krause et soit X un réseau$-t, m, s)$ en base b. Si $b^{t-m} \leq V(f)$, alors*

$$|D_N(E_f, X)| \leq sV(f)b^{-\lfloor \frac{m-t}{s} + \frac{\ln V(f)}{s \ln b} \rfloor}.$$

Toutes ces inégalités assurent une convergence bien plus lente que celle donnée par l'inégalité de Koksma-Hlawka, car on a remplacé une erreur en $\mathcal{O}\Big(\frac{(\log N)^{s-1}}{N}\Big)$ par une erreur en $\mathcal{O}\Big(\frac{(\log N)^{(s-1)/s}}{N^{1/s}}\Big)$.

2.4.3 Exemple de calcul approché d'intégrale

Nous testons les quadratures Monte Carlo et quasi-Monte Carlo sur l'intégrale suivante:

$$K = \int_{\bar{\mathcal{I}}^3} \frac{|x_1 - x_2|}{1 + x_2 x_3} \mathrm{d}\boldsymbol{x},$$

notons que $K = 5/4 - 2\log(2) + \frac{\pi^2}{24}$. Nous présentons dans la figure 2.11 les différentes erreurs des deux méthodes, dans la quadrature quasi-Monte Carlo, on utilise les suites de Halton en bases $2, 3,$ et 5 ainsi que les suites de Hammersley. Le graphe montre de meilleurs résultats pour les quadratures quasi-Monte Carlo que pour les quadratures Monte Carlo.

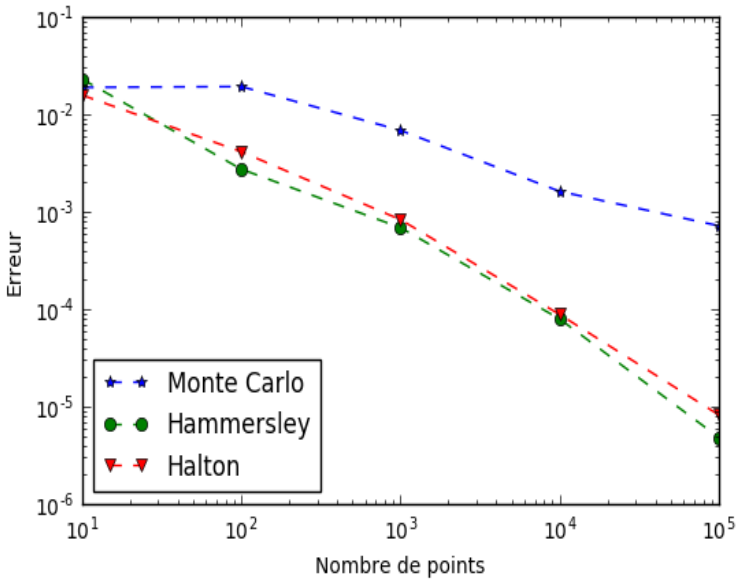


Figure 2.11: Erreurs dans l'approximation de K

Chapitre 3

Analyse numérique des méthodes de Runge-Kutta quasi-Monte Carlo

Dans ce chapitre nous décrivons une famille de méthodes apparentée à la famille des méthodes de Runge-Kutta pour résoudre un système d'équations différentielles ordinaires $y'(t) = f(t, y(t))$ où f est régulière en y mais manquant de régularité en t. Ces méthodes utilisent des évaluations Monte Carlo d'intégrales. Nous présentons le schéma d'ordre 3 qui utilise des échantillons aléatoires en dimension 3. Nous donnons des bornes d'erreur en fonction du pas de temps et de la discrépance de la suite utilisée pour les approximations Monte Carlo. Nous donnons des résultats numériques de tests faits sur un problème modèle dans lequel f subit des variations

rapides en temps. Nous concluons que l'utilisation de suites quasi-aléatoires réduit les erreurs. Ce chapitre est basé sur [5].

3.1 Introduction

Considérons le système différentiel

$$\begin{cases} y'(t) &= f(t, y(t)), \quad 0 < t < T, \\ y(0) &= y_0, \end{cases} \tag{3.1}$$

f étant régulière en espace (y) et seulement mesurable et bornée en temps (t). La méthode a été introduite par G. Stengle dans [54], [55] qui a présenté une famille de schémas numériques utilisant des nombres aléatoires pour résoudre le système (3.1) quand f varie beaucoup plus vite en temps qu'en espace. Etant donné que la régularité en t n'est pas assurée, la dépendance temporelle est considéré seulement en moyenne. En estimant les moyennes par simulation numérique utilisant des nombres aléatoires, on obtient des méthodes de Runge-Kutta Monte Carlo (RKMC).

Nous proposons dans ce chapitre de remplacer les suites pseudo-aléatoires par des suites quasi-aléatoires judicieusement choisies dans le but d'améliorer les résultats. Les approches Monte Carlo et quasi-Monte Carlo ont été comparées par plusieurs Scientifiques dans les Conférences Monte Carlo and Quasi-Monte Carlo methods [6], [11], [50]. Les méthodes de Runge-Kutta quasi-Monte Carlo (RKQMC) d'ordre 1 et 2 ont été étudiées dans des travaux antérieurs [7], [29]. Dans cette partie, nous étudions plus exactement une méthode RKQMC d'ordre 3. Le plan du chapitre

est comme suit : La méthode RKQMC d'ordre 3 est décrite au paragraphe 3.2. Au paragraphe 3.3 nous étudions la convergence de la méthode. Ensuite, dans le paragraphe 3.4 nous présentons les résultats des expériences numériques obtenus grâce au langage Python; elles permettent de comparer les méthodes RK, RKMC et RKQMC. Enfin nous donnons une conclusion au pragraphe 3.5.

3.2 Méthodes de Runge-Kutta quasi-Monte Carlo

Considérons le problème (3.1) où $y_0 \in \mathbb{R}^p$. On suppose l'existence et l'unicité d'une solution $y(t)$ telle que $y' \in L^1(0,T)$, ainsi

$$\forall t_0,\ t_0+h \in [0,T];\ y(t_0+h) = y(t_0) + \int_{t_0}^{t_0+h} y'(u)\mathrm{d}u. \tag{3.2}$$

En supposant que f est suffisamment régulière en y (on précisera les hypothèses plus loin) on a

$$\begin{aligned} y(t_0+h) = y(t_0) &+ \int_{t_0}^{t_0+h} F_1(u_1; y(t_0))\mathrm{d}u_1 \\ &+ \int_{t_0}^{t_0+h}\int_{t_0}^{u_1} F_2(u_1,u_2; y(t_0))\mathrm{d}u_2\mathrm{d}u_1 + \dots \\ &+ \int_{t_0}^{t_0+h}\int_{t_0}^{u_1}\dots\int_{t_0}^{u_{s-2}} F_{s-1}(u_1,\dots,u_{s-1}; y(t_0))\mathrm{d}u_{s-1}\dots\mathrm{d}u_1 \\ &+ \int_{t_0}^{t_0+h}\int_{t_0}^{u_1}\dots\int_{t_0}^{u_{s-1}} F_s(u_1,\dots,u_s; y(u_s))\mathrm{d}u_s\dots\mathrm{d}u_1, \end{aligned} \tag{3.3}$$

où les fonctions F_i vérifient la relation de récurrence

$$\begin{cases} F_0(y) := y, \\ F_i(u_1, \ldots, u_i; y) := D_y^1 F_{i-1}(u_1, \ldots, u_{i-1}; y).f(u_i, y), \quad 1 \leq i \leq s. \end{cases}$$

Par suite

$$\begin{aligned} y(t) = F_0(y_0)) + \int_{t_0}^{t_0+h} F_1(u_1; y(t_0))\mathrm{d}u_1 \\ + \int_{t_0}^{t_0+h} \int_{t_0}^{u_1} F_2(u_1, u_2; y(t_0))\mathrm{d}u_2\mathrm{d}u_1 + \ldots \\ + \int_{t_0}^{t_0+h} \int_{t_0}^{u_1} \ldots \int_{t_0}^{u_{s-1}} F_s(u_1, \ldots, u_s; y(t_0))\mathrm{d}u_s \ldots \mathrm{d}u_1 + \mathcal{O}(h^{s+1}), \end{aligned} \quad (3.4)$$

On se ramène au même domaine d'intégration pour toutes les intégrales $(t_0, t_0 + h)^s$ et ceci pour pouvoir faire une approximation Monte Carlo. Soit

$$\begin{cases} G_0(y) := 0, \\ G_i(u_1, \ldots, u_i; y) := G_{i-1}(u_2, \ldots, u_i; y) + G_{i-1}(u_1, u_3, \ldots, u_i; y) + \\ + \cdots + G_{i-1}(u_1, u_2, \ldots, u_{i-1}; y) + h^{i-1} F_i(u_1, \ldots, u_i; y); \ 1 \leq i \leq s. \end{cases} \quad (3.5)$$

On obtient alors

$$\begin{aligned} \int_{t_0}^{t_0+h} F_1(u_1; y)du_1 + \int_{t_0}^{t_0+h} \int_{t_0}^{u_1} F_2(u_1, u_2; y)\mathrm{d}u_2\mathrm{d}u_1 + \ldots \\ + \int_{t_0}^{t_0+h} \int_{t_0}^{u_1} \ldots \int_{t_0}^{u_{s-1}} F_s(u_1, \ldots, u_s; y)\mathrm{d}u_s \ldots \mathrm{d}u_1 \\ = \tfrac{1}{s!h^{s-1}} \int_{t_0}^{t_0+h} G_s(\underline{u}_1, \ldots, \underline{u}_s; y)\mathrm{d}u_s \ldots \mathrm{d}u_1, \end{aligned}$$

où $\{\underline{u}_1, \ldots, \underline{u}_s\} = \{u_1, \ldots, u_s\}$ avec $\underline{u}_1 \geq \cdots \geq \underline{u}_s$.

Par conséquent l'équation (3.4) peut être exprimée par

$$\begin{aligned} y(t_0 + h) = y(t_0) \\ + \frac{1}{s!h^{s-1}} \int_{t_0}^{t_0+h} \ldots \int_{t_0}^{t_0+h} G_s(\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_s; y(t_0))\mathrm{d}u_s \ldots \mathrm{d}u_1 + \mathcal{O}(h^{s+1}). \end{aligned} \quad (3.6)$$

Nous limitons notre analyse aux méthodes d'ordre trois . L'équation (3.6) s'écrit alors

$$y(t_0+h) = y(t_0) \tag{3.7}$$
$$+\frac{1}{6h^2}\int_{t_0}^{t_0+h}\int_{t_0}^{t_0+h}\int_{t_0}^{t_0+h}\Bigg(2f(\bar{u}_1,y(t_0))+2f(\bar{u}_2,y(t_0))+2f(\bar{u}_3,y(t_0))$$
$$+hD_y^1f(\bar{u}_2,y(t_0)).f(\bar{u}_1,y(t_0))+hD_y^1f(\bar{u}_3,y(t_0)).f(\bar{u}_1,y(t_0))$$
$$+hD_y^1f(\bar{u}_3,y(t_0)).f(\bar{u}_2,y(t_0))$$
$$+h^2D_y^1f(\bar{u}_3,y(t_0)).\big(D_y^1f(\bar{u}_2,y(t_0)).f(\bar{u}_1,y(t_0))\big)$$
$$+h^2D_y^2f(\bar{u}_3,y(t_0)).\big(f(\bar{u}_1,y(t_0)),f(\bar{u}_2,y(t_0))\big)\Bigg)\mathrm{d}\boldsymbol{u}+\mathcal{O}(h^4),$$

où

$$\{\bar{u}_1,\bar{u}_2,\bar{u}_3\}=\{u_1,u_2,u_3\} \text{ avec } \bar{u}_1\leq\bar{u}_2\leq\bar{u}_3. \tag{3.8}$$

Afin d'obtenir des schémas proches des méthodes de Runge-Kutta, nous combinons des développements de Taylor. L'identité suivante a été proposée par G. Stengle voir [54]:

$$2f(u_1,y)+2f(u_2,y)+2f(u_3,y) \tag{3.9}$$

$$+hD_y^1f(u_2,y).f(u_1,y)+hD_y^1f(u_3,y).f(u_1,y)$$

$$+hD_y^1f(u_3,y).f(u_2,y)$$

$$+h^2D_y^1f(u_3,y).(D_y^1f(u_2,y).f(u_1,y))$$

$$+h^2D_y^2f(u_3,y).(f(u_1,y),f(u_2,y))$$

$$=2f(u_1,y)+\Big(2-\frac{1-\kappa^2}{\alpha}\Big)f(u_2,y)+\frac{1}{\alpha}f\Big(u_2,y+\frac{\alpha}{1-\kappa}hf(u_1,y)\Big)$$

$$-\frac{\kappa^2}{\alpha}f\Big(u_2,y+\frac{\alpha}{\kappa(1-\kappa)}hf(u_1,y)\Big)+\Big(2-\frac{1-\lambda^2-\mu^2}{(1-\lambda)(1-\mu)}\Big)f(u_3,y)$$

$$-\frac{\mu^2}{(1-\lambda)(1-\mu)}f\Big(u_3,y+\frac{1-\lambda}{\mu}hf(u_1,y)\Big)-\frac{\lambda^2}{(1-\lambda)(1-\mu)}f\Big(u_3,y+\frac{1-\mu}{\lambda}hf(u_2,y)\Big)$$

$$+\frac{1}{(1-\lambda)(1-\mu)}f\Big(u_3,y+(1-\lambda)hf(u_1,y)+(1-\mu)hf(u_2,y+(1-\lambda)hf(u_1,y))\Big)+\mathcal{O}(h^3).$$

Nous utiliserons plutôt l'identité plus générale

$$2f(u_1,y)+2f(u_2,y)+2f(u_3,y) \tag{3.10}$$

$$+hD_y^1f(u_2,y).f(u_1,y)+hD_y^1f(u_3,y).f(u_1,y)$$

$$+hD_y^1f(u_3,y).f(u_2,y)$$

$$+h^2D_y^1f(u_3,y).(D_y^1f(u_2,y).f(u_1,y))$$

$$+h^2D_y^2f(u_3,y).(f(u_1,y),f(u_2,y))$$

$$=a_1f(u_1,y)+\sum_{l=1}^{L_2}a_{2,l}f(u_2,y+b_{2,l}hf(u_1,y))$$

$$+\sum_{l=1}^{L_3}a_{3,l}f\Big(u_3,y+b_{3,l}^{(1)}hf(u_1,y)+b_{3,l}^{(2)}hf(u_2,y+c_{3,l}hf(u_1,y))\Big)+\mathcal{O}(h^3),$$

où les coefficients $a_1, a_{2,l}, a_{3,l}, b_{2,l}, b^{(1)}_{3,l}, b^{(2)}_{3,l}, c_{3,l}$ vérifient les relations suivantes

$$
\begin{aligned}
&a_1 = 2,\ \sum_{l=1}^{L_2} a_{2,l} = 2,\ \sum_{l=1}^{L_3} a_{3,l} = 2,\\
&\sum_{l=1}^{L_2} a_{2,l} b_{2,l} = 1,\ \sum_{l=1}^{L_3} a_{3,l} b^{(1)}_{3,l} = 1,\ \sum_{l=1}^{L_3} a_{3,l} b^{(2)}_{3,l} = 1,\\
&\sum_{l=1}^{L_2} a_{2,l} b^2_{2,l} = 0,\ \sum_{l=1}^{L_3} a_{3,l} b^{(1)2}_{3,l} = 0,\ \sum_{l=1}^{L_3} a_{3,l} b^{(2)2}_{3,l} = 0,\\
&\sum_{l=1}^{L_3} a_{3,l} b^{(1)}_{3,l} b^{(2)}_{3,l} = 1,\ \sum_{l=1}^{L_3} a_{3,l} b^{(2)}_{3,l} c_{3,l} = 1.
\end{aligned}
$$

Ces équations ressemblent aux conditions d'ordre dans les méthodes de Runge-Kutta [4]. Nous remplaçons alors (3.7) par l'égalité suivante

$$
\begin{aligned}
y(t_0+h) = y(t_0) + \frac{1}{6h^2} \int_{t_0}^{t_0+h} \int_{t_0}^{t_0+h} \int_{t_0}^{t_0+h} \Bigg(& a_1 f(\bar{u}_1, y(t_0)) \\
& + \sum_{l=1}^{L_2} a_{2,l} f\Big(\bar{u}_2, y(t_0) + b_{2,l} h f(\bar{u}_1, y(t_0))\Big) \\
& + \sum_{l=1}^{L_3} a_{3,l} f\Big(\bar{u}_3, y(t_0) + b^{(1)}_{3,l} h f(\bar{u}_1, y(t_0)) \\
& + b^{(2)}_{3,l} h f\big(\bar{u}_2, y(t_0) + c_{3,l} h f(\bar{u}_1, y(t_0))\big)\Big)\Bigg) \mathrm{d}\mathbf{u} + \mathcal{O}(h^4).
\end{aligned} \tag{3.11}
$$

Soit $B(y, \rho)$ la boule ouverte centrée en y et de rayon ρ. Nous supposons que la fonction f vérifie les hypothèses suivantes.

Hypothèses 3.2.1. *Il existe $\tau > 0$ et $\rho > 0$ tels que*

- *La fonction $D^m_y f$ est mesurable sur*

$$
E := \bigcup_{0 \le t \le T} [t, \min(t+\tau, T)] \times B\Big(y(t), \rho\Big)
$$

  pour $0 \le m \le 3$.

- *Pour tout* $t \in [0,T]$ *la fonction* $y \mapsto D_y^m f(t,y)$ *est continue sur la boule ouverte* $B(y(t),\rho)$, *pour* $0 \le m \le 3$.

- *Il existe* $\|D_y^m f\|$ *pour* $0 \le m \le 3$ *et* $V_E(D_y^m f)$ *pour* $0 \le m \le 2$, *tels que pour tout* $t \in [0,T]$ *et tout* $y \in B(y(t),\rho)$,

  1. *la fonction* $u \mapsto D_y^m f(u,y)$ *est définie sur* $[t, \min(t+\tau, T)]$ *et est bornée par* $\|D_y^m f\|_E$ *pour* $0 \le m \le 3$,
  2. *la variation de la fonction* $u \mapsto D_y^m f(u,y)$ *sur* $[t, \min(t+\tau, T)]$ *est bornée par* $V_E(D_y^m f)$ *pour* $0 \le m \le 2$.

Considérons une partition $0 = t_0 < t_1 < \cdots < t_{n^*} = T$ de $[0,T]$ en n^* sous-intervalles de longueur $h_n := t_{n+1} - t_n$ et notons $H := \max_{0 \le n \le n^*} h_n$. Soit X un ensemble de points composé de $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N-1} \in \mathcal{I}^3$. La méthode RKQMC d'ordre 3 engendre une suite $(y_n)_{0 \le n \le n^*}$ par

$$
\begin{aligned}
y_{n+1} = y_n + \frac{h_n}{6N} \sum_{0 \le j < N} \Bigg(& a_1 f(t_n + h_n \bar{x}_{j,1}, y_n) \\
& + \sum_{l=1}^{L_2} a_{2,l} f(t_n + h_n \bar{x}_{j,2}, y_n + b_{2,l} h_n f(t_n + h_n \bar{x}_{j,1}, y_n)) \\
& + \sum_{l=1}^{L_3} a_{3,l} f\Big(t_n + h_n \bar{x}_{j,3}, y_n + b_{3,l}^{(1)} h_n f(t_n + h_n \bar{x}_{j,1}, y_n) \\
& + b_{3,l}^{(2)} h_n f(t_n + h_n \bar{x}_{j,2}, y_n + c_{3,l} h_n f(t_n + h_n \bar{x}_{j,1}, y_n))\Big)\Bigg),
\end{aligned}
\tag{3.12}
$$

où nous utilisons la même notation que (3.8): si $\boldsymbol{x} = (x_1, x_2, x_3)$, alors $\bar{\boldsymbol{x}} = (\bar{x}_1, \bar{x}_2, \bar{x}_3)$ est défini comme suit

$$
\{\bar{x}_1, \bar{x}_2, \bar{x}_3\} = \{x_1, x_2, x_3\} \text{ avec } \bar{x}_1 \le \bar{x}_2 \le \bar{x}_3. \tag{3.13}
$$

Nous avons remplacé l'intégrale triple (3.11) par son approximation quasi-Monte Carlo.

3.3 Analyse de la convergence

L'analyse de l'erreur est assez proche de celle des méthodes de Runge-Kutta [4], l'erreur globale est définie par $e_n := y_n - y(t_n)$. On définit l'erreur de troncature locale par

$$\varepsilon_n := \frac{1}{h_n}(y(t_{n+1}) - y(t_n)) - \frac{1}{6h_n^3}\int_{(t_n,t_{n+1})^3} \Bigg(a_1 f(\bar{u}_1, y(t_n)) + \sum_{l=1}^{L_2} a_{2,l} f\Big(\bar{u}_2, y(t_n) + b_{2,l} h_n f(\bar{u}_1, y(t_n))\Big) + \sum_{l=1}^{L_3} a_{3,l} f\Big(\bar{u}_3, y(t_n) + b_{3,l}^{(1)} h_n f(\bar{u}_1, y(t_n)) + b_{3,l}^{(2)} h_n f\big(\bar{u}_2, y(t_n) + c_{3,l} h_n f(\bar{u}_1, y(t_n))\big)\Big)\Bigg) \mathrm{d}\boldsymbol{u}.$$

Nous avons besoin d'un terme d'erreur complémentaire

$$\delta_n := \frac{1}{6h_n^3}\int_{(t_n,t_{n+1})^3} \Bigg(a_1 f(\bar{u}_1, y_n) + \sum_{l=1}^{L_2} a_{2,l} f\Big(\bar{u}_2, y_n + b_{2,l} h_n f(\bar{u}_1, y_n)\Big) + \sum_{l=1}^{L_3} a_{3,l} f\Big(\bar{u}_3, y_n + b_{3,l}^{(1)} h_n f(\bar{u}_1, y_n) + b_{3,l}^{(2)} h_n f\big(\bar{u}_2, y_n + c_{3,l} h_n f(\bar{u}_1, y_n)\big)\Big) - a_1 f(\bar{u}_1, y(t_n)) - \sum_{l=1}^{L_2} a_{2,l} f\Big(\bar{u}_2, y(t_n) + b_{2,l} h_n f(\bar{u}_1, y(t_n))\Big) - \sum_{l=1}^{L_3} a_{3,l} f\Big(\bar{u}_3, y(t_n) + b_{3,l}^{(1)} h_n f(\bar{u}_1, y(t_n)) + b_{3,l}^{(2)} h_n f\big(\bar{u}_2, y(t_n) + c_{3,l} h_n f(\bar{u}_1, y(t_n))\big)\Big)\Bigg) \mathrm{d}\boldsymbol{u},$$

L'erreur de l'approximation quasi-Monte Carlo est donnée par

$$
\begin{aligned}
d_n := \frac{1}{6N} \sum_{0\le j<N} \Bigg(& a_1 f(t_n + h_n \bar{x}_{j,1}, y_n) \\
& + \sum_{l=1}^{L_2} a_{2,l} f\Big(t_n + h_n \bar{x}_{j,2}, y_n + b_{2,l} h_n f(t_n + h_n \bar{x}_{j,1}, y_n)\Big) \\
& + \sum_{l=1}^{L_3} a_{3,l} f\Big(t_n + h_n \bar{x}_{j,3}, y_n + b_{3,l}^{(1)} h_n f(t_n + h_n \bar{x}_{j,1}, y_n) \\
& + b_{3l}^{(2)} h_n f\big(t_n + h_n \bar{x}_{j,2}, y_n + c_{3,l} h_n f(t_n + h_n \bar{x}_{j,1}, y_n)\big)\Big)\Bigg) \\
& - \frac{1}{6} \int_{I^3} \Bigg(a_1 f(t_n + h_n \bar{x}_1, y_n) \\
& + \sum_{l=1}^{L_3} a_{2,l} f\Big(t_n + h_n \bar{x}_2, y_n + b_{2l} h_n f(t_n + h_n \bar{x}_1, y_n)\Big) \\
& + \sum_{l=1}^{L_3} a_{3,l} f\Big(t_n + h_n \bar{x}_3, y_n + b_{3,l}^{(1)} h_n f(t_n + h_n \bar{x}_1, y_n) \\
& + b_{3,l}^{(2)} h_n f\big(t_n + h_n \bar{x}_2, y_n + c_{3,l} h_n f(t_n + h_n \bar{x}_1, y_n)\big)\Big)\Bigg) \mathrm{d}\boldsymbol{x}.
\end{aligned}
$$

On a la formule de récurrence

$$e_{n+1} = e_n - h_n \varepsilon_n + h_n \delta_n + h_n d_n. \tag{3.14}$$

À toute fonction φ définie sur $\bar{\mathcal{I}}^3$, nous associons la fonction $\bar{\varphi}$ définie par

$$\bar{\varphi}(\boldsymbol{x}) := \varphi(\bar{\boldsymbol{x}}),\ \boldsymbol{x} \in \bar{\mathcal{I}}^3$$

où $\bar{\boldsymbol{x}}$ correspond à $\boldsymbol{x}$ par (3.13). Nous avons besoin des résultats techniques suivants.

Lemme 3.3.1. *Si $\varphi \in BVHK^3$, alors $\bar{\varphi} \in BVHK^3$.*

Preuve. On a

$$V(\bar{\varphi}) = 3V^{(1)}(T_1^{\{2,3\}}\varphi) + 3V^{(2)}(\overline{T_1^3\varphi}) + V^{(3)}(\bar{\varphi}).$$

D'une part, pour toute fonction ψ de $\bar{\mathcal{I}}^2$ on a

$$V^{(2)}(\bar{\psi}) \leq V^{(1)}(T_1^1\psi) + V^{(1)}(T_1^2\psi) + 3V^{(2)}(\psi),$$

où

$$\bar{\psi}(x_1, x_2) := \psi(\min(x_1, x_2), \max(x_1, x_2)).$$

D'autre part,

$$\begin{aligned} V^{(3)}(\bar{\varphi}) \leq\ & V^{(1)}(T_1^{\{2,3\}}\varphi) + 2V^{(1)}(T_1^{\{1,3\}}\varphi) + V^{(1)}(T_1^{\{1,2\}}\varphi) \\ & + 5V^{(2)}(T_1^1\varphi) + 7V^{(2)}(T_1^2\varphi) + 4V^{(2)}(T_1^3\varphi) + 13V^{(3)}(\varphi), \end{aligned}$$

ce qui achève la démonstration. □

Si $t \in [0, T],\ h \geq 0$ et $y \in \mathbb{R}^p$, nous définissons

$$\begin{aligned} \varphi_{t,h,y}(\boldsymbol{x}) :=\ & a_1 f(t + hx_1, y) \\ & + \sum_{l=1}^{L_2} a_{2,l} f\Big(t + hx_2, y + b_{2,l} hf(t + hx_1, y)\Big) \\ & + \sum_{l=1}^{L_3} a_{3,l} f\Big(t + hx_3, y + b_{3,l}^{(1)} hf(t + hx_1, y) \\ & + b_{3,l}^{(2)} hf\big(t + hx_2, y + c_{3,l} hf(t + hx_1, y)\big)\Big), \quad \forall \boldsymbol{x} \in \bar{\mathcal{I}}^3. \end{aligned}$$

Notons

$$c^* := 1 + \max\Big(\max_{1\leq l\leq L_2} |b_{2,l}|, \max_{1\leq l\leq L_3} |b_{3,l}^{(1)}| + \max_{1\leq l\leq L_3} |b_{3,l}^{(2)}|, \max_{1\leq l\leq L_3} |c_{3,l}|\Big).$$

Lemme 3.3.2. *Si t et $t + h \in [0, T], h \leq \tau$ et $\|y - y(t)\| + hc^*\|f\|_E < \rho$, alors $\varphi_{t,h,y} \in BVHK^3$.*

Preuve. Soit

$$\alpha_{t,h,y}(\boldsymbol{x}) := f(t + hx_1, y)$$

$$\beta^b_{t,h,y}(\boldsymbol{x}) := f(t + hx_2, y + bhf(t + hx_1, y))$$

$$\gamma^{b^1,b^2,c}_{t,h,y}(\boldsymbol{x}) := f(t + hx_3, y + b^1 hf(t + hx_1, y)$$

$$+b^2 hf(t + hx_2, y + chf(t + hx_1, y))).$$

Par des développements de Taylor en y on trouve sous l'hypothèse 3.2.1

$$V^{(1)}(T_1^{\{2,3\}}\alpha_{t,h,y}) \leq V_E(f),$$

$$V^{(1)}(T_1^{\{2,3\}}\beta^b_{t,h,y}) \leq h|b|\|D^1_y f\|_E V_E(f),$$

$$V^{(1)}(T_1^{\{1,3\}}\beta^b_{t,h,y}) \leq V_E(f),$$

$$V^{(2)}(T_1^{\{3\}}\beta^b_{t,h,y}) \leq h|b|V_E(D^1_y f)V_E(f),$$

$$V^{(1)}(T_1^{\{2,3\}}\gamma^{b^1,b^2,c}_{t,h,y}) \leq h(|b^1| + h|b^2 c|\|D^1_y f\|_E)\|D^1_y f\|_E V_E(f),$$

$$V^{(1)}(T_1^{\{1,3\}}\gamma^{b^1,b^2,c}_{t,h,y}) \leq h|b^2|\|D^1_y f\|_E V_E(f),$$

$$V^{(1)}(T_1^{\{1,2\}}\gamma^{b^1,b^2,c}_{t,h,y}) \leq V_E(f),$$

$$V^{(2)}(T_1^{\{1\}}\gamma^{b^1,b^2,c}_{t,h,y}) \leq h|b^2|V_E(D^1_y f)V_E(f),$$

$$V^{(2)}(T_1^{\{2\}}\gamma^{b^1,b^2,c}_{t,h,y}) \leq h(|b^1| + h|b^2 c|\|D^1_y f\|_E)V_E(D^1_y f)V_E(f),$$

$$V^{(2)}(T_1^{\{3\}}\gamma^{b^1,b^2,c}_{t,h,y}) \leq h^2\big(|b^2|(|b^1| + h|b^2 c|\|D^1_y f\|_E)\|D^2_y\|_E V_E(f)$$

$$+|b^2 c|\|D^1_y f\|_E V_E(D^1_y f)\big)V_E(f),$$

$$V^{(3)}(\gamma^{b^1,b^2,c}_{t,h,y}) \leq h^2\big(|b^2|(|b^1| + h|b^2 c|\|D^1_y f\|_E)V_E(D^2_y f)V_E(f)$$

$$+|b^2 c|V_E(D^1_y f)^2\big)V_E(f).$$

D'où le résultat. □

Définissons

$$\phi_n := \varphi_{t_n,h_n,y_n} \text{ et } \Phi_n := \varphi_{t_n,h_n,y(t_n)}.$$

Alors

$$\varepsilon_n = \frac{1}{h_n}(y(t_{n+1}) - y(t_n)) - \frac{1}{6}\int_{\mathcal{I}^3} \bar{\Phi}_n(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \tag{3.15}$$

$$\delta_n = \frac{1}{6}\int_{\mathcal{I}^3} (\bar{\phi}_n(\boldsymbol{x}) - \bar{\Phi}_n(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}, \tag{3.16}$$

$$d_n = \frac{1}{6N}\sum_{0\leq j<N} \bar{\phi}_n(\boldsymbol{x}_j) - \frac{1}{6}\int_{\mathcal{I}^3} \bar{\phi}_n(\boldsymbol{x})\mathrm{d}\boldsymbol{x}. \tag{3.17}$$

Nous établissons à présent des estimations de chacun de ces termes.

Proposition 3.3.3. *Si $h_n \leq \tau$ et $h_n c^* \|f\|_E < \rho$, alors il existe $c_1(h_n) = \mathcal{O}(1)$ telle que*

$$\|\varepsilon_n\| \leq c_1(h_n) h_n^3.$$

Preuve. Sous l'hypothèse 3.2.1 nous avons par des développements de Taylor en y

$$\begin{aligned}
y(t_{n+1}) = y(t_n) + \frac{1}{6h_n^2}\int_{(t_n,t_{n+1})^3} \\
\Bigg(2f(\bar{u}_1, y(t_n)) + 2f(\bar{u}_2, y(t_n)) + 2f(\bar{u}_3, y(t_n)) \\
+h_n D_y^1 f(\bar{u}_2, y(t_n)).f(\bar{u}_1, y(t_n)) \\
+h_n D_y^1 f(\bar{u}_3, y(t_n)).f(\bar{u}_1, y(t_n)) \\
+h_n D_y^1 f(\bar{u}_3, y(t_n)).f(\bar{u}_2; y(t_n)) \\
+h_n^2 D_y^1 f(\bar{u}_3, y(t_n)).\big(D_y^1 f(\bar{u}_2, y(t_n)).f(\bar{u}_1, y(t_n))\big) \\
+h_n^2 D_y^2 f(\bar{u}_3, y(t_n)).\big(f(\bar{u}_1, y(t_n)), f(\bar{u}_2, y(t_n))\big) \Bigg) \mathrm{d}\mathbf{u} + h_n \varepsilon_n^*,
\end{aligned}$$

où

$$\|\varepsilon_n^*\| \leq \frac{h_n^3}{24}\big(\|D_y^1 f\|_E^3 + 4\|D_y^2 f\|_E \|D_y^1 f\|_E \|f\|_E + \|D_y^3\|_E \|f\|^2\|_E\big)\|f\|_E,$$

et pour $\boldsymbol{u} \in (t_n, t_{n+1})^3$,

$$
\begin{aligned}
a_1 f(\bar{u}_1, y(t_n)) + \sum_{l=1}^{L_2} a_{2,l} f\Big(\bar{u}_2, y(t_n) + b_{2,l} h_n f(\bar{u}_1, y(t_n))\Big) \\
+ \sum_{l=1}^{L_3} a_{3,l} f\Big(\bar{u}_3, y(t_n) + b_{3,l}^{(1)} h_n f(\bar{u}_1, y(t_n)) \\
+ b_{3,l}^{(2)} h_n f\big(\bar{u}_2, y(t_n) + c_{3,l} h_n f(\bar{u}_1, y(t_n))\big)\Big) \\
= 2f(\bar{u}_1, y(t_n)) + 2f(\bar{u}_2, y(t_n)) + 2f(\bar{u}_3, y(t_n)) \\
+ h_n D_y^1 f(\bar{u}_2, y(t_n)).f(\bar{u}_1, y(t_n)) \\
+ h_n D_y^1 f(\bar{u}_3, y(t_n)).f(\bar{u}_1, y(t_n)) \\
+ h_n D_y^1 f(\bar{u}_3, y(t_n)).f(\bar{u}_2, y(t_n)) \\
+ h_n^2 D_y^1 f(\bar{u}_3, y(t_n)).(D_y^1 f(\bar{u}_2, y(t_n)).f(\bar{u}_1, y(t_n))) \\
+ h_n^2 D_y^2 f(\bar{u}_3, y(t_n)).(f(\bar{u}_1, y(t_n)), f(\bar{u}_2, y(t_n))) + \varepsilon_n'(\mathbf{u}),
\end{aligned}
$$

où $\|\varepsilon_n'(\mathbf{u})\| \leq \mathcal{O}(h_n^3)$. Par conséquent

$$
\varepsilon_n = \varepsilon_n^* - \frac{1}{6h_n^3} \int_{(t_n, t_{n+1})^3} \varepsilon_n'(\mathbf{u}) \mathrm{d}\mathbf{u}.
$$

Ceci achève la démonstration. □

Proposition 3.3.4. *Si $h_n \leq \tau$ et $\|e_n\| + h_n c^* \|f\|_E < \rho$, alors il existe $c_2(h_n) = \mathcal{O}(1)$ tel que*

$$
\|\delta_n\| \leq c_2(h_n) \|e_n\|.
$$

Preuve. Soit

$$\begin{aligned}
\delta_{n,1}(\mathbf{u}) &:= f(\bar{u}_1, y_n) - f(\bar{u}_1, y(t_n)),\\
\delta_{n,2,l}(\mathbf{u}) &:= f(\bar{u}_2, y_n + b_{2,l}h_n f(\bar{u}_1, y_n)) - f(\bar{u}_2, y(t_n) + b_{2,l}h_n f(\bar{u}_1, y(t_n)))\\
\delta_{n,3,l}(\mathbf{u}) &:= f(\bar{u}_3, y_n + b^{(1)}_{3,l}h_n f(\bar{u}_1, y_n) + b^{(2)}_{3,l}h_n f(\bar{u}_2, y_n + c_{3,l}h_n f(\bar{u}_1, y_n)))\\
&\quad - f\big(\bar{u}_3, y(t_n) + b^{(1)}_{3,l}h_n f(\bar{u}_1, y(t_n))\\
&\quad + b^{(2)}_{3,l}h_n f(\bar{u}_2, y(t_n) + c_{3,l}h_n f(\bar{u}_1, y(t_n)))\big).
\end{aligned}$$

Sous l'hypothèse 3.2.1 et par des développements de Taylor en y, on obtient

$$|\delta_{n,1}(\mathbf{u})\| \leq \mathcal{O}(1)\|e_n\| \text{ et } \|\delta_{n,i,l}(\mathbf{u})\| \leq \mathcal{O}(1)\|e_n\|,\ i = 2, 3,$$

ce qui implique le résultat. □

Proposition 3.3.5. *Si $h_n \leq \tau$ et $\|e_n\| + h_n c^* \|f\|_E < \rho$, alors il existe $c_3(h_n) = \mathcal{O}(1)$ telle que*

$$\|d_n\| \leq c_3(h_n) D_N^*(X).$$

Preuve. En utilisant l'inégalité de Koksma-Hlawka on a

$$\|d_n\| \leq \frac{1}{6} V(\bar{\phi}_n) D_N^*(X).$$

Par les Lemmes 3.3.1 and 3.3.2, nous concluons le résultat. □

Nous terminons ce paragraphe par le résultat de convergence. Supposons que $H \leq H^*$ et posons

$$c_i := c_i(H^*), i = 1, 2, 3.$$

Proposition 3.3.6. *Si* $H \leq \tau$ *et*

$$e^{c_2 T}\|e_0\| + \frac{e^{c_2 T}-1}{c_2}\Big(c_1 H^3 + c_3 D_N^*(X)\Big) + Hc^*\|f\|_E < \rho,$$

alors

$$\|e_n\| \leq e^{c_2 t_n}\|e_0\| + \frac{e^{c_2 t_n}-1}{c_2}\Big(c_1 H^3 + c_3 D_N^*(X)\Big), 0 \leq n \leq n^*.$$

Proof. En combinant la formule de récurrence (3.14) et les propositions 3.3.3, 3.3.4 et 3.3.5 on obtient le résultat voulu. □

Il est devenu d'usage de parler de suites à faible discrépance pour celles dont la discrépance est en $\mathcal{O}((\log N)^{s-1}/N)$, en dimension s où N désigne le nombre de points considérés. Dans [46] R. Niederreiter donne un sondage sur de nombreux ensembles de points à faible discrépance. Un résultat dû à K. F. Roth [52] montre que, pour toute suite de N éléments d'un ensemble X en dimension s, on a

$$D_N^*(X) \geq \gamma_s (\log N)^{(s-1)/2}/N,$$

où la constante $\gamma_s > 0$ dépend seulement de s. D'apès l'estimation de la proposition 3.3.6 l'algorithme RKQMC utilisant des suites à faible discrépance est d'ordre 3 si $N = \mathcal{O}(H^{-3})$.

3.4 Expériences numériques

Nous appliquons l'algorithme présenté à un problème modèle proposé par G. Stengle voir [55] ensuite sur un exemple simple.

Exemple 3.4.1.

$$y'(t) = y(t) + 5\sin(\cos(kt)),\ 0 < t < 1, \tag{3.18}$$

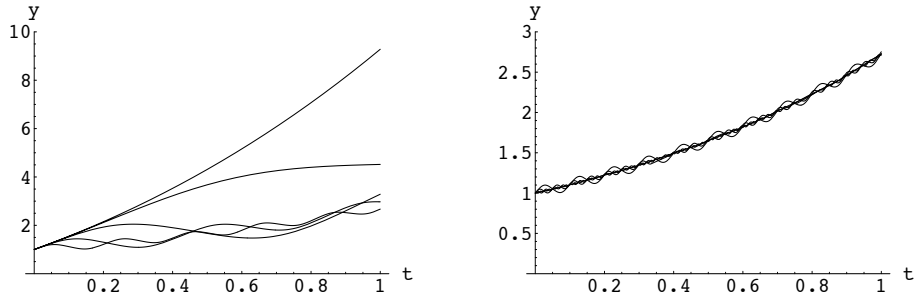
$$y(0) = 1, \tag{3.19}$$

pour $k = 2^\nu - 1,\ 1 \leq \nu \leq 20$. *La solution exacte est donnée par*

$$y(t) = e^t\Big[1 + 5t\int_0^1 e^{-rt}\sin(\cos(krt))\mathrm{d}r\Big],$$

$$y(t) = e^t + 5\sum_{m\in\mathbb{Z}} J_m(1)\frac{mk}{m^2k^2-1}\Bigg(\cos\Big[m(\pi/2-kt)\Big] - e^t\cos(m\pi/2) - \frac{1}{mk}\sin\big[m(\pi/2-kt)\big] + \frac{e^t}{mk}\sin(m\pi/2)\Bigg),$$

représentée dans la figure 3.1 pour $1 \leq \nu \leq 10$.



(a) Solution exacte pour $1 \leq \nu \leq 5$ (b) Solution exacte pour $6 \leq \nu \leq 10$

Figure 3.1: Solution exacte pour $1 \leq \nu \leq 10$

Nous résolvons le problème par les trois différents schémas présentés ci-dessous, ensuite nous comparons les erreurs obtenues:

- *La méthode de Heun avec* $10(RK10)$, $100(RK100)$ *et* $1000(RK1000)$ *pas de temps égaux, Le schéma de Heun d'ordre 3 étant décrit comme suit (voir [4]):*

$$
\begin{aligned}
t_{n,1} &= t_n, \\
t_{n,2} &= t_n + \tfrac{h_n}{3}, \\
t_{n,3} &= t_n + \tfrac{2h_n}{3}, \\
k_{n,1} &= f(t_{n,1}, y_n), \\
k_{n,2} &= f(t_{n,2}, y_n + \tfrac{h_n}{3} k_{n,1}), \\
k_{n,3} &= f(t_{n,3}, y_n + \tfrac{2h_n}{3} k_{n,2}), \\
y_{n+1} &= y_n + \tfrac{h_n}{4}(k_{n,1} + 3k_{n,3}).
\end{aligned}
$$

- *La méthode Runge Kutta Monte Carlo de G. Stengle avec 10 pas de temps égaux où à chaque pas, nous utilisons une nouvelle suite aléatoire* $(\boldsymbol{x}_j^{(n)})_{0 \le j < N}$:

$$
\begin{aligned}
t^*_{n,j,1} &= t_n + h_n \bar{x}^{(n)}_{j,1}, \\
t^*_{n,j,2} &= t_n + h_n \bar{x}^{(n)}_{j,2}, \\
t^*_{n,j,3} &= t_n + h_n \bar{x}^{(n)}_{j,3}, \\
k_{n,j,1} &= f(t^*_{n,j,1}, y_n), \\
k_{n,j,2} &= f(t^*_{n,j,2}, y_n), \\
k'_{n,j,2} &= f(t^*_{n,j,2}, y_n + h_n k_{n,j,1}), \\
k''_{n,j,2} &= f(t^*_{n,j,2}, y_n + \tfrac{h_n}{2} k_{n,j,1}), \\
k_{n,j,3} &= f(t^*_{n,j,3}, y_n + h_n k_{n,j,1}), \\
k'_{n,j,3} &= f(t^*_{n,j,3}, y_n + h_n k_{n,j,2}), \\
k''_{n,j,3} &= f(t^*_{n,j,3}, y_n + \tfrac{h_n}{2} k_{n,j,1} + \tfrac{h_n}{2} k''_{n,j,2}),
\end{aligned}
$$

et

$$y_{n+1} = y_n + \frac{1}{N}\sum_{0\le j<N} h_n\Big(\frac{1}{3}k_{n,j,1} - \frac{1}{6}k_{n,j,2} - \frac{1}{6}k'_{n,j,2} + \frac{2}{3}k''_{n,j,2} - \frac{1}{6}k_{n,j,3} - \frac{1}{6}k'_{n,j,3} + \frac{2}{3}k''_{n,j,3}\Big). \quad (3.20)$$

- *La méthode Runge Kutta quasi Monte Carlo avec 10 pas de temps égaux. À chaque pas de temps, nous utilisons dans l'équation (3.20) le même ensemble de Hammersley à 1000 points.*

  $$X = \{(\frac{j}{1000}, \phi_2(j), \phi_3(j)) : 0 \le j < 1000\},$$

  où ϕ_b est la fonction radical inverse en base b.

  Nous calculons l'erreur

  $$e = \frac{1}{10}\sum_{n=0}^{10} |e_{nm}|,$$

  où $10m$ est le nombre de pas de temps.

La figure 3.2 compare les erreurs obtenues en résolvant le problème modèle en utilisant les méthodes RK10, RK100, RK1000 et RKMC. On peut voir que les résultats de RKMC sont meilleurs que ceux de RK10 pour ($k \ge 2^5$), RK100 pour ($k \ge 2^9$) et RK1000 pour ($k \ge 2^{13}$). La figure 3.3 montre une comparaison des méthodes RK10, RK100, RK1000 et RKQMC au probème modèle. Les erreurs du schéma RKQMC se stabilisent à un niveau assez bas. La stratégie RKQMC est plus performante que le schéma RK10 pour ($k \ge 2^3$), RK100 pour ($k \ge 2^8$) et RK1000 pour ($k \ge 2^{11}$). Il est clair que pour cet exemple la stratégie quasi-Monte Carlo est nettement supérieure aux résultats de Monte Carlo.

G. Stengle [55] a noté que c'est grâce à la somme sur N dans l'équation (3.20) que le calcul RKMC ou RKQMC est avantageux. Les calculs de RKQMC utilisent 10000 sommation en comparant à 1000 pour la méthode de RK1000, mais la somme des 10000 peut être exécutée en dix étapes de temps contenant chacune 1000.

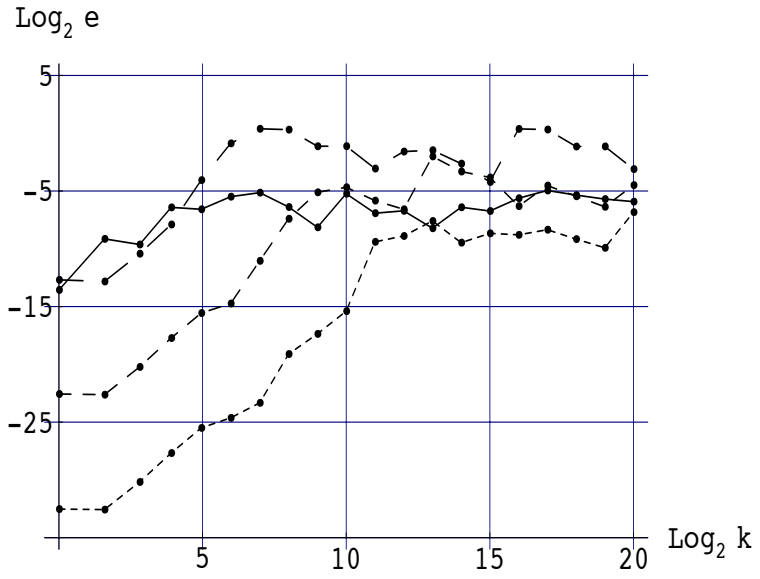


Figure 3.2: Erreurs RKMC (trait continu), RK10, RK100, RK1000 (traits pointillés)

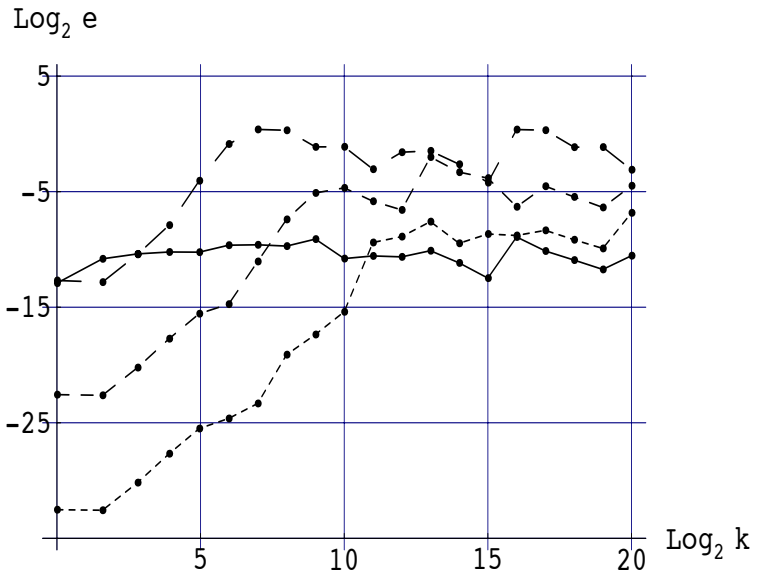


Figure 3.3: Erreurs RKQMC(trait continu), RK10, RK100, RK1000 (traits pointillés)

Exemple 3.4.2.

$$y'(t) = y(t) + \cos(1023t),\ 0 < t < 1, \tag{3.21}$$

$$y(0) = 1, \tag{3.22}$$

La solution exacte est donnée par:

$$y(t) = e^t + \frac{1}{1024}\Big(\frac{e^t}{1023} - \sin(1023t) - \frac{1}{1023}\cos(1023t)\Big).$$

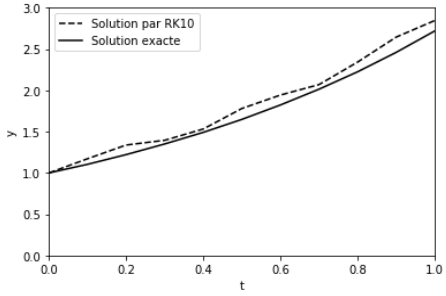


Figure 3.4: Solution exacte, solution par RK10

On compare les solutions par la méthode de Heun, Runge Kutta Monte Carlo et la méthode de Runge Kutta quasi-Monte Carlo avec 10 pas de temps égaux pour chaque méthode où l'on utilise le même ensemble de Hammersley $X = \{(\phi_2(j), \phi_3(j), \frac{j}{1000}) : 0 \leq j < 1000\}$.

Les résultats sont présentés dans le tableau ci-dessous. On remarque que l'erreur moyenne de la méthode RKQMC est plus petite que celles de la méthode RKMC et RK.

t	RK	RKMC	RKQMC
0.1	6.7364E-02	-1.3191E-03	1.6800E-03
0.2	1.1604E-01	-1.3443E-03	-7.3599E-04
0.3	4.4488E-02	-1.7922E-03	-1.4834E-03
0.4	4.0551E-02	-1.6239E-03	1.2046E-03
0.5	1.3195E-01	-3.5951E-03	8.6737E-04
0.6	1.2010E-01	-4.8042E-03	-1.7180E-03
0.7	5.5717E-02	-3.122E-03	-3.9732E-04
0.8	1.1761E-01	-2.0631E-03	1.6275E-03
0.9	1.8491E-01	-2.5099E-03	-5.3732E-04
1.0	1.2666E-01	-2.1985E-03	-1.7630E-03
Moyenne	1.0054E-01	2.4372E-03	1.2014E-03

Table 3.1: Erreurs des méthodes de RK, RKMC et RKQMC

Conclusions et perspectives

Nous avons présenté de nouvelles méthodes d'ordre 3 qui sont proches des méthodes quasi-Monte Carlo pour la résolution d'un système différentielle où la fonction f subit des variations trop rapides en temps pour être suivie en détail. Après analyse de l'erreur, nous avons constaté que les suites à faible discrépance permettent d'avoir de meilleurs résultats que les nombres pseudo-aléatoires.

Les formules du schéma présenté sont beaucoup plus compliquées que celles de la méthode Runge Kutta classique et la compléxité augmente avec l'ordre. Nos persepectives futurs sont basées sur le développement de méthodes du quatrième ordre et l'application des méthodes Monte Carlo stratifiées pour la résolution des systèmes différentielles.

Bibliographie

[1] R. Béjian, H. Faure, Discrépance de la suite de Van der Corput. Séminaire Delange-Pisot-Poitou. Théorie des nombres, tom 19, N.1, exp. N.13, pp 1-14, (1977-1979).

[2] R. Béjian, Minoration de la discrépance d'une suite quelconque sur T. Acta Arithmetica, 41, pp 185-202, (1982).

[3] E. Braaten and G. Weller, An Improved Low-Discrepancy Sequence for Multidimensional Quasi-Monte Carlo Integration. J. Comput, Phys., 33: 249-258, (1979).

[4] J.C. Butcher, The Numerical Analysis of Ordinary Differential Equations, Wiley, UK(1987).

[5] A. Chouraqui, C. Lécot and B. Djebbar, Quasi-Monte Carlo simulation of differential equations, Monte Carlo Methods and Applications, De Gruyter, Berlin, 23, pp 265-275, (2017).

[6] R. Cools and D. Nuyens, Monte Carlo and Quasi-Monte Carlo methods 2014, Springer Proc. Math. Stat. 163, Springer, Cham, (2016).

[7] I. Coulibaly and C.Lécot, A quasi-randomized Runge-Kutta method, Math.Comput., 68, pp 651-659, (1999) .

[8] R. Cranley and T.N.L. Patteron, Randomization of number theoretic methods for multiple integration, SIAM Journal of Numerical Analysis, 13(6), pp 904-914, (1976).

[9] P. Davis and P. Rabinowitz, Methods of Numerical Integration, 2*nd* ed. Academic Press, San Diego, (1984).

[10] J. Dick and F. Pillichshammer, Digital Nets and Sequences, Discrepancy Theory and Quasi-Monte Carlo integration, Cambridge University Press, Cambridge, (2010).

[11] J. Dick, F. Y. Kuo, G. W. Peters and I. H. Sloan, Monte Carlo and Quasi-Monte Carlo methods 2012, Springer Proc. Math. Stat. 65, Springer, Heidelberg, (2013).

[12] M. Evans and T. Swartz, Approximating Integrals via Monte Carlo and Deterministic Methods. Oxford University Press, Oxford, (2000).

[13] H. Faure, Discrépance de suites associées à un système de numération (en dimension un). Bulletin de la Société Mathématique de France, 109, pp. 143-182, (1981).

[14] H. Faure, Discrépance de suites associées à un système de numération (en dimension s). Acta Arithmetica, 41, pp.337-351, (1982).

[15] H. Faure, Using permutations to reduce discrepancy, Journal of Computational and Applied Mathematics 31, pp. 97-103, (1990).

[16] H. Faure and S. Tezuka, Another random scrambling of digital $(t, s)-$sequences, in K-T. Fang, F. J. Hickernell and H. Niederreiter (rds.), Monte Carlo and Quasi-Monte Carlo methods 2000, Springer, Berlin. pp. 242-256 (2002).

[17] G. S. Fishman, Monte Carlo: Concepts, Algorithms and Applications. Springer-Verlag, New York, (1996).

[18] E. Hairer and S. P. Norsett, G. Wanner, Solving Ordinary Differential Equations, Springer Verlag, Berlin(2008).

[19] J. H. Halton, On the efficiency of certain of certain quasi-random sequences of points in evaluating multi-dimensional integrals, Numerische Mathematik, 2, pp 196, (1960)

[20] J.M. Hammersley, Monte Carlo methods for solving multivariable problems, Annals of the New York Academy of Sciences, 86, pp. 844-874 (1960)

[21] J.M. Hammersley and D.C. Handscomb, Monte Carlo Methods. Methuen, London, (1964)

[22] E. Hlawka, Funktionen von beschrankter Variation in der Theorie der Gleichverteilung, Ann .Mat. Pura Appl, 54, pp. 325-333, (1961).

[23] E. W. Hobson, The Theory of Functions of a real Variable and the Theory of Fourier's Series,.3rd. ed. Cambridge University Press, Cambridge, (1950).

[24] H. H. Hong and F. J. Hickernell, Algorithm 823: implementing scrambled digital sequences, ACM Transactions on Mathematical Software, 29, pp. 95-109, (2003)

[25] D. E. Knuth. The Art of Computer Programming, Vol. 2: Seminumerical Algorithms, 2nd ed. Addison-Wesley, Reading, (1981).

[26] J. F. Koksma, Een algemeene stelling uit de theorie der gelijkmage verdeeling modulo 1, Mathematica B(Zutphen), 11:7-11, (1942/43).

[27] L. Kuipers and H. Niederreiter, Uniform Distribution of Sequences, Wiley, New York (1974).

[28] C. Lécot, Quasi-randomized numerical methods for systems with coefficients of bounded variation, Elsevier Science. 55 (2001), 113-121.

[29] C. Lécot, Error bounds for quasi-Monte Carlo integration with nets. Mathematics of Computation, 65, pp. 179-187, (1996).

[30] P. L'Ecuyer, Good parameters and implementations for combined multiple recursive random number generators, Operations Research, 47, pp. 159-164, (1999).

[31] D. H. Lehmer, Mathematical methods in large-scale computing units. In: Proceedings of the Second Symposium on Large-Scale Digital Calculating Machinery (Cambridge, Mass., 1949). Harvard University Press, Cambridge, Mass., pp. 141-146, (1951).

[32] R. Lidl and H. Niederreiter, Introduction to Finite Fields and their Applications, rev. ed. Cambridge University Press, Cambridge, (1994).

[33] R. Lidl and H. Niederreiter, Finite Fields, rev. ed. CambridgeUniversity Press, Cambridge, (1997).

[34] N. Madras, Lectures on Monte Carlo Methods, AMS, Providence, (2002).

[35] J. Matousẽk, On the L_2-discrepancy for anchored boxes, Journal of Complexity, 14. pp. 527-556, (1998).

[36] J. Matousẽk, Geometric Discrepancy: An illustrated Guide, Springer, Berlin, (1999).

[37] M. Matsumoto and T.Nishimura, Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator, ADM Transaction on Modeling and Computer Simulation, Vol-8, NÂ°1, January, pp. 3-30, (1998).

[38] N. Metropolis and S.ULAM, The Monte Carlo method. J. Amer. Statist. Assoc., 44, 335-341, (1949).

[39] W. J. Morokoff and R. E. Caflich, Quasi-random sequences and their discrepancies, SIAM J.Sci.Comput., 6:1251-1279, (1994).

[40] H. Niederreiter, Discrepancy and convex programming. Annali di Matematica Pura ed Applicata, 93, pp. 89-97, (1972).

[41] H. Niederreiter, Methods for estimating discrepancy. In : S. K. Zaremba (Ed.), Application of number Theory to Numerical Analysis. Academic Press, New york, pp. 203-236, (1972).

[42] H. Niederreiter and J.M. Willss, Diskrepanz und Distanz von Maben besuglich konvexer und Jordanscher Mengen. Mathematische Zeitschrift, 144, pp. 125-134, (1975). Berichtigung, ibid. 148, p. 99 (1976).

[43] H. Niederreiter, Quasi-Monte Carlo methods and pseudo-random numbers, Bull. Amer. Math. Soc., 84:957-1041,(1978).

[44] H. Niederreiter, Point sets and sequences with small discrepancy. Monatsh. Math, 104, pp. 273-337, (1987).

[45] H. Niederreiter, Low-discrepancy and low-dispersion sequences. Journal of Number Theory, 30, pp. 51-70, (1988).

[46] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods, Society for industrial and applied mathematics, Philadelphia (1992).

[47] A. Owen, Randomly permuted $(t, m, s)-$ nets and $(t, s)-$ sequences, in H.Niederreiter and P.J.-S.Shiue (eds.), Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing, Lecture Notes in Statistics, 106, Springer, New York, pp. 299-317, (1995).

[48] A. Owen, Monte Carlo variance of scrambled net quadrature, SIAM Journal on Numerical Analysis, 34, 1884-1910, (1997).

[49] A. Owen, Variance with alternative scrambling of digital nets, ACM Transactions on modeling and Computer Simulation, 13, pp. 363-378, (2003).

[50] L. Plaskota and H. Wozniakowski, Monte Carlo and Quasi-Monte Carlo Methods 2010, Springer Proc. Math. Stat. 23, Springer, Heidelberg, (2012).

[51] P. D. Proinov, Discrepancy and integration of continuous functions, J. Approx. Theory, 52:121-131, (1988).

[52] K. F. Roth, On irregularities of distribution, Mathematika, 2, pp. 73-79, (1954).

[53] W. M. Schmidt, Lectures On Irregularities Of Distribution, K. P. Puthran at Tata Press Limited, India (1977).

[54] G. Stengle, Numerical methods for systems with measurable coefficients, Applied Mathematics 3, 25-29, (1990).

[55] G. Stengle, Error analysis of a randomized numerical method, Numer. Math. 70, 119-128, (1995).

[56] S. Tezuka, Uniform Random Numbers: Theory and Practice, Kluwer, Boston (1995).

[57] S. Tezuka and H. Faure, I-binomial scrambling of digital nets and sequences, Journal of Complexity, 19, pp. 744-757 (2003).

[58] B. Tuffin, A new permutation choice in Halton sequences, Monte Carlo and Quasi-Monte Carlo methods, pp. 427-435, (1996).

[59] H.Ventsel, Théorie des probabilités, Editions mir. Moscou. (1973).

[60] J. Von Neumann, Memorandum on the program of the high speed computer. Technical report, Institute of Advanced Study, Princeton, (1945).

[61] S. K. Zaremba, Funktionen von beschrankter Variation in der Theorie der Gleichverteilung. Annali di Matematica Pura ed Applicata, 54, pp. 325-333, (1961).

[62] S. K. Zaremba, Some applications of multidimensional integration by parts, Ann. Polon. Math. 21 ,85-96, (1968).