Quasi-Monte Carlo on Product Spaces

Kinjal Basu

Department of Statistics Stanford University



Joint work with Prof. Art Owen

Będlewo, Poland 29th April, 2015



Overview

- Introduction
- Scrambled Geometric Net
 - Function ϕ
 - Nested Uniform Scrambling
- Results
 - Results in L^2 not requiring smoothness
 - Scrambled Net Variance for smooth functions
- Proof Idea
 - Using ANOVA and Multiresolution
 - Variance and gain coefficients
 - Technical Challenges
- Summary



Introduction

• Problem : Quasi-Monte Carlo integration over product spaces of the form \mathcal{X}^s where $\mathcal{X} \subseteq \mathbb{R}^d$.

$$\mu = \frac{1}{\mathsf{vol}(\mathcal{X})^s} \int_{\mathcal{X}^s} f(\mathbf{x}) \, \mathrm{d}\mathbf{x}.$$

Usual Approach for s=1

• Measure preserving transformation $\tau:[0,1]^d\to\mathcal{X}$.



Usual Approach for s = 1

- Measure preserving transformation $\tau:[0,1]^d\to\mathcal{X}$.
- $\boldsymbol{u} \sim \mathsf{U}[0,1)^d$ then $\boldsymbol{x} = \tau(\boldsymbol{u}) \sim \mathsf{U}(\mathcal{X})$.



Usual Approach for s=1

- Measure preserving transformation $\tau:[0,1]^d \to \mathcal{X}$.
- $\boldsymbol{u} \sim \mathsf{U}[0,1)^d$ then $\boldsymbol{x} = \tau(\boldsymbol{u}) \sim \mathsf{U}(\mathcal{X})$.
- Now use

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} f(\mathbf{x}_i) = \frac{1}{n} \sum_{i=1}^{n} f(\tau(\mathbf{u}_i))$$



Usual Approach for s = 1

- Measure preserving transformation $\tau:[0,1]^d \to \mathcal{X}$.
- $\boldsymbol{u} \sim \mathsf{U}[0,1)^d$ then $\boldsymbol{x} = \tau(\boldsymbol{u}) \sim \mathsf{U}(\mathcal{X})$.
- Now use

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} f(\mathbf{x}_i) = \frac{1}{n} \sum_{i=1}^{n} f(\tau(\mathbf{u}_i))$$

• This is fine for MC. For QMC we often find that $f \circ \tau \notin BVHK$

[Pillards and Cools (2005), Fang and Wang (1994)]



General Space \mathcal{X}^s

• Our estimates are equal weight rules

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} f(\boldsymbol{x}_i), \quad \text{where} \quad \boldsymbol{x}_i = \phi(\boldsymbol{u}_i)$$

for random points $u_i \in [0,1]^s$.



General Space \mathcal{X}^s

• Our estimates are equal weight rules

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} f(\boldsymbol{x}_i), \text{ where } \boldsymbol{x}_i = \phi(\boldsymbol{u}_i)$$

for random points $u_i \in [0,1]^s$.

• Most interesting case : d = 2, such as, triangles, spherical triangles and discs.



Overview of Results

• For any
$$f \in L^2(\mathcal{X}^s)$$
, $Var(\hat{\mu}) = o(1/n)$.



Overview of Results

- For any $f \in L^2(\mathcal{X}^s)$, $Var(\hat{\mu}) = o(1/n)$.
- For each finite n, $Var(\hat{\mu})$ is bounded by a finite multiple of the Monte Carlo variance, uniformly over all $f \in L^2(\mathcal{X}^s)$.



Overview of Results

- For any $f \in L^2(\mathcal{X}^s)$, $Var(\hat{\mu}) = o(1/n)$.
- For each finite n, $Var(\hat{\mu})$ is bounded by a finite multiple of the Monte Carlo variance, uniformly over all $f \in L^2(\mathcal{X}^s)$.
- Under smoothness conditions on f and a sphericity constraint on the partitioning of $\mathcal X$ we show

$$\operatorname{Var}(\hat{\mu}) = O\left(\frac{(\log n)^{s-1}}{n^{1+2/d}}\right).$$



Overview

- Introduction
- Scrambled Geometric Net
 - Function ϕ
 - Nested Uniform Scrambling
- Results
 - Results in L^2 not requiring smoothness
 - Scrambled Net Variance for smooth functions
- Proof Idea
 - Using ANOVA and Multiresolution
 - Variance and gain coefficients
 - Technical Challenges
- Summary



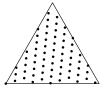


Figure: Kronecker Triangular Lattice using $3\pi/8$

• Kronecker Points. Uses badly approximable numbers.



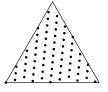


Figure: Kronecker Triangular Lattice using $3\pi/8$

- Kronecker Points. Uses badly approximable numbers.
- Convergence rate $O(\log n/n)$



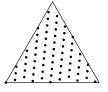
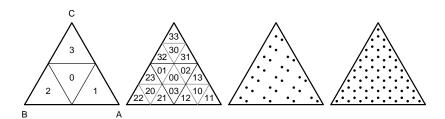


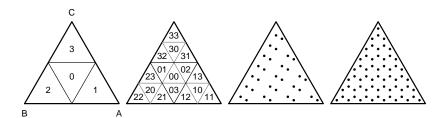
Figure: Kronecker Triangular Lattice using $3\pi/8$

- Kronecker Points. Uses badly approximable numbers.
- Convergence rate $O(\log n/n)$
- Fails for d > 2 and also for d = 2, s > 1 due to Littlewood conjecture.



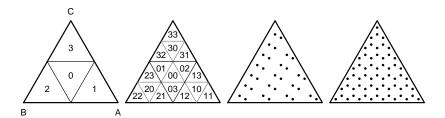






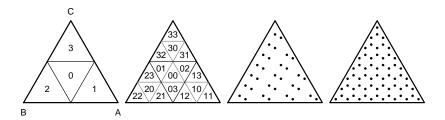
• Triangular van der Corput





- Triangular van der Corput
- $n \ge 0$ in a base 4 representation $n = \sum_{k \ge 1} d_k 4^{k-1}$ where $d_k \in \{0, 1, 2, 3\}$
- $\bullet \ x_n = 0.d_1d_2\dots$





- Triangular van der Corput
- $n \ge 0$ in a base 4 representation $n = \sum_{k \ge 1} d_k 4^{k-1}$ where $d_k \in \{0, 1, 2, 3\}$
- $\bullet \ x_n = 0.d_1d_2\ldots$
- Discrepancy: $O(1/\sqrt{n})$. RMSE under randomization: O(1/n).



Splits on the Triangle

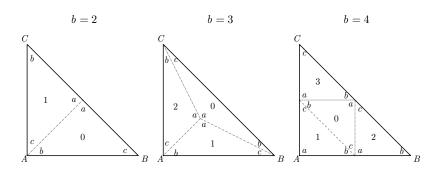


Figure: Splits of a triangle \mathcal{X} for bases b=2, 3 and 4. The subtriangles \mathcal{X}_i are labeled by the digit $j \in \mathbb{Z}_b$.



Recursive Splits on the Triangle

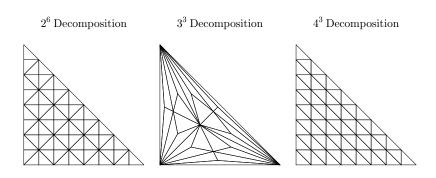


Figure: The base b splits from previous figure carried out to k = 6 or 3 or 4 levels.



Sphericity Constraint

Definition 1

Let \mathbb{X} be a recursive split of $\mathcal{X} \in \mathbb{R}^d$ in base b. Then \mathbb{X} satisfies the *sphericity condition* if there exists $C < \infty$ such that $\operatorname{diam}(\mathcal{X}_{a_1,\ldots,a_k}) \leq Cb^{-k/d}$ holds for all cells $\mathcal{X}_{a_1,\ldots,a_k}$ in \mathbb{X} .



Sphericity Constraint

Definition 1

Let \mathbb{X} be a recursive split of $\mathcal{X} \in \mathbb{R}^d$ in base b. Then \mathbb{X} satisfies the *sphericity condition* if there exists $C < \infty$ such that $\operatorname{diam}(\mathcal{X}_{a_1,\ldots,a_k}) \leq Cb^{-k/d}$ holds for all cells $\mathcal{X}_{a_1,\ldots,a_k}$ in \mathbb{X} .

Definition 2

A recursive split $\mathbb X$ in base b is convergent if for every infinite sequence $a_1,a_2,a_3,\dots\in\mathbb Z_b$, the cells $\mathcal X_{a_1,a_2,\dots a_K}$ converges a point as $K\to\infty$. That point is denoted $\lim_{K\to\infty}\mathcal X_{a_1,a_2,\dots,a_K}$.



Sphericity Constraint

Definition 1

Let \mathbb{X} be a recursive split of $\mathcal{X} \in \mathbb{R}^d$ in base b. Then \mathbb{X} satisfies the *sphericity condition* if there exists $C < \infty$ such that $\operatorname{diam}(\mathcal{X}_{a_1,\ldots,a_k}) \leq Cb^{-k/d}$ holds for all cells $\mathcal{X}_{a_1,\ldots,a_k}$ in \mathbb{X} .

Definition 2

A recursive split $\mathbb X$ in base b is convergent if for every infinite sequence $a_1,a_2,a_3,\dots\in\mathbb Z_b$, the cells $\mathcal X_{a_1,a_2,\dots a_K}$ converges a point as $K\to\infty$. That point is denoted $\lim_{K\to\infty}\mathcal X_{a_1,a_2,\dots,a_K}$.

 $\phi(x) = \lim_{K \to \infty} \mathcal{X}_{x_1, x_2, \dots, x_K}$ where x has the base b representation $0.x_1x_2...$



Splitting on the Disc

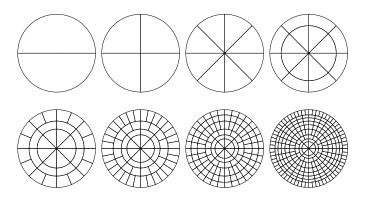


Figure: A recursive binary equal area splitting of the unit disk, keeping the aspect ratio close to unity.



• For
$$a \in [0,1)$$
 let $a = (0.a_1a_2...)_b$.



- For $a \in [0,1)$ let $a = (0.a_1a_2...)_b$.
- Random permutations to the digits a_k yielding $x_k \in \mathbb{Z}_b$ and deliver $x = (0.x_1x_2...)_b$.



- For $a \in [0,1)$ let $a = (0.a_1a_2...)_b$.
- Random permutations to the digits a_k yielding $x_k \in \mathbb{Z}_b$ and deliver $x = (0.x_1x_2...)_b$.
- $x_{k+1} = \pi_{\bullet a_1 a_2 \dots a_k}(a_{k+1})$ where all of these permutations are independent and uniform.



- For $a \in [0,1)$ let $a = (0.a_1a_2...)_b$.
- Random permutations to the digits a_k yielding $x_k \in \mathbb{Z}_b$ and deliver $x = (0.x_1x_2...)_b$.
- $x_{k+1} = \pi_{\bullet a_1 a_2 \dots a_k}(a_{k+1})$ where all of these permutations are independent and uniform.
- For $\mathbf{a} = (a_1, \dots, a_s) \in [0, 1)^s$, we apply nested uniform scramble component-wise.



- For $a \in [0,1)$ let $a = (0.a_1a_2...)_b$.
- Random permutations to the digits a_k yielding $x_k \in \mathbb{Z}_b$ and deliver $x = (0.x_1x_2...)_b$.
- $x_{k+1} = \pi_{\bullet a_1 a_2 \dots a_k}(a_{k+1})$ where all of these permutations are independent and uniform.
- For $\mathbf{a} = (a_1, \dots, a_s) \in [0, 1)^s$, we apply nested uniform scramble component-wise.
- A nested uniform scramble of $a_1, \ldots, a_n \in [0, 1)^s$ applies the same set of permutations to the digits of all n of those points.



Scrambled Geometric nets in \mathcal{X}^s via splitting

• For $s \ge 1$, let a_i be (t, m, s)-net in base b



Scrambled Geometric nets in \mathcal{X}^s via splitting

- For $s \ge 1$, let a_i be (t, m, s)-net in base b
- Let $u_i \in [0,1]^s$ nested uniform scrambling.

Scrambled Geometric nets in \mathcal{X}^s via splitting

- For $s \ge 1$, let a_i be (t, m, s)-net in base b
- Let $u_i \in [0,1]^s$ nested uniform scrambling.
- $x_{ij} = \phi(u_{ij})$ for j = 1, ..., s. Then $\mathbf{x}_i \in \mathcal{X}^s$ and we use

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} f(\mathbf{x}_i).$$

[Niederreiter (1992), Dick and Pillichshammer (2010), Owen (1995).]



Overview

- Introduction
- Scrambled Geometric Net
 - Function ϕ
 - Nested Uniform Scrambling
- Results
 - Results in L^2 not requiring smoothness
 - Scrambled Net Variance for smooth functions
- Proof Idea
 - Using ANOVA and Multiresolution
 - Variance and gain coefficients
 - Technical Challenges
- Summary



Results

Theorem 1

Let u_1, \ldots, u_n be a nested uniform scramble of a (t, m, s)-net in base $b \ge 2$ and let $x_i = \phi(u_i)$ componentwise. Then for any $f \in L^2(\mathcal{X}^{1:s})$,

$$\mathbb{E}(\hat{\mu}) = \mu$$
$$\operatorname{Var}(\hat{\mu}) = o\left(\frac{1}{n}\right)$$

as $n \to \infty$.



Results

Theorem 2

Under the conditions of Theorem 1

$$\operatorname{Var}(\hat{\mu}) \leq b^t \left(\frac{b+1}{b-1}\right)^{s-1} \frac{\sigma^2}{n},$$

where $\sigma^2 = \text{Var}(f(\mathbf{x}))$ for $\mathbf{x} \sim \mathbf{U}(\mathcal{X}^{1:s})$. If t = 0, then $\text{Var}(\hat{\mu}) \leq e\sigma^2/n \doteq 2.718\sigma^2/n$.



Main Theorem

Theorem: B. and Owen (2015b)

Let $\boldsymbol{u}_1,\ldots,\boldsymbol{u}_n$ be the points of a randomized (t,m,s)-net in base b. Let $\boldsymbol{x}_i=\phi(\boldsymbol{u}_i)\in\mathcal{X}^{1:s}$ for $i=1,\ldots,n$ where ϕ is the componentwise application of the transformation from convergent recursive splits in base b. Then for a smooth f on $\mathcal{X}^{1:s}$,

$$\operatorname{Var}(\hat{\mu}) = O\left(\frac{(\log n)^{s-1}}{n^{1+2/d}}\right).$$

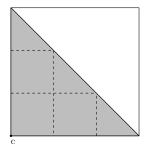


Sobol' Extensible ${\mathcal X}$

• A closed set $\mathcal{X} \subseteq \mathbb{R}^m$ with non-empty interior is said to be *Sobol' extensible* if there exists a point $c \in \mathcal{X}$ such that $c \in \mathcal{X}$ implies $\text{rect}[c, c] \subseteq \mathcal{X}$. The point $c \in \mathcal{X}$ is called the anchor.



Sobol Extensible \mathcal{X}



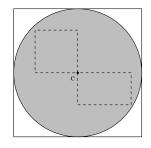
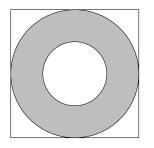


Figure: Sobol' extensible regions. At left, \mathcal{X} is the triangle with vertices $(0,0), (0,\sqrt{2}), (\sqrt{2},0)$ and the anchor is $\boldsymbol{c}=(0,0)$. At right, \mathcal{X} is a circular disk centered its anchor \boldsymbol{c} . The dashed lines depict some rectangular hulls joining selected points to the anchor.



Non Sobol Extensible \mathcal{X}



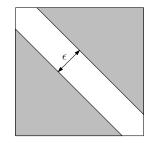


Figure: Non-Sobol' extensible regions. At left, \mathcal{X} is an annular region centered at the origin. At right, \mathcal{X} is the unit square exclusive of an ϵ -wide strip centered on the diagonal.



Smoothness Condition

• Let $\mathcal{X} \subseteq \mathbb{R}^m$ for $m \in \mathbb{N}$ be Sobol extensible. The function $f: \mathcal{X} \to \mathbb{R}$ is said to be smooth if $\partial^{1:m} f$ is continuous on \mathcal{X} .



Smoothness Condition

- Let $\mathcal{X} \subseteq \mathbb{R}^m$ for $m \in \mathbb{N}$ be Sobol extensible. The function $f: \mathcal{X} \to \mathbb{R}$ is said to be smooth if $\partial^{1:m} f$ is continuous on \mathcal{X} .
- For general \mathcal{X} , f is said to be smooth if $f \in C^m(\mathcal{X})$.



Overview

- Introduction
- Scrambled Geometric Net
 - Function ϕ
 - Nested Uniform Scrambling
- Results
 - Results in L^2 not requiring smoothness
 - Scrambled Net Variance for smooth functions
- Proof Idea
 - Using ANOVA and Multiresolution
 - Variance and gain coefficients
 - Technical Challenges
- Summary

• If \mathbb{X} is convergent and $f \in L^2(\mathcal{X})$ then,

$$f(\mathbf{x}) = \langle f, \varphi \rangle \varphi(\mathbf{x}) + \sum_{k=1}^{\infty} \sum_{t=0}^{b^k - 1} \sum_{c=0}^{b - 1} \langle f, \psi_{ktc} \rangle \psi_{ktc}(\mathbf{x}).$$



• If \mathbb{X} is convergent and $f \in L^2(\mathcal{X})$ then,

$$f(\mathbf{x}) = \langle f, \varphi \rangle \varphi(\mathbf{x}) + \sum_{k=1}^{\infty} \sum_{t=0}^{b^k - 1} \sum_{c=0}^{b - 1} \langle f, \psi_{ktc} \rangle \psi_{ktc}(\mathbf{x}).$$

• ψ_{ktc} are Haar wavelets in base b.



• Multiresolution of $L^2(\mathcal{X}^{1:s})$ is

$$f(\mathbf{x}) = \sum_{u \subseteq 1:s} \sum_{\kappa \mid u} \sum_{\tau \mid u, \kappa} \sum_{\gamma \mid u} \langle \psi_{u\kappa\tau\gamma}, f \rangle \psi_{u\kappa\tau\gamma}(\mathbf{x})$$
$$= \mu + \sum_{|u| > 0} \sum_{\kappa \mid u} \sum_{\tau \mid u, \kappa} \sum_{\gamma \mid u} \langle \psi_{u\kappa\tau\gamma}, f \rangle \psi_{u\kappa\tau\gamma}(\mathbf{x}).$$

The sum over κ is over all possible values of κ given the subset u.



• Multiresolution of $L^2(\mathcal{X}^{1:s})$ is

$$f(\mathbf{x}) = \sum_{u \subseteq 1:s} \sum_{\kappa \mid u} \sum_{\tau \mid u, \kappa} \sum_{\gamma \mid u} \langle \psi_{u\kappa\tau\gamma}, f \rangle \psi_{u\kappa\tau\gamma}(\mathbf{x})$$
$$= \mu + \sum_{|u| > 0} \sum_{\kappa \mid u} \sum_{\tau \mid u, \kappa} \sum_{\gamma \mid u} \langle \psi_{u\kappa\tau\gamma}, f \rangle \psi_{u\kappa\tau\gamma}(\mathbf{x}).$$

The sum over κ is over all possible values of κ given the subset u.

• $\psi_{u\kappa\tau\gamma}$ are tensor products of Haar wavelets in base b.



Variance expression

•

$$\operatorname{Var}(\hat{\mu}) = \mathbb{E}\left(\frac{1}{n^2} \sum_{i=1}^n \sum_{i'=1}^n \sum_{|u|>0} \sum_{\kappa|u} \sum_{\tau|u,\kappa} \sum_{\gamma|u} \sum_{|u'|>0} \sum_{\kappa'|u'} \sum_{\tau'|u',\kappa'} \sum_{\gamma'|u'} \sum_{\kappa'|u'|>0} \sum_{\kappa'|u'|} \sum_{\tau'|u',\kappa'} \sum_{\gamma'|u'|} \langle f, \psi_{u'\kappa'\tau'\gamma'} \rangle \psi_{u\kappa\tau\gamma}(\mathbf{x}_i) \psi_{u'\kappa'\tau'\gamma'}(\mathbf{x}_{i'})\right).$$



Variance expression

•

•

$$\operatorname{Var}(\hat{\mu}) = \mathbb{E}\left(\frac{1}{n^2} \sum_{i=1}^n \sum_{i'=1}^n \sum_{|u|>0} \sum_{\kappa|u} \sum_{\tau|u,\kappa} \sum_{\gamma|u} \sum_{|u'|>0} \sum_{\kappa'|u'} \sum_{\tau'|u',\kappa'} \sum_{\gamma'|u'} \left(f, \psi_{u\kappa\tau\gamma}\rangle\langle f, \psi_{u'\kappa'\tau'\gamma'}\rangle\psi_{u\kappa\tau\gamma}(\mathbf{x}_i)\psi_{u'\kappa'\tau'\gamma'}(\mathbf{x}_{i'})\right).$$

$$\operatorname{Var}(\hat{\mu}) = \sum_{|u|>0} \sum_{\kappa|u} \operatorname{Var}\left(\frac{1}{n} \sum_{i=1}^{n} \nu_{u\kappa}(\mathbf{x}_i)\right),$$

where
$$\nu_{u\kappa}(\mathbf{x}) = \sum_{\tau|u,\kappa} \sum_{\gamma|u} \langle f, \psi_{u\kappa\tau\gamma} \rangle \psi_{u\kappa\tau\gamma}(\mathbf{x})$$
 with $\nu_{\emptyset,()} = \mu$.



Simplification of Variance expression

- $\sigma_{\mu\kappa}^2 = \int_{\chi_{1:s}} \nu_{\mu\kappa}^2(x) \, \mathrm{d} x$.
- The multiresolution-based ANOVA decomposition is

$$\sigma^2 = \int_{\mathcal{X}^{1:s}} (f(\mathbf{x}) - \mu)^2 \, \mathrm{d}\mathbf{x} = \sum_{|u| > 0} \sum_{\kappa |u|} \sigma_{u\kappa}^2$$



Simplification of Variance expression

- $\bullet \ \sigma_{u\kappa}^2 = \int_{\mathcal{X}^{1:s}} \nu_{u\kappa}^2(x) \, \mathrm{d} \boldsymbol{x}.$
- The multiresolution-based ANOVA decomposition is

$$\sigma^2 = \int_{\mathcal{X}^{1:s}} (f(\mathbf{x}) - \mu)^2 d\mathbf{x} = \sum_{|u| > 0} \sum_{\kappa | u} \sigma_{u\kappa}^2$$

From the equidistribution properties of a_i

$$\operatorname{Var}(\hat{\mu}) = \frac{1}{n} \sum_{|u|>0} \sum_{\kappa|u} \Gamma_{u,\kappa} \sigma_{u\kappa}^{2}$$

$$\leq \frac{b^{t}}{n} \left(\frac{b+1}{b-1}\right)^{s} \sum_{|u|>0} \sum_{|\kappa|+|u|>m-t} \sigma_{u\kappa}^{2}.$$



• Bound $\sigma_{u\kappa}^2$.

$$\sigma_{u\kappa}^{2} = \int_{\mathcal{X}^{1:s}} \nu_{u\kappa}^{2}(\mathbf{x}) \, d\mathbf{x}$$

$$= \sum_{\tau \mid u, \kappa} \sum_{\gamma, \gamma' \mid u} \langle f, \psi_{u\kappa\tau\gamma} \rangle \langle f, \psi_{u\kappa\tau\gamma'} \rangle \int_{\mathcal{X}^{1:s}} \psi_{u\kappa\tau\gamma}(\mathbf{x}) \psi_{u\kappa\tau\gamma'}(\mathbf{x}) \, d\mathbf{x}$$

$$= \sum_{\tau \mid u, \kappa} \sum_{\gamma, \gamma' \mid u} \langle f, \psi_{u\kappa\tau\gamma} \rangle \langle f, \psi_{u\kappa\tau\gamma'} \rangle \prod_{j \in u} (1_{\gamma_{j} = \gamma'_{j}} - b^{-1})$$

• Bound $|\langle f, \psi_{u\kappa\tau\gamma} \rangle|$



• The coefficients $|\langle f, \psi_{u\kappa\tau\gamma} \rangle|$ decay quickly.



- The coefficients $|\langle f, \psi_{u\kappa\tau\gamma} \rangle|$ decay quickly.
- ullet Extend the region from \mathcal{X}^s to a bounding box
- Extend the integrand to that bounding box.



- The coefficients $|\langle f, \psi_{u\kappa\tau\gamma} \rangle|$ decay quickly.
- Extend the region from \mathcal{X}^s to a bounding box
- Extend the integrand to that bounding box.
- Sobol's low variance extension given by

$$\widetilde{f}(\mathbf{x}) = \sum_{u \in 1:m} \int_{[\mathbf{c}_u, \mathbf{x}_u]} \partial^u f(\mathbf{c}_{-u}: \mathbf{y}_u) 1_{\mathbf{c}_{-u}: \mathbf{y}_u \in \mathcal{X}} d\mathbf{y}_u$$

• Whitney extension requiring extra smoothness.



Summary

- RQMC available for \mathcal{X}^s and $\mathcal{X} \subset \mathbb{R}^d$
- Root mean squared error of $O(n^{-1/2-1/d}(\log(n))^{(s-1)/2})$.
- We can extend it to unequal d_j by taking $d = \max_{j \in 1:s} d_j$.
- QMC and composition mapping, $O(n^{-1} \log(n)^{sd-1})$
- Advantage for d=1 and 2
- Convergence guarantee as well as error estimates.
- Given additional smoothness perhaps higher order nets would work.

[Dick and Baldeaux (2009)]



Thank you!

A very big thank you to

- Prof. Leszek Plaskota and all other organizers of the conference.
- Prof. Art Owen for his discussions and support.
- National Science Foundation (NSF)
- Banach Center for organizing this conference.