Optimal order digital nets and sequences

Takashi Goda

University of Tokyo

Joint work with Kosuke Suzuki and Takehito Yoshiki

RICAM Discrepancy Workshop, November 2018

This talk is

- NOT directly about discrepancy...
- BUT about how the idea from discrepancy problem can be used to obtain some results in numerical integration.

This talk is

- NOT directly about discrepancy...
- BUT about how the idea from discrepancy problem can be used to obtain some results in numerical integration.
- The following is a list of our papers which this talk is based on.
  - TG, K. Suzuki, T. Yoshiki: Optimal order quadrature error bounds for infinite-dimensional higher-order digital sequences, Found. Comput. Math. 18 (2018) 433–458.
  - TG, K. Suzuki, T. Yoshiki: Optimal order quasi-Monte Carlo integration in weighted Sobolev spaces of arbitrary smoothness, IMA J. Numer. Anal. 37 (2017), 505–518.
  - TG, K. Suzuki, T. Yoshiki: An explicit construction of optimal order quasi-Monte Carlo rules for smooth integrands, SIAM J. Numer. Anal. 54 (2016), 2664–2683.

L_p-discrepancy

Definition

For an N-element point set $P \subset [0,1]^s$, the L_p-discrepancy is defined by

$$(L_p(P))^p := \int_{[0,1]^s} \left| \frac{1}{N} \sum_{\mathbf{x} \in P} \mathbf{1}_{\mathbf{x} \in [\mathbf{0}, \mathbf{y})} - \lambda([\mathbf{0}, \mathbf{y})) \right|^p \, \mathrm{d}\mathbf{y},$$

where λ denotes the Lebesgue measure in $\mathbb{R}^s$.

L_p-discrepancy

Definition

For an N-element point set $P \subset [0,1]^s$, the L_p-discrepancy is defined by

$$(L_p(P))^p := \int_{[0,1]^s} \left| \frac{1}{N} \sum_{\mathbf{x} \in P} \mathbf{1}_{\mathbf{x} \in [\mathbf{0},\mathbf{y})} - \lambda([\mathbf{0},\mathbf{y})) \right|^p \mathrm{d}\mathbf{y},$$

where λ denotes the Lebesgue measure in $\mathbb{R}^s$.

- In case of $p = 2$,

$$(L_2(P))^2 = \frac{1}{3^s} - \frac{2}{N} \sum_{\mathbf{x} \in P} \prod_{j=1}^{s} \frac{1 - x_j^2}{2} + \frac{1}{N^2} \sum_{\mathbf{x},\mathbf{y} \in P} \prod_{j=1}^{s} \min(1 - x_j, 1 - y_j).$$

Quasi-Monte Carlo (QMC) integration

Problem

Approximate/Estimate

$$I(f) := \int_{[0,1]^s} f(\mathbf{x})\,\mathrm{d}\mathbf{x},$$

where $s \in \mathbb{N}$ and $f\colon [0,1]^s \to \mathbb{R}$ is integrable.

Quasi-Monte Carlo (QMC) integration

Problem

Approximate/Estimate

$$I(f) := \int_{[0,1]^s} f(\mathbf{x})\,\mathrm{d}\mathbf{x},$$

where $s \in \mathbb{N}$ and $f\colon [0,1]^s \to \mathbb{R}$ is integrable.

- Choose an N-element point set $P \subset [0,1]^s$.
- Approximate $I(f)$ by

$$Q_P(f) := \frac{1}{N} \sum_{\mathbf{x} \in P} f(\mathbf{x}).$$

- For an infinite sequence of points $\mathcal{S} = \{\mathbf{x}_n \mid n \geq 0\}$, the first N elements of $\mathcal{S}$ are used as P.

Worst-case error

Definition

For a function space V with norm $\|\cdot\|_V$,

$$e^{\mathrm{wor}}(V,P) := \sup_{\|f\|_V \le 1} |Q_P(f) - I(f)|.$$

Worst-case error

Definition

For a function space V with norm $\|\cdot\|_V$,

$$e^{\mathrm{wor}}(V,P) := \sup_{\|f\|_V \le 1} |Q_P(f) - I(f)|.$$

- We want to construct a good P depending on V.
- When V is an RKHS with kernel K,

$$(e^{\mathrm{wor}}(V,P))^2 = \int_{[0,1]^s}\int_{[0,1]^s} K(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} - \frac{2}{N}\sum_{\mathbf{x}\in P}\int_{[0,1]^s} K(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{y} + \frac{1}{N^2}\sum_{\mathbf{x},\mathbf{y}\in P} K(\mathbf{x},\mathbf{y}).$$

L_2-discrepancy (again)

- Let us consider an RKHS with

$$K_{1,s}^*(\mathbf{x}, \mathbf{y}) = \prod_{j=1}^{s} \min(1 - x_j, 1 - y_j).$$

The L_2-discrepancy of P is same as the worst-case error in this RKHS.

L_2-discrepancy (again)

- Let us consider an RKHS with

$$K^*_{1,s}(\mathbf{x}, \mathbf{y}) = \prod_{j=1}^{s} \min(1 - x_j, 1 - y_j).$$

The L_2-discrepancy of P is same as the worst-case error in this RKHS.

- This RKHS is

$$V = H^*_{1,s} = \bigotimes_{j=1}^{s} H^*_1$$

where

$$H^*_1 = \left\{ f\colon [0,1] \to \mathbb{R} \mid f(1) = 0, f^{(1)} \in L_2 \right\}.$$

Optimal order QMC for L_2-discrepancy

- QMC point sets, which achieve the optimal order worst-case error in this RKHS, also achieve the optimal order L_2-discrepancy:

$$L_2(P) \asymp \frac{(\log N)^{(s-1)/2}}{N}.$$

Optimal order QMC for L_2-discrepancy

- QMC point sets, which achieve the optimal order worst-case error in this RKHS, also achieve the optimal order L_2-discrepancy:

$$L_2(P) \asymp \frac{(\log N)^{(s-1)/2}}{N}.$$

- For $s = 1$ and $s = 2$, there are many explicit constructions of such point sets.
- But for $s \geq 3$, we only know the results from

1. Chen and Skriganov (2002), Skriganov (2006)
2. Dick and Pillichshammer (2014), Dick, Hinrichs, Markhasin and Pillichshammer (2017): higher order digital nets/sequences
3. Levin (2018): Halton sequences

Sobolev spaces of our interest

- In this work we consider functions of higher smoothness:

$$V = H_{\alpha,s} = \bigotimes_{j=1}^{s} H_\alpha,$$

where $\alpha \geq 2$ and

$$H_\alpha = \left\{ f\colon [0,1] \to \mathbb{R} \mid f^{(r)}\colon \text{abs. conti. for } 0 \leq r \leq \alpha - 1, f^{(\alpha)} \in L_2 \right\}.$$

Sobolev spaces of our interest

- In this work we consider functions of higher smoothness:

$$V = H_{\alpha,s} = \bigotimes_{j=1}^{s} H_\alpha,$$

where $\alpha \geq 2$ and

$$H_\alpha = \left\{ f : [0,1] \to \mathbb{R} \mid f^{(r)} \text{: abs. conti. for } 0 \leq r \leq \alpha - 1, f^{(\alpha)} \in L_2 \right\}.$$

- $H_{\alpha,s}$ coincides with an RKHS with kernel

$$K_{\alpha,s}(\mathbf{x}, \mathbf{y}) = \prod_{j=1}^{s} \left[\sum_{r=0}^{\alpha} \frac{B_r(x_j)B_r(y_j)}{(r!)^2} + (-1)^{\alpha+1} \frac{B_{2\alpha}(|x_j - y_j|)}{(2\alpha)!} \right],$$

where B_r denotes the Bernoulli poly. of degree r (Wahba, 1990).

Higher order convergence

Known results (Dick, 2008; Baldeaux and Dick, 2009)

Higher order digital nets/sequences achieve

$$e^{\mathrm{wor}}(H_{\alpha,s}, P) \ll \frac{(\log N)^{c(\alpha,s)}}{N^{\alpha}}.$$

- $c(\alpha, s) = \alpha s$ is obtained for order α digital nets.
- The best possible exponent of $\log N$ term is $(s-1)/2$ (for linear quadrature algorithm).

Higher order convergence

Known results (Dick, 2008; Baldeaux and Dick, 2009)

Higher order digital nets/sequences achieve

$$e^{\mathrm{wor}}(H_{\alpha,s}, P) \ll \frac{(\log N)^{c(\alpha,s)}}{N^{\alpha}}.$$

- $c(\alpha, s) = \alpha s$ is obtained for order α digital nets.
- The best possible exponent of $\log N$ term is $(s-1)/2$ (for linear quadrature algorithm).

Aim of this talk

Prove that higher order digital nets/sequences achieve the *optimal* order of convergence

$$e^{\mathrm{wor}}(H_{\alpha,s}, P) \ll \frac{(\log N)^{(s-1)/2}}{N^{\alpha}}.$$

Main result

Theorem (G., Suzuki and Yoshiki, 2018)

For $\alpha \geq 2$, order $2\alpha + 1$ digital nets/sequences in prime base b achieve

$$e^{\mathrm{wor}}(H_{\alpha,s}, P) \asymp \frac{(\log N)^{(s-1)/2}}{N^\alpha}$$

for $N = b^m$ with $m \in \mathbb{N}$.

In the rest of this talk

- I focus on

1. nets, and
2. an upper bound. (Remark: A lower bound follows from the "bumping function" argument by Bakhvalov (1959).)

In the rest of this talk

- I focus on

1. nets, and
2. an upper bound. (Remark: A lower bound follows from the "bumping function" argument by Bakhvalov (1959).)

- I want to

1. introduce higher order digital nets, and
2. show a sketch of our proof for the main result by highlighting an analogy to the proof by Dick and Pillichshammer (2014), who proved the optimal order L_2-discrepancy bound for order 3 digital nets.

Digital nets

Definition (Niederreiter, 1992)

- For prime b, $\mathbb{F}_b$ denotes the b-element field.

Digital nets

Definition (Niederreiter, 1992)

- For prime b, $\mathbb{F}_b$ denotes the b-element field.
- For $s, m, n \in \mathbb{N}$, let $C_1, \ldots, C_s \in \mathbb{F}_b^{n \times m}$.

Digital nets

Definition (Niederreiter, 1992)

- For prime b, $\mathbb{F}_b$ denotes the b-element field.
- For $s, m, n \in \mathbb{N}$, let $C_1, \ldots, C_s \in \mathbb{F}_b^{n\times m}$.
- For $0 \le h < b^m$, write $h = \eta_0 + \eta_1 b + \cdots + \eta_{m-1} b^{m-1}$.

Digital nets

Definition (Niederreiter, 1992)

- For prime b, $\mathbb{F}_b$ denotes the b-element field.
- For $s, m, n \in \mathbb{N}$, let $C_1, \ldots, C_s \in \mathbb{F}_b^{n\times m}$.
- For $0 \le h < b^m$, write $h = \eta_0 + \eta_1 b + \cdots + \eta_{m-1} b^{m-1}$.
- For $1 \le j \le s$, let

$$x_{h,j} = \frac{\xi_{1,h,j}}{b} + \cdots + \frac{\xi_{n,h,j}}{b^n} \in [0,1],$$

where

$$(\xi_{1,h,j}, \ldots, \xi_{n,h,j})^\top = C_j \cdot (\eta_0, \ldots, \eta_{m-1})^\top \in \mathbb{F}_b^n.$$

Digital nets

Definition (Niederreiter, 1992)

- For prime b, $\mathbb{F}_b$ denotes the b-element field.
- For $s, m, n \in \mathbb{N}$, let $C_1, \ldots, C_s \in \mathbb{F}_b^{n\times m}$.
- For $0 \le h < b^m$, write $h = \eta_0 + \eta_1 b + \cdots + \eta_{m-1} b^{m-1}$.
- For $1 \le j \le s$, let

$$x_{h,j} = \frac{\xi_{1,h,j}}{b} + \cdots + \frac{\xi_{n,h,j}}{b^n} \in [0,1],$$

where

$$(\xi_{1,h,j}, \ldots, \xi_{n,h,j})^\top = C_j \cdot (\eta_0, \ldots, \eta_{m-1})^\top \in \mathbb{F}_b^n.$$

- Set $\mathbf{x}_h = (x_{h,1}, \ldots, x_{h,s}) \in [0,1]^s$.

Digital nets

Definition (Niederreiter, 1992)

- For prime b, $\mathbb{F}_b$ denotes the b-element field.
- For $s, m, n \in \mathbb{N}$, let $C_1, \ldots, C_s \in \mathbb{F}_b^{n\times m}$.
- For $0 \le h < b^m$, write $h = \eta_0 + \eta_1 b + \cdots + \eta_{m-1} b^{m-1}$.
- For $1 \le j \le s$, let

$$x_{h,j} = \frac{\xi_{1,h,j}}{b} + \cdots + \frac{\xi_{n,h,j}}{b^n} \in [0,1],$$

where

$$(\xi_{1,h,j}, \ldots, \xi_{n,h,j})^\top = C_j \cdot (\eta_0, \ldots, \eta_{m-1})^\top \in \mathbb{F}_b^n.$$

- Set $\mathbf{x}_h = (x_{h,1}, \ldots, x_{h,s}) \in [0,1]^s$.

We call $P = \{\mathbf{x}_h \colon 0 \le h < b^m\}$ a *digital net* over $\mathbb{F}_b$ with generating matrices $C_1, \ldots, C_s$.

Dual net

Definition

The dual net of P with $C_1, \ldots, C_s$ is defined by

$$P^{\perp} = \left\{ (k_1, \ldots, k_s) \in \mathbb{N}_0^s \colon C_1^{\top} \vec{k}_1 \oplus \cdots \oplus C_s^{\top} \vec{k}_s = \mathbf{0} \in \mathbb{F}_b^m \right\},$$

where we write

$$\vec{k} = (\kappa_0, \ldots, \kappa_{n-1}) \in \mathbb{F}_b^n$$

for $k = \kappa_0 + \kappa_1 b + \cdots$.

NRT metric and (t, m, s)-nets

- For $k = \kappa_1 b^{a_1 - 1} + \cdots + \kappa_v b^{a_v - 1}$ with $\kappa_i \in \{1, \ldots, b-1\}$ and $a_1 > a_2 > \cdots$, the NRT metric is

$$\mu_1(k) = a_1 \quad \text{and} \quad \mu_1(k_1, \ldots, k_s) = \sum_{j=1}^{s} \mu_1(k_j).$$

NRT metric and (t, m, s)-nets

- For $k = \kappa_1 b^{a_1 - 1} + \cdots + \kappa_v b^{a_v - 1}$ with $\kappa_i \in \{1, \ldots, b-1\}$ and $a_1 > a_2 > \cdots$, the NRT metric is

$$\mu_1(k) = a_1 \quad \text{and} \quad \mu_1(k_1, \ldots, k_s) = \sum_{j=1}^{s} \mu_1(k_j).$$

Definition

We call P a digital (t, m, s)-net if

$$\mu_1(P^\perp) := \min_{\mathbf{k} \in P^\perp \setminus \{\mathbf{0}\}} \mu_1(\mathbf{k}) \geq m - t + 1.$$

(Constructions: Sobol', Faure, Niederreiter, Tezuka, Niederreiter-Xing, ...)

Dick metric and higher order nets

- Let $\alpha \in \mathbb{N}$.
- For $k = \kappa_1 b^{a_1 - 1} + \cdots + \kappa_v b^{a_v - 1}$ with $\kappa_i \in \{1, \ldots, b-1\}$ and $a_1 > a_2 > \cdots$, the Dick metric is

$$\mu_\alpha(k) = a_1 + \cdots + a_{\min(v,\alpha)} \quad \text{and} \quad \mu_\alpha(k_1, \ldots, k_s) = \sum_{j=1}^{s} \mu_\alpha(k_j).$$

$\alpha = 1$: the NRT metric.

Dick metric and higher order nets

- Let $\alpha \in \mathbb{N}$.
- For $k = \kappa_1 b^{a_1 - 1} + \cdots + \kappa_v b^{a_v - 1}$ with $\kappa_i \in \{1, \ldots, b-1\}$ and $a_1 > a_2 > \cdots$, the Dick metric is

$$\mu_\alpha(k) = a_1 + \cdots + a_{\min(v,\alpha)} \quad \text{and} \quad \mu_\alpha(k_1, \ldots, k_s) = \sum_{j=1}^{s} \mu_\alpha(k_j).$$

$\alpha = 1$: the NRT metric.

Definition

We call P an order α digital (t, m, s)-net if

$$\mu_\alpha(P^\perp) := \min_{\mathbf{k} \in P^\perp \setminus \{\mathbf{0}\}} \mu_\alpha(\mathbf{k}) \geq \alpha m - t + 1.$$

Explicit construction

- Define $\mathcal{D}_\alpha : [0,1]^\alpha \to [0,1]$ by

$$\begin{cases} x_1 = & (0.\xi_{1,1}\xi_{2,1}\xi_{3,1}\ldots)_b \\ x_2 = & (0.\xi_{1,2}\xi_{2,2}\xi_{3,2}\ldots)_b \\ \vdots & \\ x_\alpha = & (0.\xi_{1,\alpha}\xi_{2,\alpha}\xi_{3,\alpha}\ldots)_b. \end{cases} \mapsto (0.\underbrace{\xi_{1,1}\xi_{1,2}\ldots\xi_{1,\alpha}}_{\alpha}\underbrace{\xi_{2,1}\xi_{2,2}\ldots\xi_{2,\alpha}}_{\alpha}\ldots)_b.$$

- For $\mathbf{x} \in [0,1]^{\alpha s}$, we write

$$\mathcal{D}_\alpha(\mathbf{x}) = (\mathcal{D}_\alpha(x_1,\ldots,x_\alpha),\ldots,\mathcal{D}_\alpha(x_{\alpha(s-1)+1},\ldots,x_{\alpha s})) \in [0,1]^s.$$

Explicit construction

- Define $\mathcal{D}_\alpha : [0,1]^\alpha \to [0,1]$ by

$$\begin{cases} x_1 = & (0.\xi_{1,1}\xi_{2,1}\xi_{3,1}\ldots)_b \\ x_2 = & (0.\xi_{1,2}\xi_{2,2}\xi_{3,2}\ldots)_b \\ \vdots & \\ x_\alpha = & (0.\xi_{1,\alpha}\xi_{2,\alpha}\xi_{3,\alpha}\ldots)_b. \end{cases} \mapsto (0.\underbrace{\xi_{1,1}\xi_{1,2}\ldots\xi_{1,\alpha}}_{\alpha}\underbrace{\xi_{2,1}\xi_{2,2}\ldots\xi_{2,\alpha}}_{\alpha}\ldots)_b.$$

- For $\mathbf{x} \in [0,1]^{\alpha s}$, we write

$$\mathcal{D}_\alpha(\mathbf{x}) = (\mathcal{D}_\alpha(x_1,\ldots,x_\alpha),\ldots,\mathcal{D}_\alpha(x_{\alpha(s-1)+1},\ldots,x_{\alpha s})) \in [0,1]^s.$$

Theorem (Dick, 2007)

Let P be a digital $(t, m, \alpha s)$-net. Then $\mathcal{D}_\alpha(P)$ is an order α digital (t_α, m, s)-net with $t_\alpha \le \alpha \min\{m, t + \lfloor s(\alpha-1)/2 \rfloor\}$.

Another look at construction

- Let P be a digital $(t, m, \alpha s)$-net with $C_1, \ldots, C_{\alpha s} \in \mathbb{F}_b^{m \times m}$. We write

$$C_1 = \begin{pmatrix} \mathbf{c}_{1,1} \\ \vdots \\ \mathbf{c}_{m,1} \end{pmatrix}, \ldots, C_\alpha = \begin{pmatrix} \mathbf{c}_{1,\alpha} \\ \vdots \\ \mathbf{c}_{m,\alpha} \end{pmatrix}, C_{\alpha+1} = \begin{pmatrix} \mathbf{c}_{1,\alpha+1} \\ \vdots \\ \mathbf{c}_{m,\alpha+1} \end{pmatrix} \ldots.$$

Another look at construction

- Let P be a digital $(t, m, \alpha s)$-net with $C_1, \dots, C_{\alpha s} \in \mathbb{F}_b^{m \times m}$. We write

$$C_1 = \begin{pmatrix} \mathbf{c}_{1,1} \\ \vdots \\ \mathbf{c}_{m,1} \end{pmatrix}, \dots, C_\alpha = \begin{pmatrix} \mathbf{c}_{1,\alpha} \\ \vdots \\ \mathbf{c}_{m,\alpha} \end{pmatrix}, C_{\alpha+1} = \begin{pmatrix} \mathbf{c}_{1,\alpha+1} \\ \vdots \\ \mathbf{c}_{m,\alpha+1} \end{pmatrix} \dots.$$

- $\mathcal{D}_\alpha(P)$ is a digital net with $D_1, \dots, D_s \in \mathbb{F}_b^{\alpha m \times m}$ where

$$D_1 = \begin{pmatrix} \mathbf{c}_{1,1} \\ \vdots \\ \mathbf{c}_{1,\alpha} \\ \vdots \\ \mathbf{c}_{m,1} \\ \vdots \\ \mathbf{c}_{m,\alpha} \end{pmatrix}, D_2 = \begin{pmatrix} \mathbf{c}_{1,\alpha+1} \\ \vdots \\ \mathbf{c}_{1,2\alpha} \\ \vdots \\ \mathbf{c}_{m,\alpha+1} \\ \vdots \\ \mathbf{c}_{m,2\alpha} \end{pmatrix}, \dots.$$

Toward the proof: Walsh functions

Definition

- For $k = \kappa_0 + \kappa_1 b + \cdots$, the k-th Walsh function wal_k is defined by

$$\mathrm{wal}_k(x) := \exp\left[\frac{2\pi i}{b}(\kappa_0\xi_1 + \kappa_1\xi_2 + \cdots)\right],$$

  where $x = \xi_1/b + \xi_2/b^2 + \cdots$.

- For vectors $\mathbf{k} \in \mathbb{N}_0^s$, we define

$$\mathrm{wal}_{\mathbf{k}}(\mathbf{x}) := \prod_{j=1}^{s} \mathrm{wal}_{k_j}(x_j).$$

ONB

- The set $\{\mathrm{wal}_{\mathbf{k}} : \mathbf{k} \in \mathbb{N}_0^s\}$ forms an ONB in $L_2([0,1]^s)$.
- Therefore, in general, we have

$$K(\mathbf{x},\mathbf{y}) = \sum_{\mathbf{k},\mathbf{l}\in\mathbb{N}_0^s} \hat{K}(\mathbf{k},\mathbf{l})\mathrm{wal}_{\mathbf{k}}(\mathbf{x})\overline{\mathrm{wal}_{\mathbf{l}}(\mathbf{y})},$$

where $\hat{K}(\mathbf{k},\mathbf{l})$ is the $(\mathbf{k},\mathbf{l})$-th Walsh coefficient

$$\hat{K}(\mathbf{k},\mathbf{l}) = \int_{[0,1]^s}\int_{[0,1]^s} K(\mathbf{x},\mathbf{y})\overline{\mathrm{wal}_{\mathbf{k}}(\mathbf{x})}\mathrm{wal}_{\mathbf{l}}(\mathbf{y})\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y}.$$

Character property

Lemma

Let P be a digital net with $C_1, \ldots, C_s \in \mathbb{F}_b^{n \times m}$. We have

$$\sum_{\mathbf{x} \in P} \mathrm{wal}_{\mathbf{k}}(\mathbf{x}) = \begin{cases} b^m = N & \text{if } \mathbf{k} \in P^{\perp}, \\ 0 & \text{otherwise.} \end{cases}$$

Worst-case error for digital nets

Lemma

For an RKHS V *with kernel* K *and a digital net* P*,*

$$(e^{\mathrm{wor}}(V,P))^2 = \sum_{\mathbf{k},\mathbf{l}\in P^{\perp}\setminus\{\mathbf{0}\}} \hat{K}(\mathbf{k},\mathbf{l}).$$

Worst-case error for digital nets

Lemma

For an RKHS V with kernel K and a digital net P,

$$(e^{\mathrm{wor}}(V,P))^2 = \sum_{\mathbf{k},\mathbf{l}\in P^\perp\setminus\{\mathbf{0}\}} \hat{K}(\mathbf{k},\mathbf{l}).$$

Proof: Recall the following explicit formula

$$(e^{\mathrm{wor}}(V,P))^2 = \int_{[0,1]^s}\int_{[0,1]^s} K(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} - \frac{2}{N}\sum_{\mathbf{x}\in P}\int_{[0,1]^s} K(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{y} + \frac{1}{N^2}\sum_{\mathbf{x},\mathbf{y}\in P} K(\mathbf{x},\mathbf{y}).$$

Worst-case error for digital nets (continued)

The first term is

$$\int_{[0,1]^s}\int_{[0,1]^s} K(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} = \hat{K}(\mathbf{0},\mathbf{0}).$$

Worst-case error for digital nets (continued)

The first term is

$$\int_{[0,1]^s}\int_{[0,1]^s} K(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} = \hat{K}(\mathbf{0},\mathbf{0}).$$

The second term is

$$\frac{2}{N}\sum_{\mathbf{x}\in P}\int_{[0,1]^s} K(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{y}$$
$$= \frac{1}{N}\sum_{\mathbf{x}\in P}\int_{[0,1]^s} K(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{y} + \frac{1}{N}\sum_{\mathbf{x}\in P}\int_{[0,1]^s} K(\mathbf{y},\mathbf{x})\,\mathrm{d}\mathbf{y}$$

Worst-case error for digital nets (continued)

The first term is

$$\int_{[0,1]^s}\int_{[0,1]^s} K(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} = \hat{K}(\mathbf{0},\mathbf{0}).$$

The second term is

$$\begin{aligned}
&\frac{2}{N}\sum_{\mathbf{x}\in P}\int_{[0,1]^s} K(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&= \frac{1}{N}\sum_{\mathbf{x}\in P}\int_{[0,1]^s} K(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{y} + \frac{1}{N}\sum_{\mathbf{x}\in P}\int_{[0,1]^s} K(\mathbf{y},\mathbf{x})\,\mathrm{d}\mathbf{y} \\
&= \frac{1}{N}\sum_{\mathbf{x}\in P}\sum_{\mathbf{k}\in\mathbb{N}_0^s} \hat{K}(\mathbf{k},\mathbf{0})\mathrm{wal}_{\mathbf{k}}(\mathbf{x}) + \frac{1}{N}\sum_{\mathbf{x}\in P}\sum_{\mathbf{l}\in\mathbb{N}_0^s} \hat{K}(\mathbf{0},\mathbf{l})\overline{\mathrm{wal}_{\mathbf{l}}(\mathbf{x})}
\end{aligned}$$

Worst-case error for digital nets (continued)

The first term is

$$\int_{[0,1]^s}\int_{[0,1]^s} K(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} = \hat{K}(\mathbf{0},\mathbf{0}).$$

The second term is

$$\begin{aligned}
&\frac{2}{N}\sum_{\mathbf{x}\in P}\int_{[0,1]^s} K(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&= \frac{1}{N}\sum_{\mathbf{x}\in P}\int_{[0,1]^s} K(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{y} + \frac{1}{N}\sum_{\mathbf{x}\in P}\int_{[0,1]^s} K(\mathbf{y},\mathbf{x})\,\mathrm{d}\mathbf{y} \\
&= \frac{1}{N}\sum_{\mathbf{x}\in P}\sum_{\mathbf{k}\in\mathbb{N}_0^s} \hat{K}(\mathbf{k},\mathbf{0})\mathrm{wal}_{\mathbf{k}}(\mathbf{x}) + \frac{1}{N}\sum_{\mathbf{x}\in P}\sum_{\mathbf{l}\in\mathbb{N}_0^s} \hat{K}(\mathbf{0},\mathbf{l})\overline{\mathrm{wal}_{\mathbf{l}}(\mathbf{x})} \\
&= \sum_{\mathbf{k}\in P^\perp} \hat{K}(\mathbf{k},\mathbf{0}) + \sum_{\mathbf{l}\in P^\perp} \hat{K}(\mathbf{0},\mathbf{l}).
\end{aligned}$$

Worst-case error for digital nets (continued)

Finally, the third term is

$$\frac{1}{N^2}\sum_{\mathbf{x},\mathbf{y}\in P} K(\mathbf{x},\mathbf{y}) = \frac{1}{N^2}\sum_{\mathbf{x},\mathbf{y}\in P}\sum_{\mathbf{k},\mathbf{l}\in\mathbb{N}_0^s}\hat{K}(\mathbf{k},\mathbf{l})\mathrm{wal}_{\mathbf{k}}(\mathbf{x})\overline{\mathrm{wal}_{\mathbf{l}}(\mathbf{y})}$$

Worst-case error for digital nets (continued)

Finally, the third term is

$$\frac{1}{N^2}\sum_{\mathbf{x},\mathbf{y}\in P} K(\mathbf{x},\mathbf{y}) = \frac{1}{N^2}\sum_{\mathbf{x},\mathbf{y}\in P}\sum_{\mathbf{k},\mathbf{l}\in\mathbb{N}_0^s} \hat{K}(\mathbf{k},\mathbf{l})\mathrm{wal}_{\mathbf{k}}(\mathbf{x})\overline{\mathrm{wal}_{\mathbf{l}}(\mathbf{y})}$$

$$= \sum_{\mathbf{k},\mathbf{l}\in\mathbb{N}_0^s} \hat{K}(\mathbf{k},\mathbf{l})\frac{1}{N}\sum_{\mathbf{x}\in P}\mathrm{wal}_{\mathbf{k}}(\mathbf{x})\overline{\frac{1}{N}\sum_{\mathbf{y}\in P}\mathrm{wal}_{\mathbf{l}}(\mathbf{y})}$$

Worst-case error for digital nets (continued)

Finally, the third term is

$$\begin{aligned}
\frac{1}{N^2}\sum_{\mathbf{x},\mathbf{y}\in P} K(\mathbf{x},\mathbf{y}) &= \frac{1}{N^2}\sum_{\mathbf{x},\mathbf{y}\in P}\sum_{\mathbf{k},\mathbf{l}\in\mathbb{N}_0^s} \hat{K}(\mathbf{k},\mathbf{l})\mathrm{wal}_{\mathbf{k}}(\mathbf{x})\overline{\mathrm{wal}_{\mathbf{l}}(\mathbf{y})} \\
&= \sum_{\mathbf{k},\mathbf{l}\in\mathbb{N}_0^s} \hat{K}(\mathbf{k},\mathbf{l})\frac{1}{N}\sum_{\mathbf{x}\in P}\mathrm{wal}_{\mathbf{k}}(\mathbf{x})\overline{\frac{1}{N}\sum_{\mathbf{y}\in P}\mathrm{wal}_{\mathbf{l}}(\mathbf{y})} \\
&= \sum_{\mathbf{k},\mathbf{l}\in P^\perp} \hat{K}(\mathbf{k},\mathbf{l}).
\end{aligned}$$

Worst-case error for digital nets (continued)

Finally, the third term is

$$\frac{1}{N^2}\sum_{\mathbf{x},\mathbf{y}\in P} K(\mathbf{x},\mathbf{y}) = \frac{1}{N^2}\sum_{\mathbf{x},\mathbf{y}\in P}\sum_{\mathbf{k},\mathbf{l}\in\mathbb{N}_0^s} \hat{K}(\mathbf{k},\mathbf{l})\mathrm{wal}_{\mathbf{k}}(\mathbf{x})\overline{\mathrm{wal}_{\mathbf{l}}(\mathbf{y})}$$

$$= \sum_{\mathbf{k},\mathbf{l}\in\mathbb{N}_0^s} \hat{K}(\mathbf{k},\mathbf{l})\frac{1}{N}\sum_{\mathbf{x}\in P}\mathrm{wal}_{\mathbf{k}}(\mathbf{x})\overline{\frac{1}{N}\sum_{\mathbf{y}\in P}\mathrm{wal}_{\mathbf{l}}(\mathbf{y})}$$

$$= \sum_{\mathbf{k},\mathbf{l}\in P^\perp} \hat{K}(\mathbf{k},\mathbf{l}).$$

Altogether we have

$$(e^{\mathrm{wor}}(V,P))^2 = \hat{K}(\mathbf{0},\mathbf{0}) - \sum_{\mathbf{k}\in P^\perp}\hat{K}(\mathbf{k},\mathbf{0}) - \sum_{\mathbf{l}\in P^\perp}\hat{K}(\mathbf{0},\mathbf{l}) + \sum_{\mathbf{k},\mathbf{l}\in P^\perp}\hat{K}(\mathbf{k},\mathbf{l})$$

$$= \sum_{\mathbf{k},\mathbf{l}\in P^\perp\setminus\{\mathbf{0}\}} \hat{K}(\mathbf{k},\mathbf{l}).$$

Learning from L_2-discrepancy for digital nets

- For a digital net P,

$$(L_2(P))^2 = \sum_{\mathbf{k},\mathbf{l} \in P^\perp \setminus \{\mathbf{0}\}} \hat{K}^*_{1,s}(\mathbf{k},\mathbf{l}),$$

  where

$$K^*_{1,s}(\mathbf{x},\mathbf{y}) = \prod_{j=1}^{s} \min(1 - x_j, 1 - y_j).$$

- Constructions of digital nets with optimal order L_2-discrepancy
  - ▷ Chen and Skriganov (2002)
  - ▷ Dick and Pillichshammer (2014): order 3 digital nets

Learning from L_2-discrepancy for digital nets

- For a digital net P,

$$(L_2(P))^2 = \sum_{\mathbf{k},\mathbf{l} \in P^\perp \setminus \{\mathbf{0}\}} \hat{K}^*_{1,s}(\mathbf{k},\mathbf{l}),$$

where

$$K^*_{1,s}(\mathbf{x},\mathbf{y}) = \prod_{j=1}^{s} \min(1-x_j, 1-y_j).$$

- Constructions of digital nets with optimal order L_2-discrepancy
  - ▷ Chen and Skriganov (2002)
  - ▷ Dick and Pillichshammer (2014): order 3 digital nets

Common idea

is to **exploit the decay and the sparsity of** $\hat{K}^*_{1,s}$**.**

Decay: $\hat{K}^*_{1,s}(\mathbf{k},\mathbf{l}) \ll b^{-\mu_1(\mathbf{k})-\mu_1(\mathbf{l})}$

Sparsity: $\hat{K}^*_{1,s}(\mathbf{k},\mathbf{l}) = 0$ for many $(\mathbf{k},\mathbf{l})$

Rough sketch: Optimal order L_2-discrepancy bounds

P: an order 3 digital (t, m, s)-net

$$(L_2(P))^2 = \sum_{\mathbf{k},\mathbf{l} \in P^{\perp} \setminus \{\mathbf{0}\}} \hat{K}^*_{1,s}(\mathbf{k}, \mathbf{l})$$

(Dick and Pillichshammer, 2014)

Rough sketch: Optimal order L_2-discrepancy bounds

P: an order 3 digital (t, m, s)-net

$$(L_2(P))^2 = \sum_{\mathbf{k},\mathbf{l} \in P^\perp \setminus \{\mathbf{0}\}} \hat{K}^*_{1,s}(\mathbf{k}, \mathbf{l}) \ll \sum_{\substack{\mathbf{k},\mathbf{l} \in P^\perp \setminus \{\mathbf{0}\} \\ \hat{K}^*_{1,s}(\mathbf{k},\mathbf{l}) \neq 0}} b^{-\mu_1(\mathbf{k}) - \mu_1(\mathbf{l})}$$

(Dick and Pillichshammer, 2014)

Rough sketch: Optimal order L_2-discrepancy bounds

P: an order 3 digital (t, m, s)-net

$$
\begin{aligned}
&(L_2(P))^2 \\
&= \sum_{\mathbf{k},\mathbf{l}\in P^\perp\setminus\{\mathbf{0}\}} \hat{K}^*_{1,s}(\mathbf{k},\mathbf{l}) \ll \sum_{\substack{\mathbf{k},\mathbf{l}\in P^\perp\setminus\{\mathbf{0}\} \\ \hat{K}^*_{1,s}(\mathbf{k},\mathbf{l})\neq 0}} b^{-\mu_1(\mathbf{k})-\mu_1(\mathbf{l})} \\
&= \sum_{z=2\mu_1(P^\perp)}^{\infty} b^{-z} \left| \{(\mathbf{k},\mathbf{l}) \in (P^\perp \setminus \{\mathbf{0}\})^2 : \mu_1(\mathbf{k}) + \mu_1(\mathbf{l}) = z, \hat{K}^*_{1,s}(\mathbf{k},\mathbf{l}) \neq 0\} \right|
\end{aligned}
$$

(Dick and Pillichshammer, 2014)

Rough sketch: Optimal order L_2-discrepancy bounds

P: an order 3 digital (t, m, s)-net

$$(L_2(P))^2$$
$$= \sum_{\mathbf{k},\mathbf{l} \in P^\perp \setminus \{\mathbf{0}\}} \hat{K}^*_{1,s}(\mathbf{k},\mathbf{l}) \ll \sum_{\substack{\mathbf{k},\mathbf{l} \in P^\perp \setminus \{\mathbf{0}\} \\ \hat{K}^*_{1,s}(\mathbf{k},\mathbf{l}) \neq 0}} b^{-\mu_1(\mathbf{k})-\mu_1(\mathbf{l})}$$
$$= \sum_{z=2\mu_1(P^\perp)}^{\infty} b^{-z} \left| \{(\mathbf{k},\mathbf{l}) \in (P^\perp \setminus \{\mathbf{0}\})^2 : \mu_1(\mathbf{k}) + \mu_1(\mathbf{l}) = z, \hat{K}^*_{1,s}(\mathbf{k},\mathbf{l}) \neq 0\} \right|$$
$$\ll \frac{m^{s-1}}{b^{2\mu_1(P^\perp)}} \ll \frac{m^{s-1}}{b^{2m}} = \frac{(\log_b N)^{s-1}}{N^2}.$$

(Dick and Pillichshammer, 2014)

If only the decay of $\hat{K}_{\alpha,s}$ is considered

$$(e^{\mathrm{wor}}(H_{\alpha,s}, P))^2 = \sum_{\mathbf{k},\mathbf{l} \in P^\perp \setminus \{\mathbf{0}\}} \hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l}),$$

where

$$K_{\alpha,s}(\mathbf{x}, \mathbf{y}) = \prod_{j=1}^{s} \left[\sum_{r=0}^{\alpha} \frac{B_r(x_j) B_r(y_j)}{(r!)^2} + (-1)^{\alpha+1} \frac{B_{2\alpha}(|x_j - y_j|)}{(2\alpha)!} \right].$$

Lemma (Baldeaux and Dick, 2009)

For $\mathbf{k}, \mathbf{l} \in \mathbb{N}_0^s$,

$$\hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l}) \ll b^{-\mu_\alpha(\mathbf{k}) - \mu_\alpha(\mathbf{l})}.$$

Rough sketch: Nearly optimal order error bounds

P: an order α digital (t, m, s)-net

$$
\begin{aligned}
&(e^{\mathrm{wor}}(H_{\alpha,s}, P))^2 \\
&= \sum_{\mathbf{k},\mathbf{l} \in P^{\perp} \setminus \{\mathbf{0}\}} \hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l})
\end{aligned}
$$

Rough sketch: Nearly optimal order error bounds

P: an order α digital (t, m, s)-net

$$(e^{\mathrm{wor}}(H_{\alpha,s}, P))^2 = \sum_{\mathbf{k},\mathbf{l} \in P^{\perp} \setminus \{\mathbf{0}\}} \hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l}) \ll \sum_{\mathbf{k},\mathbf{l} \in P^{\perp} \setminus \{\mathbf{0}\}} b^{-\mu_\alpha(\mathbf{k}) - \mu_\alpha(\mathbf{l})}$$

Rough sketch: Nearly optimal order error bounds

P: an order α digital (t, m, s)-net

$$
\begin{aligned}
&(e^{\mathrm{wor}}(H_{\alpha,s}, P))^2 \\
&= \sum_{\mathbf{k},\mathbf{l} \in P^{\perp} \setminus \{\mathbf{0}\}} \hat{K}_{\alpha,s}(\mathbf{k},\mathbf{l}) \ll \sum_{\mathbf{k},\mathbf{l} \in P^{\perp} \setminus \{\mathbf{0}\}} b^{-\mu_\alpha(\mathbf{k})-\mu_\alpha(\mathbf{l})} \\
&\quad = \sum_{z=2\mu_\alpha(P^{\perp})}^{\infty} b^{-z} \left| \{(\mathbf{k},\mathbf{l}) \in (P^{\perp} \setminus \{\mathbf{0}\})^2 : \mu_\alpha(\mathbf{k}) + \mu_\alpha(\mathbf{l}) = z\} \right|
\end{aligned}
$$

Rough sketch: Nearly optimal order error bounds

P: an order α digital (t, m, s)-net

$$
\begin{aligned}
&(e^{\mathrm{wor}}(H_{\alpha,s}, P))^2 \\
&= \sum_{\mathbf{k},\mathbf{l}\in P^\perp\setminus\{\mathbf{0}\}} \hat{K}_{\alpha,s}(\mathbf{k},\mathbf{l}) \ll \sum_{\mathbf{k},\mathbf{l}\in P^\perp\setminus\{\mathbf{0}\}} b^{-\mu_\alpha(\mathbf{k})-\mu_\alpha(\mathbf{l})} \\
&\quad = \sum_{z=2\mu_\alpha(P^\perp)}^{\infty} b^{-z} \left|\{(\mathbf{k},\mathbf{l}) \in (P^\perp \setminus \{\mathbf{0}\})^2 : \mu_\alpha(\mathbf{k}) + \mu_\alpha(\mathbf{l}) = z\}\right| \\
&\quad \ll \frac{m^{2\alpha s}}{b^{2\mu_\alpha(P^\perp)}} \ll \frac{m^{2\alpha s}}{b^{2\alpha m}} = \frac{(\log_b N)^{2\alpha s}}{N^{2\alpha}}.
\end{aligned}
$$

- The exponent of $\log N$ terms comes from the counting step with respect to μ_α-metric.

Effect of switching the metric

- Counting along μ_α-metric gives

$$\left|\{\mathbf{k} \in P^\perp \setminus \{\mathbf{0}\} : \mu_\alpha(\mathbf{k}) = z\}\right| \ll b^{z-\alpha m} m^{\alpha s} \quad \text{for } z \geq \mu_\alpha(P^\perp),$$

Effect of switching the metric

- Counting along μ_α-metric gives

$$\left|\{\mathbf{k} \in P^\perp \setminus \{\mathbf{0}\} : \mu_\alpha(\mathbf{k}) = z\}\right| \ll b^{z-\alpha m} m^{\alpha s} \quad \text{for } z \geq \mu_\alpha(P^\perp),$$

while counting along μ_1-metric gives

$$\left|\{\mathbf{k} \in P^\perp \setminus \{\mathbf{0}\} : \mu_1(\mathbf{k}) = z\}\right| \ll b^{z-m} m^{s-1} \quad \text{for } z \geq \mu_1(P^\perp).$$

Effect of switching the metric

- Counting along μ_α-metric gives

$$\left|\{\mathbf{k} \in P^\perp \setminus \{\mathbf{0}\} : \mu_\alpha(\mathbf{k}) = z\}\right| \ll b^{z-\alpha m} m^{\alpha s} \quad \text{for } z \geq \mu_\alpha(P^\perp),$$

while counting along μ_1-metric gives

$$\left|\{\mathbf{k} \in P^\perp \setminus \{\mathbf{0}\} : \mu_1(\mathbf{k}) = z\}\right| \ll b^{z-m} m^{s-1} \quad \text{for } z \geq \mu_1(P^\perp).$$

- We want to

1. switch the metric from μ_α to μ_1, and
2. exploit the sparsity of $\hat{K}_{\alpha,s}$ to deal with the double sum $\sum_{\mathbf{k},\mathbf{l} \in P^\perp \setminus \{\mathbf{0}\}}$.

Analogy to L_2-discrepancy problem

$$(e^{\mathrm{wor}}(H_{\alpha,s}, P))^2 = \sum_{\mathbf{k},\mathbf{l} \in P^{\perp} \setminus \{\mathbf{0}\}} \hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l}).$$

Exploit the decay and the sparsity of $\hat{K}_{\alpha,s}$

Decay: $\hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l}) \ll b^{-\mu_\alpha(\mathbf{k})-\mu_\alpha(\mathbf{l})}$ (Baldeaux and Dick, 2009)

Sparsity: $\hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l}) = 0$ for many $(\mathbf{k}, \mathbf{l})$ (G., Suzuki and Yoshiki, 2018)

Analogy to L_2-discrepancy problem $+\ \epsilon$

- In order to switch the metric in the counting step, we need:

Additional tricks

Propagation: P: order $2\alpha+1$ net $\Rightarrow$ P: order 1 net (Dick, 2008)

Interpolation: $\mu_\alpha(\mathbf{k}) \geq A\mu_{2\alpha+1}(\mathbf{k}) + B\mu_1(\mathbf{k})$

where $A = \dfrac{\alpha-1}{2\alpha}$ and $B = \dfrac{\alpha+1}{2\alpha}$ (G., Suzuki and Yoshiki, 2017).

- The interpolation inequality on μ_α is nothing but Jensen's inequality with respect to α.

Remark: how sparse is $\hat{K}_{\alpha,s}$?

- Let
$$k = \sum_{i=1}^{v} \kappa_i b^{c_i - 1} \quad \text{and} \quad l = \sum_{i=1}^{w} \lambda_i b^{d_i - 1},$$
such that $c_1 > c_2 > \cdots$, $d_1 > d_2 > \cdots$ and $\kappa_i, \lambda_i \in \{1, \ldots, b-1\}$.

Remark: how sparse is $\hat{K}_{\alpha,s}$?

- Let
$$k = \sum_{i=1}^{v} \kappa_i b^{c_i - 1} \quad \text{and} \quad l = \sum_{i=1}^{w} \lambda_i b^{d_i - 1},$$
such that $c_1 > c_2 > \cdots$, $d_1 > d_2 > \cdots$ and $\kappa_i, \lambda_i \in \{1, \ldots, b-1\}$.
- For $p, q \in \mathbb{N}_0$, we write
$$k^{(p)} = \sum_{i=p+1}^{v} \kappa_i b^{c_i - 1} \quad \text{and} \quad l^{(q)} = \sum_{i=q+1}^{w} \lambda_i b^{d_i - 1}$$
where the empty sum is set to 0.

Remark: how sparse is $\hat{K}_{\alpha,s}$?

- Let
$$k = \sum_{i=1}^{v} \kappa_i b^{c_i - 1} \quad \text{and} \quad l = \sum_{i=1}^{w} \lambda_i b^{d_i - 1},$$
such that $c_1 > c_2 > \cdots$, $d_1 > d_2 > \cdots$ and $\kappa_i, \lambda_i \in \{1, \ldots, b-1\}$.
- For $p, q \in \mathbb{N}_0$, we write
$$k^{(p)} = \sum_{i=p+1}^{v} \kappa_i b^{c_i - 1} \quad \text{and} \quad l^{(q)} = \sum_{i=q+1}^{w} \lambda_i b^{d_i - 1}$$
where the empty sum is set to 0.
- We say "(k, l) is of type (p, q)" if $k^{(p)} = l^{(q)}$ and $\kappa_p b^{c_p - 1} \neq \lambda_q b^{d_q - 1}$.

Remark: how sparse is $\hat{K}_{\alpha,s}$?

- Let
$$k = \sum_{i=1}^{v} \kappa_i b^{c_i - 1} \quad \text{and} \quad l = \sum_{i=1}^{w} \lambda_i b^{d_i - 1},$$
such that $c_1 > c_2 > \cdots$, $d_1 > d_2 > \cdots$ and $\kappa_i, \lambda_i \in \{1, \ldots, b-1\}$.
- For $p, q \in \mathbb{N}_0$, we write
$$k^{(p)} = \sum_{i=p+1}^{v} \kappa_i b^{c_i - 1} \quad \text{and} \quad l^{(q)} = \sum_{i=q+1}^{w} \lambda_i b^{d_i - 1}$$
where the empty sum is set to 0.
- We say "(k, l) is of type (p, q)" if $k^{(p)} = l^{(q)}$ and $\kappa_p b^{c_p - 1} \neq \lambda_q b^{d_q - 1}$.

Lemma (G., Suzuki and Yoshiki, 2018)

If there exists one index $j \in \{1, \ldots, s\}$ such that (k_j, l_j) is of type (p_j, q_j) with $p_j + q_j > 2\alpha$, then $\hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l}) = 0$.

Rough sketch: Optimal order error bounds

P: an order $2\alpha + 1$ digital (t, m, s)-net

$$(e^{\mathrm{wor}}(H_{\alpha,s}, P))^2 = \sum_{\mathbf{k},\mathbf{l} \in P^{\perp} \setminus \{\mathbf{0}\}} \hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l})$$

Rough sketch: Optimal order error bounds

P: an order $2\alpha + 1$ digital (t, m, s)-net

$$(e^{\mathrm{wor}}(H_{\alpha,s}, P))^2 = \sum_{\mathbf{k},\mathbf{l} \in P^\perp \setminus \{\mathbf{0}\}} \hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l}) \ll \sum_{\substack{\mathbf{k},\mathbf{l} \in P^\perp \setminus \{\mathbf{0}\} \\ \hat{K}_{\alpha,s}(\mathbf{k},\mathbf{l}) \neq 0}} b^{-\mu_\alpha(\mathbf{k}) - \mu_\alpha(\mathbf{l})}$$

Rough sketch: Optimal order error bounds

P: an order $2\alpha+1$ digital (t,m,s)-net

$$(e^{\mathrm{wor}}(H_{\alpha,s},P))^2 = \sum_{\mathbf{k},\mathbf{l}\in P^\perp\setminus\{\mathbf{0}\}} \hat{K}_{\alpha,s}(\mathbf{k},\mathbf{l}) \ll \sum_{\substack{\mathbf{k},\mathbf{l}\in P^\perp\setminus\{\mathbf{0}\} \\ \hat{K}_{\alpha,s}(\mathbf{k},\mathbf{l})\neq 0}} b^{-\mu_\alpha(\mathbf{k})-\mu_\alpha(\mathbf{l})}$$

$$\leq b^{-2A\mu_{2\alpha+1}(P^\perp)} \sum_{\substack{\mathbf{k},\mathbf{l}\in P^\perp\setminus\{\mathbf{0}\} \\ \hat{K}_{\alpha,s}(\mathbf{k},\mathbf{l})\neq 0}} b^{-B(\mu_1(\mathbf{k})+\mu_1(\mathbf{l}))}$$

Rough sketch: Optimal order error bounds

P: an order $2\alpha+1$ digital (t,m,s)-net

$$(e^{\mathrm{wor}}(H_{\alpha,s},P))^2 = \sum_{\mathbf{k},\mathbf{l}\in P^\perp\setminus\{\mathbf{0}\}} \hat{K}_{\alpha,s}(\mathbf{k},\mathbf{l}) \ll \sum_{\substack{\mathbf{k},\mathbf{l}\in P^\perp\setminus\{\mathbf{0}\} \\ \hat{K}_{\alpha,s}(\mathbf{k},\mathbf{l})\neq 0}} b^{-\mu_\alpha(\mathbf{k})-\mu_\alpha(\mathbf{l})}$$

$$\leq b^{-2A\mu_{2\alpha+1}(P^\perp)} \sum_{\substack{\mathbf{k},\mathbf{l}\in P^\perp\setminus\{\mathbf{0}\} \\ \hat{K}_{\alpha,s}(\mathbf{k},\mathbf{l})\neq 0}} b^{-B(\mu_1(\mathbf{k})+\mu_1(\mathbf{l}))}$$

$$= b^{-2A\mu_{2\alpha+1}(P^\perp)} \sum_{z=2\mu_1(P^\perp)}^{\infty} b^{-Bz}$$

$$\times \left| \{(\mathbf{k},\mathbf{l}) \in (P^\perp\setminus\{\mathbf{0}\})^2 : \mu_1(\mathbf{k})+\mu_1(\mathbf{l}) = z, \hat{K}_{\alpha,s}(\mathbf{k},\mathbf{l})\neq 0\} \right|$$

Rough sketch: Optimal order error bounds

P: an order $2\alpha+1$ digital (t,m,s)-net

$$(e^{\mathrm{wor}}(H_{\alpha,s},P))^2 = \sum_{\mathbf{k},\mathbf{l}\in P^\perp\setminus\{\mathbf{0}\}} \hat{K}_{\alpha,s}(\mathbf{k},\mathbf{l}) \ll \sum_{\substack{\mathbf{k},\mathbf{l}\in P^\perp\setminus\{\mathbf{0}\} \\ \hat{K}_{\alpha,s}(\mathbf{k},\mathbf{l})\neq 0}} b^{-\mu_\alpha(\mathbf{k})-\mu_\alpha(\mathbf{l})}$$

$$\leq b^{-2A\mu_{2\alpha+1}(P^\perp)} \sum_{\substack{\mathbf{k},\mathbf{l}\in P^\perp\setminus\{\mathbf{0}\} \\ \hat{K}_{\alpha,s}(\mathbf{k},\mathbf{l})\neq 0}} b^{-B(\mu_1(\mathbf{k})+\mu_1(\mathbf{l}))}$$

$$= b^{-2A\mu_{2\alpha+1}(P^\perp)} \sum_{z=2\mu_1(P^\perp)}^{\infty} b^{-Bz}$$

$$\times \left|\{(\mathbf{k},\mathbf{l})\in (P^\perp\setminus\{\mathbf{0}\})^2 : \mu_1(\mathbf{k})+\mu_1(\mathbf{l})=z, \hat{K}_{\alpha,s}(\mathbf{k},\mathbf{l})\neq 0\}\right|$$

$$\ll \frac{m^{s-1}}{b^{2A\mu_{2\alpha+1}(P^\perp)+2B\mu_1(P^\perp)}} \ll \frac{m^{s-1}}{b^{2\alpha m}} = \frac{(\log_b N)^{s-1}}{N^{2\alpha}}.$$

Some comments

- Again I want to stress similarity of the proofs. In the proof of Dick and Pillichshammer (2014), we see

$$\left|\{(\mathbf{k},\mathbf{l}) \in (P^\perp \setminus \{\mathbf{0}\})^2 : \mu_1(\mathbf{k}) + \mu_1(\mathbf{l}) = z, \hat{K}^*_{1,s}(\mathbf{k},\mathbf{l}) \neq 0\}\right|.$$

In our proof, we see

$$\left|\{(\mathbf{k},\mathbf{l}) \in (P^\perp \setminus \{\mathbf{0}\})^2 : \mu_1(\mathbf{k}) + \mu_1(\mathbf{l}) = z, \hat{K}_{\alpha,s}(\mathbf{k},\mathbf{l}) \neq 0\}\right|.$$

Some comments

- Again I want to stress similarity of the proofs. In the proof of Dick and Pillichshammer (2014), we see

$$\left|\{(\mathbf{k},\mathbf{l}) \in (P^{\perp} \setminus \{\mathbf{0}\})^2 : \mu_1(\mathbf{k}) + \mu_1(\mathbf{l}) = z, \hat{K}^*_{1,s}(\mathbf{k},\mathbf{l}) \neq 0\}\right|.$$

  In our proof, we see

$$\left|\{(\mathbf{k},\mathbf{l}) \in (P^{\perp} \setminus \{\mathbf{0}\})^2 : \mu_1(\mathbf{k}) + \mu_1(\mathbf{l}) = z, \hat{K}_{\alpha,s}(\mathbf{k},\mathbf{l}) \neq 0\}\right|.$$

- On order $2\alpha + 1$ digital *sequences*: The optimal order error bound holds only for a geometric sequence of # of points $N = b^1, b^2, \ldots$. This cannot be improved for any extensible QMC rule (Owen, 2016).

Summary

- Order $2\alpha + 1$ digital nets/sequences achieve the optimal order of convergence for numerical integration in $H_{\alpha,s} = \bigotimes_{j=1}^{s} H_\alpha$.
- Explicit construction of higher order digital nets and sequences is possible. We obtain an optimal order quadrature rule which is extensible both in the number of points and the dimension.
- The main idea of the proof comes from the L_2-discrepancy problem together with some additional tricks (interpolation and propagation).

"Deep" learning from L_2-discrepancy problem?

- After the work of Dick and Pillichshammer (2014), it has been proven that order 2 (instead of 3) digital sequences achieve the best possible order of L_p-discrepancy for all $1 < p < \infty$ by
  - Dick, Hinrichs, Markhasin and Pillichshammer (2017) Israel J. Math.
- The proof is based on *Haar functions* instead of Walsh functions.

"Deep" learning from L_2-discrepancy problem?

- After the work of Dick and Pillichshammer (2014), it has been proven that order 2 (instead of 3) digital sequences achieve the best possible order of L_p-discrepancy for all $1 < p < \infty$ by
  - Dick, Hinrichs, Markhasin and Pillichshammer (2017) Israel J. Math.
- The proof is based on *Haar functions* instead of Walsh functions.
- Can we learn further from their result? Is order $2\alpha + 1$ best possible?? Does Haar analysis help us???
- How about other function spaces (e.g., Besov and Triebel-Lizorkin)? For the case of smoothness less than 2, please refer to
  - Hinrichs, Markhasin, Oettershagen and Ullrich (2016) Numer. Math.
- We do not have any progress on these questions so far...

Thank you for your attention!