# Optimal Randomized Algorithms for Integration on Function Spaces with underlying ANOVA decomposition

Michael Gnewuch[1]

University of Kaiserslautern, Germany

October 16, 2013

Based on Joint Work with

Jan Baldeaux (UTS Sydney) & Josef Dick (UNSW Sydney)

[1]Supported by the German Science Foundation DFG under Grant GN 91/3-1 and by the Australian Research Council ARC

# ANOVA Decomposition ($\infty$-Variate Functions)

Sequence space $[0,1]^{\mathbb{N}}$, endowed with probability measure $\mathrm{d}\mathbf{x} = \otimes_{j\in\mathbb{N}} \mathrm{d}x_j$. For $f \in L^2([0,1]^{\mathbb{N}})$, $u \subset_f \mathbb{N}$:

$$f_\emptyset(\mathbf{x}) := \int_{[0,1]^{\mathbb{N}}} f(\mathbf{y})\,\mathrm{d}\mathbf{y}\,,$$

$$f_u(\mathbf{x}) := \int_{[0,1]^{[\mathbb{N}]\setminus u}} f(\mathbf{x}_u, \mathbf{y}_{\mathbb{N}\setminus u})\,\mathrm{d}\mathbf{y}_{\mathbb{N}\setminus u} - \sum_{v\subsetneq u} f_v(\mathbf{x})\,,$$

where $\mathbf{x}_u = (x_j)_{j\in u}$, $\mathbf{y}_{\mathbb{N}\setminus u} = (y_j)_{j\in\mathbb{N}\setminus u}$. Then

$$\int_0^1 f_u(\mathbf{x})\,\mathrm{d}x_j = 0 \quad \text{for } j \in u.$$

This implies

$$f = \sum_{u\subset_f \mathbb{N}} f_u \quad \text{in } L^2([0,1]^{\mathbb{N}})\,, \quad \text{and} \quad \mathrm{Var}(f) = \sum_{u\subset_f\mathbb{N}} \mathrm{Var}(f_u).$$

# Function Spaces of Integrands

Construction of spaces of integrands $f : [0,1]^{\mathbb{N}} \to \mathbb{R}$:

- Reproducing kernel Hilbert space $H = H(k)$ of univariate functions $f : [0,1] \to \mathbb{R}$ with $\int_0^1 f(x)\,\mathrm{d}x = 0$.
- Hilbert spaces $H_u$ of multivariate functions $f_u : [0,1]^u \mapsto \mathbb{R}$:

$$H_u := \widehat{\bigotimes}_{j\in u} H \quad \text{for } u \subset_f \mathbb{N}, \quad \text{where} \quad H_\emptyset = \operatorname{span}\{1\}.$$

- Hilbert space $\mathcal{H}_\gamma$ of functions of infinitely many variables: Weights $\gamma = (\gamma_u)_{u\subset_f\mathbb{N}}$ with $\sum_{u\subset_f\mathbb{N}} \gamma_u < \infty$

$$\mathcal{H}_\gamma := \left\{ \sum_{u\subset_f\mathbb{N}} f_u \,\middle|\, f_u \in H_u,\ \|f\|^2_{\mathcal{H}_\gamma} := \sum_{u\subset_f\mathbb{N}} \gamma_u^{-1}\|f_u\|^2_{H_u} < \infty \right\},$$

where $\mathcal{H}_\gamma \subset L^2([0,1]^{\mathbb{N}})$ and $f = \sum_{u\subset_f\mathbb{N}} f_u$ ANOVA decomposition.

# Weights

Product weights $\gamma$ [Sloan & Woźniakowski'98]:
Let $\gamma_1 \geq \gamma_2 \geq \gamma_3 \geq \cdots \geq 0$. Then $\gamma_u := \prod_{j \in u} \gamma_j$.

Finite-order weights $\gamma$ of order $\omega$ [Dick, Sloan, Wang & Woźniakowski'06]: $\gamma_u = 0$ for all $|u| > \omega$.

Finite-intersection weights $\gamma$ of degree $\rho$: Finite-order weights with

$$|\{u \subset_f \mathbb{N} \,|\, \gamma_u > 0,\, u \cap v \neq \emptyset\}| \leq 1 + \rho \quad \text{for all } v \subset_f \mathbb{N},\, \gamma_v > 0.$$

(Subclass of finite-intersection weights are the "finite-diameter weights" proposed by Creutzig.)

$$\text{decay} := \sup \left\{ p \in \mathbb{R} \,\middle|\, \sum_{u \subset_f \mathbb{N}} \gamma_u^{1/p} < \infty \right\}.$$

# Integration, Algorithms & Cost Model

Integration functional $I$ on $\mathcal{H}_\gamma$ : $\quad I(f) := \int_{[0,1]^{\mathbb{N}}} f(\mathbf{x})\,\mathrm{d}\mathbf{x}$

Admissable randomized algorithms:

$$Q_n(f) = \sum_{i=1}^{n} \alpha_i f(\mathbf{t}_{v_i}^{(i)}; \mathbf{a}), \quad \text{where } \mathbf{t}_{v_i}^{(i)} \in [0,1]^{v_i},\ v_i \subset_f \mathbb{N},\ a = 1/2$$

Nested Subspace Sampling [Creutzig, Dereich, Müller-Gronbach, Ritter'09]: Fix $s \geq 1$.

$$\mathrm{cost}_{\mathrm{nest}}(Q_n) := \sum_{i=1}^{n} \big(\max v_i\big)^s$$

Unrestricted Subspace Sampling [Kuo, Sloan, Wasilkowski, Woźniakowski'10]: Fix $s \geq 1$.

$$\mathrm{cost}_{\mathrm{unr}}(Q_n) := \sum_{i=1}^{n} |v_i|^s$$

# Randomized Setting

Error criterion: (worst case) randomized error

$$e^{\mathrm{ran}}(Q;\mathcal{H}_\gamma)^2 := \sup_{\|f\|_{\mathcal{H}_\gamma}\le 1} \mathbb{E}\big((I(f)-Q(f))^2\big)$$

$N$th minimal randomized error: $\mathrm{mod} \in \{\mathrm{nest}, \mathrm{unr}\}$,

$$e^{\mathrm{ran}}_{\mathrm{mod}}(N) := \inf\{e^{\mathrm{ran}}(Q;\mathcal{H}_\gamma) \,|\, Q \text{ adm. rand. alg., } \mathrm{cost}_{\mathrm{mod}}(Q) \le N\}.$$

"Convergence order" of $e^{\mathrm{ran}}_{\mathrm{mod}}(N)$:

$$\lambda^{\mathrm{ran}}_{\mathrm{mod}} := \sup\left\{t>0 \,|\, \sup_{N\in\mathbb{N}} e^{\mathrm{ran}}_{\mathrm{mod}}(N)\cdot N^t < \infty\right\}.$$

# Nested Subspace Sampling: Multilevel Algorithms

For levels $k = 1, \ldots, m$: $v_k := \{1, \ldots, 2^k\}$. $n_1 \geq n_2 \geq n_3 \geq \cdots$.

(Unbiased) RQMC Algorithms:

$$Q_{v_k}(g) := \frac{1}{n_k} \sum_{j=1}^{n_k} g(\mathbf{t}_{v_k}^{(j,k)}), \qquad \mathbf{t}_{v_k}^{(j,k)} \in [0,1]^{v_k}.$$

Projections: $\Psi_{v_k} f(\mathbf{x}) := f(\mathbf{x}_{v_k}; \mathbf{a})$ for $k \geq 1$ and $\Psi_{v_0} f(\mathbf{x}) := 0$.

RQMC-Multilevel Algorithm:

$$\mathcal{Q}_m^{\mathrm{ML}}(f) := \sum_{k=1}^{m} Q_{v_k}(\Psi_{v_k} f - \Psi_{v_{k-1}} f).$$

Cost: $$\mathrm{cost}_{\mathrm{nest}}(\mathcal{Q}_m^{\mathrm{ML}}) = \mathrm{cost}_{\mathrm{unr}}(\mathcal{Q}_m^{\mathrm{ML}}) \leq \sum_{k=1}^{m} 2\, n_k\, 2^{ks}.$$

# Nested Subspace Sampling: Multilevel Algorithms

Projections: $\Psi_{v_k} f(\mathbf{x}) = f(\mathbf{x}_{v_k}; \mathbf{a})$ for $k \geq 1$ and $\Psi_{v_0} f(\mathbf{x}) = 0$.

RQMC-ML Algo.: $\mathcal{Q}_m^{\mathrm{ML}}(f) = \sum_{k=1}^{m} Q_{v_k}(\Psi_{v_k} f - \Psi_{v_{k-1}} f)$.

Then

$$\mathbb{E}\left(Q_m^{\mathrm{ML}}(f)\right) = \sum_{k=1}^{m} I(\Psi_{v_k} f - \Psi_{v_{k-1}} f) = I(\Psi_{v_m} f),$$

and

$$\begin{aligned}&\mathbb{E}\left(\left(I(f) - Q_m^{\mathrm{ML}}(f)\right)^2\right)\\ =&\left|I(f) - I(\Psi_{v_m} f)\right|^2 + \sum_{k=1}^{m} \mathrm{Var}\left(Q_{v_k}(\Psi_{v_k} f - \Psi_{v_{k-1}} f)\right).\end{aligned}$$

# Multilevel Algorithms

Multilevel Monte Carlo algorithms were introduced in the context of integral equations and parametric integration by Heinrich (1998) and Heinrich and Sindambiwe (1999) and in the context of stochastic differential equations by Giles (2008). Multilevel quasi-Monte Carlo algorithms were tested by Giles and Waterhouse (2009).

Multilevel Monte Carlo and quasi-Monte Carlo algorithms have been studied in a number of papers, see, e.g., the web page of Mike Giles

http://people.maths.ox.ac.uk/gilesm/mlmc_community.html

for more recent information.

# Nested Subspace Sampling

Unanchored reproducing kernel $k$ of $H$: For $x, y \in [0,1]$:

$$k(x,y) = \frac{1}{3} + \frac{x^2 + y^2}{2} - \max\{x,y\}$$

$H = H(k)$ consists of functions $f \in L^2([0,1])$ with $f$ absolutely continuous, $f^{(1)} \in L^2([0,1])$, and $\int_0^1 f(x)\,\mathrm{d}x = 0$.

$k$ induces ANOVA decomposition on $\mathcal{H}_\gamma$:

$$f = \sum_{u \subset_f \mathbb{N}} f_u, \quad f_u \in H_u, \quad \text{and} \quad \int_0^1 f_u(\mathbf{x})\,\mathrm{d}x_j = 0 \quad \text{if } j \in u.$$

# Nested Subspace Sampling: Product Weights

**Theorem** [Baldeaux, G.'12]. $\gamma$ *product weights*, $\mathrm{decay} > 1$. *Then*

$$\mathrm{decay} \geq 1 + 3s: \quad \lambda^{\mathrm{ran}}_{\mathrm{nest}} = 3/2$$

$$1 + 3s \geq \mathrm{decay} > 1: \quad \lambda^{\mathrm{ran}}_{\mathrm{nest}} = \frac{\mathrm{decay} - 1}{2s}$$

(Upper error bound via multilevel algorithms based on scrambled polynomial lattice rules (scrambling: Owen'95; polynomial lattice rules: Niederreiter'92); lower error bound holds for general randomized algorithms.)

Comparison with previously known results for $s = 1$:

[Hickernell, Niu, Müller-Gronbach, Ritter'10]: Multilevel algorithms $\tilde{\mathcal{Q}}^{\mathrm{ML}}_m$ based on scrambled Niederreiter $(t, m, s)$-nets:
$\mathrm{decay} \geq 11$: $\lambda^{\mathrm{ran}}_{\mathrm{nest}} = 3/2$

[Baldeaux'11]: Multilevel algorithms $\hat{\mathcal{Q}}^{\mathrm{ML}}_m$ based on scrambled polynomial lattice rules:
$\mathrm{decay} \geq 10$: $\lambda^{\mathrm{ran}}_{\mathrm{nest}} = 3/2$

# Nested Subspace Sampling: Finite-Intersection Weights

**Theorem** [Baldeaux, G.'12]. $\gamma$ *be finite-intersection weights,* $\mathrm{decay} > 1$. *Then*

$$\mathrm{decay} \geq 1 + 3s: \quad \lambda^{\mathrm{ran}}_{\mathrm{nest}} = 3/2$$

$$1 + 3s \geq \mathrm{decay} > 1: \quad \lambda^{\mathrm{ran}}_{\mathrm{nest}} = \frac{\mathrm{decay} - 1}{2s}$$

(Upper error bound achieved by multilevel algorithms based on scrambled polynomial lattice rules; lower error bound holds for general randomized algorithms.)

# Unrestricted Subspace Sampling: CDAs (alias MDMs)

Anchored decomposition:

$$f_{\emptyset,\mathbf{a}} := f(\mathbf{a}) \quad \text{and} \quad f_{u,\mathbf{a}}(\mathbf{x}) := f(\mathbf{x}_u; \mathbf{a}) - \sum_{v \subsetneq u} f_{v,\mathbf{a}}(\mathbf{x}).$$

A changing dimension algorithm (or multivariate decomposition method) $Q^{\mathrm{CD}}$ is of the form

$$Q^{\mathrm{CD}}(f) = \sum_{u \subset_f \mathbb{N}} Q_{u,n_u}(f_{u,\mathbf{a}}),$$

$Q_{u,n_u}$ using $n_u$ samples to approximate $\int_{[0,1]^u} f_{u,\mathbf{a}}(\mathbf{x}_u)\,\mathrm{d}\mathbf{x}_u$.

$Q^{\mathrm{CD}}$ is linear if $Q_{u,n_u}$s are linear: $f_{u,\mathbf{a}}(\mathbf{x}) = \sum_{v \subseteq u} (-1)^{|u \setminus v|} f(\mathbf{x}_v; \mathbf{a})$ [Kuo, Sloan, Wasilkowski, Woźniakowski'10a]

Cost for evaluating $f_{u,\mathbf{a}}$ in unrestricted model: $O(2^{|u|}|u|^s)$.

## Changing Dimension Algorithms

Changing dimension algorithms (alias "multivariate decomposition methods") for infinite-dimensional integration were introduced in [Kuo, Sloan, Wasilkowski, Woźniakowski'10] and refined in [Plaskota & Wasilkowski'11].

These algorithms have also been adapted to infinite-dimensional approximation problems, see the papers of Wasilkowski and of Wasilkowski & Woźniakowski.

A similar idea was used for multivariate integration in [Griebel & Holtz'10] ("dimension-wise quadrature methods").

# Unrestricted Subspace Sampling

Unanchored reproducing kernel $k_\chi$ of smoothness $\chi$: $x, y \in [0,1]$

$$k_\chi(x,y) = \sum_{\tau=1}^{\chi} \frac{B_\tau(x)}{\tau!}\frac{B_\tau(y)}{\tau!} + (-1)^{\chi+1}\frac{B_{2\chi}(|x-y|)}{(2\chi)!},$$

where $B_\tau$ is Bernoulli polynomial of degree $\tau$.

$H = H(k_\chi)$ consists of functions $f \in L^2([0,1])$ with $f, f^{(1)}, \ldots, f^{(\chi-1)}$ absolutely continuous, $f^{(\chi)} \in L^2([0,1])$, and $\int_0^1 f(x)\,\mathrm{d}x = 0$.

$k_\chi$ induces ANOVA decomposition on $\mathcal{H}_\gamma$:

$$f = \sum_{u \subset_f \mathbb{N}} f_u, \quad f_u \in H_u, \quad \text{and} \quad \int_0^1 f_u(\mathbf{x})\,\mathrm{d}x_j = 0 \quad \text{if } j \in u.$$

# Unrestricted Subspace Sampling: Product Weights and Finite-Intersection Weights

**Theorem** [Dick, G.'13]. $\gamma$ *product weights or finite-intersection weights,* $\mathrm{decay} > 1$. *Then*

$$\mathrm{decay} \geq 2(\chi + 1): \quad \lambda^{\mathrm{ran}}_{\mathrm{unr}} = \chi + 1/2$$

$$2(\chi + 1) \geq \mathrm{decay} > 1: \quad \lambda^{\mathrm{ran}}_{\mathrm{unr}} = \frac{\mathrm{decay} - 1}{2}$$

(Upper error bounds achieved by changing dimension algorithms based on interlaced scrambled polynomial lattice rules [Dick'11; Goda & Dick'13]; lower error bounds holds for (rather) general randomized algorithms.)

Result of theorem still holds if cost of function evaluation in points with $k$ active variables costs $O(e^{\sigma k})$ for some $\sigma \in (0, \infty)$!

# Generalizations

- Similar results hold more generally for spaces $\mathcal{H}_\gamma$ induced by more general kernels $k : D \times D \to \mathbb{R}$, $D \subseteq \mathbb{R}$, $\rho$ probability measure on $D$ if

$$\int_D k(x, y)\rho(\mathrm{d}x) = 0 \quad \text{for all } y \in D \text{ (ANOVA Case)}.$$

- We have also results for $1 > s \geq 0$.

- The results can also be transferred to non-ANOVA settings, as the anchored or alternative unanchored settings [Dick, G., Hefter, Hinrichs, Ritter]
  - for product weights (relying on results from [Hefter, Ritter 13])
  - for more general weights (work in progress).

# Thank you for your attention!