HAPPY BIRTHDAY STEFAN

Liberating the Dimension for Weighted Integration in the Worst Case and Randomized Settings

G. W. Wasilkowski

University of Kentucky, USA

joint work with

L. Plaskota

University of Warsaw

There are computational problems with $d = \infty$ variables, e.g.,

path integrals

Such problems can be approximated by problems with finite d and many existing results could be and are used.

There are computational problems with $d = \infty$ variables, e.g.,

path integrals

Such problems can be approximated by problems with finite d and many existing results could be and are used.

HOWEVER!

Many results are **IRRELEVANT**

especially negative results

LIBERATION SO FAR

Initial attempts for Feynman–Kac–type of integrals in

W. and Woźniakowski 1996 and Plaskota, W., and Woźniakowski 2000

Recent approaches for integration:

Creutzing, Dereich, Müller-Gronbach, and Ritter 2009,

Gnewuch 10,

Hickernell, Müller-Gronbach, Niu, and Ritter 2010,

Hickernell and Wang 2001,

Kuo, Sloan, W., and Woźniakowski 2009a,

Niu and Hickernell 2010,

Niu and Hickernell, Müller-Gronbach, and Ritter 2010

HOWEVER: Special Spaces and No Sharp Bounds

TO BE PRESENTED: General Spaces and Sharp Bounds

QUASI-RKHS $\mathcal{F}$

Following Kuo, Sloan, W., and Woźniakowski 2009a

Let $D \subseteq \mathbb{R}$ and H be a reproducing kernel Hilbert space (RKHS) of functions f:

$$f : D \to \mathbb{R} \quad \text{and} \quad f(x) = \langle f, K(\cdot, x) \rangle_H \quad \text{where } K \text{ is the kernel}$$

ASSUMPTION: $K(a, a) = 0$ for an anchor $a \in D$

EXAMPLE (Wiener kernel):

$K(x, y) = \min(x, y)$ with $a = 0,$ where $D = [0, 1]$ or $D = [0, \infty)$

∞-variate domain: $\mathcal{D} = D^\infty$, $\boldsymbol{x} = [x_1, x_2, \dots] \in \mathcal{D}$

∞-variate functions:

$$f(\boldsymbol{x}) = \sum_{\mathfrak{u} \subset \mathbb{N}} f_{\mathfrak{u}}(\boldsymbol{x}),$$

where $f_{\mathfrak{u}}$ depends on variables in $\mathfrak{u}$, a finite subset of $\mathbb{N}$:

$$f_{\mathfrak{u}} \in H_{\mathfrak{u}} \quad \text{with kernel} \quad K_{\mathfrak{u}}(\boldsymbol{x}, \boldsymbol{y}) = \prod_{j \in \mathfrak{u}} K(x_j, y_j)$$

SPACE $\mathcal{F}$: completion with respect to inner-product:

$$\langle f, g \rangle_{\mathcal{F}} = \sum_{\mathfrak{u}} \frac{1}{\gamma_{\mathfrak{u}}} \cdot \langle f_{\mathfrak{u}}, g_{\mathfrak{u}} \rangle_{H_{\mathfrak{u}}} \quad \text{for} \quad f = \sum_{\mathfrak{u}} f_{\mathfrak{u}} \qquad g = \sum_{\mathfrak{u}} g_{\mathfrak{u}}$$

For simplicity, we present results only for **product weights:**

$$\gamma_{\mathfrak{u}} = \prod_{j \in \mathfrak{u}} \gamma_j$$

$\mathcal{F}$ is RKHS iff $\displaystyle\sum_{\mathfrak{u}} \gamma_{\mathfrak{u}} \cdot \sup_{x \in D} (K(x,x))^{|\mathfrak{u}|} < \infty$

iff $\displaystyle\sum_{j=1}^{\infty} \gamma_j < \infty$ and $\displaystyle\sup_{x \in D} K(x,x) < \infty$

Otherwise,

function sampling $L_{\boldsymbol{x}}(f) = f(\boldsymbol{x})$ is
DISCONTINUOUS! for some $\boldsymbol{x}$

For Wiener kernel and $D = [0, \infty)$, $\mathcal{F}$ is **NOT** RKHS.
It is only Quasi-RKHS

HOWEVER:

Function sampling $L_{\boldsymbol{x}}(f) = f(\boldsymbol{x})$ is always continuous if $\boldsymbol{x}$ has finitely many active variables

Active variables: x_j is active if $x_j \neq a$

Sampling points used by our algorithms: Given $\boldsymbol{x}$ and $\mathfrak{u}$,

$$(\boldsymbol{x};\mathfrak{u}) = [y_1, y_2, \dots] \quad \text{with} \quad y_j = \begin{cases} x_j & \text{if } j \in \mathfrak{u}, \\ a & \text{otherwise.} \end{cases}$$

$(\boldsymbol{x};\mathfrak{u})$ has $|\mathfrak{u}|$ active variables

For any finite $\mathfrak{u}$ and any $\boldsymbol{x} \in \mathcal{D}$, $L_{(\boldsymbol{x};\mathfrak{u})}$ is continuous,

$$\|L_{(\boldsymbol{x};\mathfrak{u})}\|^2 = \sum_{\mathfrak{v}\subseteq\mathfrak{u}} \gamma_{\mathfrak{v}} \cdot K_{\mathfrak{v}}(\boldsymbol{x},\boldsymbol{x}) < \infty.$$

Integration Problem

Given probability density function ρ, approximate

$$\mathcal{I}(f) = \int_{\mathcal{D}} f(\boldsymbol{x}) \cdot \rho_\infty(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$$
$$= \lim_{d\to\infty} \int_{D^d} f(x_1,\dots,x_d,\boldsymbol{a}) \prod_{j=1}^{d} \rho(x_j)\,\mathrm{d}(x_1,\dots,x_d).$$

It is continuous iff

$$\prod_{j=1}^{\infty}(1+\gamma_j\cdot C_0) < \infty \quad \text{with} \quad C_0 = \int_D \int_D K(x,y)\cdot\rho(x)\cdot\rho(y)\,\mathrm{d}x\,\mathrm{d}y,$$

Hence **STANDING ASSUMPTION:**

$$C_0 < \infty \quad \text{and} \quad \sum_{j=1}^{\infty}\gamma_j < \infty$$

Algorithms: $\mathcal{A}_n(f) = \sum_{i=1}^{n} f(\boldsymbol{x}_i; \mathbf{u}_i) \cdot a_i$

Errors:

$e(\mathcal{A}_n; \mathcal{F}) := \sup_{\|f\|_{\mathcal{F}} \le 1} |\mathcal{I}(f) - \mathcal{A}_n(f)|$ if $\mathcal{A}_n$ deterministic

$e(\mathcal{A}_n; \mathcal{F}) := \sup_{\|f\|_{\mathcal{F}} \le 1} \sqrt{\mathbb{E}|\mathcal{I}(f) - \mathcal{A}_n(f)|^2}$ if $\mathcal{A}_n$ randomized

Algorithms: $\mathcal{A}_n(f) = \sum_{k=1}^{n} f(\boldsymbol{x}_i; \mathfrak{u}_i) \cdot a_i$

Errors:

$e(\mathcal{A}_n; \mathcal{F}) := \sup_{\|f\|_{\mathcal{F}} \le 1} |\mathcal{I}(f) - \mathcal{A}_n(f)|$ if $\mathcal{A}_n$ deterministic

$e(\mathcal{A}_n; \mathcal{F}) := \sup_{\|f\|_{\mathcal{F}} \le 1} \sqrt{\mathbb{E}|\mathcal{I}(f) - \mathcal{A}_n(f)|^2}$ if $\mathcal{A}_n$ randomized

Cost of sampling $f(\boldsymbol{x}; \mathfrak{u})$: $\$(|\mathfrak{u}|)$, $\$$ is a cost function:

$$\$: [0, \infty) \to [1, \infty) \quad \text{is monotonic}$$

For instance, cost of computing $f(x_1, a, x_3, a, \dots)$ equals $\$(2)$

Cost of $\mathcal{A}_n$: $\mathrm{cost}(\mathcal{A}_n) := \sum_{i=1}^{n} \$(|\mathfrak{u}_i|)$

ε-Complexity: the minimal cost among algorithms with errors $\le \varepsilon$:

$$\mathrm{comp}(\varepsilon;\mathcal{F}) := \inf\left\{\mathrm{cost}(\mathcal{A}) \;\;:\;\; e(\mathcal{A},\mathcal{F}) \le \varepsilon\right\}$$

ε-Complexity: the minimal cost among algorithms with errors $\leq \varepsilon$:

$$\mathrm{comp}(\varepsilon;\mathcal{F}) = \inf\{\mathrm{cost}(\mathcal{A}) \ : \ e(\mathcal{A},\mathcal{F}) \leq \varepsilon\}$$

Polynomial Tractability: if there are C and p such that

$$\mathrm{comp}(\varepsilon,\mathcal{F}) \leq C \cdot \varepsilon^{-p} \quad \text{for all } \varepsilon \in (0,1)$$

The smallest such p is the exponent of tractability:

p^{wor} if only deterministic algorithms are allowed,

and p^{ran} in general

Weak Tractability: if $\lim_{\varepsilon\to 0} \varepsilon \cdot \ln(\mathrm{comp}(\varepsilon;\mathcal{F})) = 0$

i.e., complexity is **NOT** exponential in $1/\varepsilon$

One of Results: Using Smolyak's Construction

Based on Plaskota and W. 2010, and using techniques from Kuo, Sloan, W., and Woźniakowski 2009a and W. and Woźniakowski 2010b

ASSUMPTION 1: $\gamma_j \le c_1 \cdot j^{-\beta}$ for $\beta > 1$

ASSUMPTION 2: There are C_1, p such that for every $n = 0, 1, \ldots$ there is a quadrature Q_n that uses n function samples and

$$e(Q_n; H) \le \frac{C_1}{(n+1)^{1/p}}$$

On the ASSUMPTION 2

From Hinrichs 2010:

ASSUMPTION 2 holds with $p = 2$ for any K and ρ

From Plaskota, W., and Zhao 2009: Assume $\int_D \sqrt{K(x,x)} \cdot \rho(x)\,\mathrm{d}x < \infty$

Let λ_n be the eigenvalues of

$$W(f)(x) = \int_D f(y) \cdot \frac{K(y,x)}{\sqrt{K(y,y)}} \cdot \rho(y)\,\mathrm{d}y$$

If $\lambda_n = \mathcal{O}\left(n^{-q}\right)$ then $p = \dfrac{2}{q+1-\delta}$

Always $q \geq 1$ and hence $p \leq \dfrac{2}{2-\delta} \sim 1$

Let $D = [0,1]$, $\rho \equiv 1$

For Wiener kernel $K(x,y) = \min(x,y)$

$$p = \begin{cases} \frac{2}{3} & \text{for randomized algs.} \\ 1 & \text{for deterministic algs.} \end{cases}$$

For r-folded Wiener kernel

$$K(x,y) = \int_0^{\min(x,y)} \frac{(x-t)^{r-1} \cdot (y-t)^{r-1}}{[(r-1)!]^2} \, \mathrm{d}t$$

$$p = \begin{cases} \frac{2}{2 \cdot r+1} & \text{for randomized algs.} \\ \frac{1}{r} & \text{for deterministic algs.} \end{cases}$$

Given $\mathfrak{u} \neq \emptyset$ Smolyak's construction Smolyak [64] yields cubatures $Q_{\mathfrak{u},n}$ such that

$$e(Q_{\mathfrak{u},n}; H_{\mathfrak{u}}) \leq \frac{c_2 \cdot C_2^{|\mathfrak{u}|}}{(n+1)^{1/p}} \cdot \left[1 + \frac{\ln(n+1)}{|\mathfrak{u}|-1}\right]^{(|\mathfrak{u}|-1)\cdot\alpha}$$

for all $n = 0, 1, \ldots$ and some $\alpha \geq 0$ (see W. and Woźniakowski 1995)

$Q_{\mathfrak{u},n}$ are **deterministic** if Q_n are

Define

$$L(k;\gamma) := \sum_{j=1}^{\infty} \gamma_j^k \quad \text{and} \quad L_k := -1 + \prod_{j=1}^{\infty}(1+\gamma_j^k)$$

Due **ASSUMPTION 1**, $L(k;\gamma),\ L_k < \infty$ for all $k > \beta^{-1}$

Hence, for any $a < 1 - 1/\beta$

$$L_{1-a} < \infty$$

Define

$$n_{\mathfrak{u}} := \begin{cases} 0 & \text{if } \gamma_{\mathfrak{u}}^a \cdot L_{1-a} \cdot c_2 \cdot C_2^{|\mathfrak{u}|} \le \varepsilon^2 \\ \left\lceil \left[\gamma_{\mathfrak{u}}^a \cdot L_{1-a} \cdot c_2 \cdot C_2^{|\mathfrak{u}|} \cdot \varepsilon^{-2}\right]^p \right\rceil & \text{otherwise} \end{cases}$$

and

$$\mathcal{Q}_{\varepsilon}^{\mathrm{CD}}(f) := f(\boldsymbol{a}) + \sum_{\mathfrak{u};n_{\mathfrak{u}}\ge 1} Q_{n_{\mathfrak{u}},\mathfrak{u}}(f_{\mathfrak{u}})$$

Lemma 1: Let

$$d(\varepsilon) := \max\{|\mathfrak{u}| \ : \ n_{\mathfrak{u}} \geq 1\}$$

be the largest number of **active variables** used by $\mathcal{Q}_\varepsilon^{\mathrm{CD}}$. Then

$$d(\varepsilon) \leq c \cdot \frac{\ln(1/\varepsilon)}{\ln(\ln(1/\varepsilon))} = o(\ln(1/\varepsilon))$$

$\mathcal{Q}_\varepsilon^{\mathrm{CD}}$ uses samples of $f_{\mathfrak{u}}$. Due to Kuo, Sloan, W., and Woźniakowski 2009b cost of sampling $f_{\mathfrak{u}}$ is bounded by $2^{|\mathfrak{u}|} \cdot \$(|\mathfrak{u}|)$

Hence

$$\mathrm{cost}(\mathcal{Q}_\varepsilon^{\mathrm{CD}}) \leq \$(0) + 2^{d(\varepsilon)} \cdot \$(d(\varepsilon)) \cdot \sum_{\mathfrak{u};n_{\mathfrak{u}} \geq 1} n_{\mathfrak{u}}$$

THM. 1 Let **ASSUMPTIONS 1 and 2** hold.

If $\gamma_n = \mathcal{O}(n^{-\beta})$ with $\beta > 1$ then

$$e(\mathcal{Q}_\varepsilon^{\mathrm{CD}}; \mathcal{F}) = \varepsilon \cdot (1 + o(1))$$

and

$$\mathrm{cost}(\mathcal{Q}_\varepsilon^{\mathrm{CD}}) \le \frac{c_\delta \cdot \$(d(\varepsilon))}{\varepsilon^{\max(p\,,\,2/(\beta-1)+\delta)}}.$$

COROLLARY 1

Polynomial Tractability with tractability exponent

$$p^{\mathrm{sett}} \le \max\left(p\,,\,\frac{2}{\beta-1}\right) \quad \text{even for} \quad \$(d) = \mathcal{O}\left(\varepsilon^{k\cdot d}\right)$$

Weak Tractability even for

$$\$(d) = \mathcal{O}\left(e^{e^{k\cdot d}}\right)$$

OPTIMALITY

Due to a lower bound from Kuo, Sloan, W., and Woźniakowski 2009a

THM. 2 (Worst Case Setting)

If $\gamma_n = \Theta\left(n^{-\beta}\right)$ $(\beta > 1)$ and the exponent $1/p$ in **ASSUMPTION 2** is sharp, then

$$p^{\mathrm{wor}} = \max\left(p\,,\, \frac{2}{\beta - 1}\right)$$

for ALL $\$$ such that

$$\Omega(d+1) \le \$(d) \le \Theta\left(e^{k \cdot d}\right)$$

THM. 3 Let **ASSUMPTIONS 1 and 2** hold.
If $\gamma_n = \mathcal{O}(r^n)$ with $r \in (0,1)$ then (as before)

$$e(\mathcal{Q}_\varepsilon^{\mathrm{CD}}; \mathcal{F}) = \varepsilon \cdot (1 + o(1)),$$

However now

$$d(\varepsilon) = \mathcal{O}\left(\sqrt{\ln(1/\varepsilon)}\right)$$

and therefore

$$\mathrm{cost}(\mathcal{Q}_\varepsilon^{\mathrm{CD}}) \leq \frac{c \cdot \$(d(\varepsilon))}{\varepsilon^{\mathrm{exponent}}}$$

where

$$\mathrm{exponent} = p + \frac{\mathcal{O}(\ln(\ln(1/\varepsilon)))}{\sqrt{\ln(\varepsilon)}} = p + o(1)$$

COROLLARY 2

Polynomial Tractability with tractability exponent

$$p^{\mathrm{trct}} \le p \quad \text{even for} \quad \$(d) = \mathcal{O}\left(\varepsilon^{k \cdot d^r}\right) \quad \text{with} \quad r < 2.$$

Weak Tractability even for

$$\$(d) = \mathcal{O}\left(e^{e^{k \cdot d^2}}\right)$$

Clearly, if the exponent $1/p$ in **ASSUMPTION 2** is sharp, then

$$p^{\mathrm{wor}} = p$$

for ALL \$ such that $\Omega(d+1) \le \$(d) \le \Theta\left(e^{k \cdot d^r}\right)$

HAPPY BIRTHDAY STEFAN