∞-Variate Integration

G. W. Wasilkowski
Department of Computer Science
University of Kentucky

Presentation based on papers co-authored with

A. Gilbert, M. Gnewuch, M. Hefter,
A. Hinrichs, P. Kritzer, F. Y. Kuo,
F. Pillichshammer, L. Plaskota,
K. Ritter, I. H. Sloan, H. Woźniakowski

INFORMATION-BASED COMPLEXITY approach
very nicely explained in Houman Owhadi's tutorial:
many thanks Houman

Briefly: instead of working with specific functions,
IBC deals with problems on whole spaces of functions
and tries to determine
the **complexity**, i.e., the minimal cost,
and (almost) **optimal algorithms**.

This is done in various settings including:
worst case, average case and randomized settings

"Classical" Integration Problem

Given a normed space F of functions on $[0,1]^d$

approximate $I(f) = \int_{D^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$

by cubatures $Q_n(f) = \sum_{j=1}^{n} f(\boldsymbol{t}_{n,j}) \cdot a_{n,j}$

with small error $\|I - Q_n\|$ and

(if possible) small cost

Classical Methods are Extremely Bad!!!! even for Finitely Many Variables.

(Product) Trapezoidal T_n with n samples has error

$$\mathrm{error}(T_n; d = 2) \simeq \frac{1}{n} \qquad d = 2 \text{ variables}$$

$$\mathrm{error}(T_n; d = 360) \simeq \frac{1}{n^{2/360}} \qquad d = 360 \text{ variables}$$

E.g., for 360 variables, it needs

$n \sim 20^{180}$ to get only 1 digit of accuracy

$$20^{180} =$$

1532495540865888858358347027150309183618739122
1836021760000000000000000000000000000000000000
0000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000
00000

For Classical (Isotropic) Spaces

One **Cannot Do Better**

“Curse of Dimensionality”

For Classical (Isotropic) Spaces

One **Cannot Do Better**

“Curse of Dimensionality”

To Break this Curse

Different Spaces Are Needed

Weighted Spaces
treat different variables differently

Motivating Example

Compute expectation $\mathbb{E}(g(\boldsymbol{X}(t_0)))$ for

stochastic process $\boldsymbol{X}(t) = \sum_{j=1}^{\infty} x_j \, \xi_j(t)$

Equivalent to computing an integral of

$$f(x_1, x_2, \dots) = g\left(\sum_{j=1}^{\infty} x_j \, \xi_j(t_0)\right)$$

In $$g\left(\sum_{j=1}^{\infty} x_j\, \xi_j(t_0)\right)$$

“importance” of x_j

is quantized by the size of $|\xi_j(t_0)|$.

The larger $|\xi_j(t_0)|$ the more important x_j.

Although there are results for quite general spaces and problems

we present results for **Integration** over a special space $\mathcal{F}$

$\mathcal{F}$ is the γ-weighted Banach space of functions with dominating mixed derivatives of order one bounded in L_p-norm

Notation:

$\mathfrak{w}$ finite subsets of $\mathbb{N}_+$

listing the "variables in action"

Given $\boldsymbol{x} = (x_1, x_2, \dots)$, $\boldsymbol{x}_\mathfrak{w} = (x_j : j \in \mathfrak{w})$

$$[\boldsymbol{x}_\mathfrak{w}; \mathbf{0}] = (y_1, y_2, \dots) \quad \text{with} \quad y_j = \begin{cases} x_j & \text{if } j \in \mathfrak{w}, \\ 0 & \text{if } j \notin \mathfrak{w} \end{cases}$$

$$f^{(\mathfrak{w})} = \frac{\partial^{|\mathfrak{w}|}}{\partial \boldsymbol{x}_\mathfrak{w}} f = \prod_{j \in \mathfrak{w}} \frac{\partial}{\partial x_j} f$$

Domain: $D^{\mathbb{N}}$ set of sequences $(x_j)_{j\in\mathbb{N}}$ with $x_j \in D$;
for simplicity $D = [0,1]$.

$\mathcal{F}$ the Banach space of $f : D^{\mathbb{N}} \to \mathbb{R}$
endowed with the norm

$$\|f\|_{\mathcal{F}} = \left(\sum_{\mathfrak{w}} \gamma_{\mathfrak{w}}^{-p} \, \|f^{(\mathfrak{w})}([\cdot_{\mathfrak{w}}; \mathbf{0}])\|_{L_p(D^{|\mathfrak{w}|})}^p \right)^{1/p} < \infty$$

Here $p \in [1,\infty]$ and $\gamma_{\mathfrak{w}} \geq 0$ are *weights*

For simplicity

Product Weights introduced by

[Sloan and Woźniakowski 1998]:

$$\gamma_{\mathfrak{w}} = c \prod_{j \in \mathfrak{w}} \gamma_j \qquad (\gamma_j = j^{-\beta})$$

For 'motivating example' we have

$$\gamma_{\mathfrak{w}} \simeq \prod_{j \in \mathfrak{w}} |\xi_j(t_0)|^{\alpha}$$

Integration Problem

APPROXIMATE:

$$\mathcal{I}(f) \;:=\; \int_{D^{\mathbb{N}}} f(\boldsymbol{x})\,\mathrm{d}^{\mathbb{N}}\boldsymbol{x}$$
$$= \lim_{d\to\infty} \int_{D^d} f(x_1,\ldots,x_d,0,0,\ldots)\,\mathrm{d}[x_1,\ldots,x_d]$$

Integration Problem

APPROXIMATE:

$$\mathcal{I}(f) \;:=\; \int_{D^{\mathbb{N}}} f(\boldsymbol{x})\,\mathrm{d}^{\mathbb{N}}\boldsymbol{x}$$

$$=\; \lim_{d\to\infty} \int_{D^d} f(x_1,\dots,x_d,0,0,\dots)\,\mathrm{d}[x_1,\dots,x_d]$$

We have

$$\|\mathcal{I}\| \;=\; \left(\sum_{\mathfrak{w}} \gamma_{\mathfrak{w}}^{p^*}/(1+p^*)^{|\mathfrak{w}|}\right)^{1/p^*} \qquad \left(=\max_{\mathfrak{w}} \gamma_{\mathfrak{w}} \text{ for } p=1\right)$$

ASSUMPTIONS:

(A1) $$\|\mathcal{I}\| < \infty$$

(A2) We can sample f at points $[\boldsymbol{x}_{\mathfrak{w}};\mathbf{0}]$

For 'motivating example' we have:

$$f([\boldsymbol{x}_{\mathfrak{w}};\mathbf{0}]) = g\left(\sum_{j\in\mathfrak{w}} x_j\,\xi(t_0)\right)$$

How to Cope with So Many Variables?

(i) **Truncate the dimension**, i.e.,

approximate the integral of

$$f(x_1,\dots,x_k,0,0,\dots)$$

or (even better)

(ii) use

Multivariate Decomposition Method

Low Truncation Dimension

[Kritzer, Pillichshammer, W. 2016]$\subset$ [Hinrichs, Kritzer, Pillichshammer, W.]

Let

$$f_k(x_1,\dots,x_k) = f(x_1,\dots,x_k,0,0,\dots)$$

Given the **error demand** $\varepsilon > 0$,

$\mathbf{dim}^{\mathbf{trnc}}(\varepsilon)$ ε-**truncation dimension**

the smallest k such that

$$|\mathcal{I}(f)-\mathcal{I}(f_k)| \le \varepsilon\,\|f\|_{\mathcal{F}} \quad \text{for all } f\in\mathcal{F}$$

Our concept of **Truncation Dimension**

is different than the one in Statistics!!!

If

$$|\mathcal{I}(f)-\mathcal{I}(f_k)| \le \varepsilon\,\|f\|_{\mathcal{F}} \quad \text{and} \quad |\mathcal{I}(f_k)-Q_k(f_k)| \le \varepsilon\,\|f\|_{\mathcal{F}}$$

then

$$|\mathcal{I}(f)-Q_k(f_k)| \le 2\,\varepsilon\,\|f\|_{\mathcal{F}}$$

Hence

the smaller $\mathbf{dim}(\varepsilon)$ the better

Special Case: $\gamma_{\mathfrak{w}} = c \prod_{j\in\mathfrak{w}} j^{-\beta}$

$$\mathbf{dim}^{\mathbf{trnc}}(\varepsilon) \le \min\left\{\ell \ : \ \sum_{j=\ell+1} j^{-\beta\, p^*} \le \frac{p^*+1}{c^{p^*}} \ln(1/(1-\varepsilon^{p^*}))\right\}$$

$$= O\left(\varepsilon^{-1/(\beta-1+1/p)}\right)$$

for $p > 1$ and

$$\mathbf{dim}^{\mathbf{trnc}}(\varepsilon) = \left\lceil \left(\frac{c}{\varepsilon}\right)^{1/\beta} \right\rceil - 1$$

for $p = 1$

Specific Values of $\mathbf{dim}^{\mathbf{trnc}}(\varepsilon)$ for $p = 1$ and $\gamma_{\mathfrak{w}} = \prod_{j\in\mathfrak{w}} j^{-\beta}$

ε	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	
$\mathbf{dim}^{\mathbf{trnc}}(\varepsilon)$	2	9	31	99	316	$\beta = 2$
$\mathbf{dim}^{\mathbf{trnc}}(\varepsilon)$	2	4	9	21	46	$\beta = 3$
$\mathbf{dim}^{\mathbf{trnc}}(\varepsilon)$	1	3	5	9	17	$\beta = 4$

For instance, for the error demand $\varepsilon = 10^{-3}$ with $\beta = 4$,

only five variables instead of ∞-many!

Worst Case Error of QMC or Sparse Grids Methods is:

$$\leq O\left(\frac{\ln^4 n}{n}\right)$$

Multivariate Decomposition Method

MDM replaces

one ∞-variate integral

by

only few integrals

each with

only few variables

Introduced in [Kuo, Sloan, W., and Woźniakowski 2010]

Any function f has the unique

anchored decomposition

$$f(\boldsymbol{x}) = \sum_{\mathfrak{w}} f_{\mathfrak{w}}(\boldsymbol{x}_{\mathfrak{w}}),$$

where $f_{\mathfrak{w}}$ depends only on x_j with $j \in \mathfrak{w}$
and vanishes if $x_j = 0$.

General Idea:

- Select the **"most important"** $\mathfrak{w}$'s
- Approximate Integrals of $f_{\mathfrak{w}}$ **only** for the selected $\mathfrak{w}$'s

More precisely:

Given the **error demand** $\varepsilon > 0$,
construct an **"active set"** $\mathbf{Act}(\varepsilon)$
of subsets $\mathfrak{w}$ such that

$$\left| \mathcal{I} \left(\sum_{\mathfrak{w} \notin \mathbf{Act}(\varepsilon)} f_{\mathfrak{w}} \right) \right| \le \frac{\varepsilon}{2} \, \|f\|_{\mathcal{F}} \quad \text{for all } f \in \mathcal{F}.$$

Do nothing for integrals of $f_{\mathfrak{w}}$ with $\mathfrak{w} \notin \mathbf{Act}(\varepsilon)$

For $\mathfrak{w}\in\mathbf{Act}(\varepsilon)$, choose $n_{\mathfrak{w}}$ and cubatures $Q_{\mathfrak{w},n_{\mathfrak{w}}}$

to approximate integrals of $f_{\mathfrak{w}}$ such that

$$\sum_{\mathfrak{w}\in\mathbf{Act}(\varepsilon)} |\mathcal{I}(f_{\mathfrak{w}}) - Q_{\mathfrak{w},n_{\mathfrak{w}}}(f_{\mathfrak{w}})| \le \frac{\varepsilon}{2}\,\|f\|_{\mathcal{F}} \quad \text{for all } f \in \mathcal{F}.$$

The cubatures $Q_{\mathfrak{w},n_{\mathfrak{w}}}$ could be **QMC** or **Sparse Grids**

Then the **MDM** given by

$$\mathcal{Q}_\varepsilon(f) := \sum_{\mathfrak{w}\in\mathbf{Act}(\varepsilon)} Q_{\mathfrak{w},n_{\mathfrak{w}}}(f_{\mathfrak{w}})$$

has the **"worst case error"** bounded by ε, i.e.,

$$|\mathcal{I}(f) - \mathcal{Q}_\varepsilon(f)| \le \varepsilon \, \|f\|_{\mathcal{F}} \quad \text{for all } f \in \mathcal{F}.$$

How about the **COST**?

The number of integrals to approximate is small:

$$\mathbf{card}(\varepsilon) := |\mathbf{Act}(\varepsilon)| = \mathcal{O}\left(\frac{1}{\varepsilon}\right)$$

Each $f_{\mathfrak{w}}$ depends on only $|\mathfrak{w}|$ variables.

The largest number of variables is also small:

$$\mathbf{dim}(\varepsilon) := \max\{|\mathfrak{w}| \;:\; \mathfrak{w} \in \mathbf{Act}(\varepsilon)\} = \mathcal{O}(???)$$

The number of integrals to approximate is small:

$$\mathbf{card}(\varepsilon) := |\mathbf{Act}(\varepsilon)| = \mathcal{O}\left(\frac{1}{\varepsilon}\right)$$

Each $f_{\mathfrak{w}}$ depends on only $|\mathfrak{w}|$ variables.

The largest number of variables is also small:

$$\mathbf{dim}(\varepsilon) := \max\{|\mathfrak{w}| \ : \ \mathfrak{w} \in \mathbf{Act}(\varepsilon)\} = \mathcal{O}\left(\frac{\ln(1/\varepsilon)}{\ln(\ln(1/\varepsilon))}\right)$$

[Plaskota and W. 2011]

$\dim(\varepsilon)$ is like

Superposition Dimension in Statistics

Very efficient algorithm to construct $\mathbf{Act}(\varepsilon)$ in [Gilbert and W. 2016]

Specific Values of $\mathrm{dim}(\varepsilon)$ and $\mathrm{card}(\varepsilon)$ for $p = 1$ and $\gamma_{\mathfrak{w}} = \prod_{j\in\mathfrak{w}} j^{-\beta}$

ε	10^{-1}		10^{-2}		10^{-3}		10^{-4}		10^{-5}		
	2	6	3	22	4	113	4	534	5	2424	$\beta = 2$
	2	6	2	8	3	22	3	68	4	192	$\beta = 3$
	1	2	2	6	2	10	3	26	3	50	$\beta = 4$

For instance, for $\varepsilon = 10^{-3}$ with $\beta = 4$

it is sufficient to approximate

10 integrals with at most 2 variables!

Active Set $\mathrm{Act}(10^{-3})$

For $\beta = 4$

$\emptyset$,

$\{1\}, \{2\}, \{3\}, \{4\}, \{5\},$

$\{1,2\}, \{1,3\}, \{1,4\}, \{1,5\}$

For $\beta = 3$

$\emptyset$,

$\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{7\}, \{8\}, \{9\},$

$\{1,2\}, \{1,3\}, \{1,4\}, \{1,5\}, \{1,6\}, \{1,7\}, \{1,8\}, \{1,9\}, \{2,3\}, \{2,4\},$

$\{1,2,3\}, \{1,2,4\}$

For $\beta = 2$

31 of integrals with 1 variable

54 of integrals with 2 variables

26 of integrals with 3 variables

2 of integrals with 4 variables

$$\emptyset,$$
$$\{1\}, \ldots, \{31\},$$
$$\{1,2\}, \ldots, \{1,31\}, \{2,3\}, \ldots, \{2,15\},$$
$$\{3,4\}, \ldots, \{3,10\}, \{4,5\}, \{4,6\}, \{4,7\}, \{5,6\},$$
$$\{1,2,3\}, \ldots, \{1,2,15\}, \{1,3,4\}, \ldots, \{1,3,10\},$$
$$\{1,4,5\}, \{1,4,6\}, \{1,4,7\}, \{1,5,6\}, \{2,3,4\}, \{2,3,5\},$$
$$\{1,2,3,4\}, \{1,2,3,5\}$$

REMARK:

We do **NOT** know $f_{\mathfrak{w}}$ terms. However, we can sample them. Indeed due to [Kuo, Sloan, W., and Woźniakowski 2010b]

$$f_{\mathfrak{w}}(\boldsymbol{x}_{\mathfrak{w}}) = \sum_{\mathfrak{v}\subseteq\mathfrak{w}} (-1)^{|\mathfrak{w}|-|\mathfrak{v}|} f([\boldsymbol{x}_{\mathfrak{v}};\mathbf{0}])$$

requires

$2^{|\mathfrak{w}|}$ samples of f

REMARK:

We do **NOT** know $f_{\mathfrak{w}}$ terms. However, we can sample them. Indeed due to [Kuo, Sloan, W., and Woźniakowski 2010b]

$$f_{\mathfrak{w}}(\boldsymbol{x}_{\mathfrak{w}}) = \sum_{\mathfrak{v}\subseteq\mathfrak{w}} (-1)^{|\mathfrak{w}|-|\mathfrak{v}|} f([\boldsymbol{x}_{\mathfrak{v}};\mathbf{0}])$$

requires

$$2^{|\mathfrak{w}|} \quad \text{samples of } f$$

but from [Plaskota and W. 2011]

$$2^{|\mathfrak{w}|} = O\left(\varepsilon^{\frac{-1}{\ln(\ln(1/\varepsilon))}}\right) \quad \text{is small for modest } \varepsilon.$$

How About ANOVA Spaces?

$$f \in \mathcal{F}^{\mathrm{AVOVA}} \text{ iff}$$

$$f(\boldsymbol{x}) = \sum_{\mathfrak{w}} f_{\mathfrak{w},\mathrm{A}}(\boldsymbol{x}_{\mathfrak{w}})$$

with $\int_D f_{\mathfrak{w},\mathrm{A}}(\boldsymbol{x}_{\mathfrak{w}})\,\mathrm{d}x_j = 0$ and

$$\|f\|_{\mathcal{F}^{\mathrm{ANOVA}}} = \left(\sum_{\mathfrak{w}} \gamma_{\mathfrak{w}}^{-p}\,\|f_{\mathfrak{w},\mathrm{A}}^{(\mathfrak{w})}\|_{L_p}^{p}\right)^{1/p} < \infty$$

ANOVA decomposition terms $f_{\mathfrak{w},\mathrm{A}}$
cannot be sampled, i.e.,
low truncation dimension and MDM might
not be applicable.

Even worse: the 'easiest' (constant) term
is **NOT known**;
it is the integral we want:

$$f_{\emptyset,\mathrm{A}} = \mathcal{I}(f)$$

HOWEVER

If the spaces are EQUIVALENT, then

efficient algorithms for anchored spaces

are also efficient for ANOVA spaces

This motivated the study of

Equivalence of anchored and ANOVA Spaces

For product weights $\gamma_{\mathfrak{w}} = \prod_{j \in \mathfrak{w}} j^{-\beta}$

$$\mathcal{F} = \mathcal{F}^{\mathrm{ANOVA}} \quad \text{as sets.}$$

For the imbedding $\imath : \mathcal{F} \hookrightarrow \mathcal{F}^{\mathrm{ANOVA}}$ we have

$$\|\imath\| = \|\imath^{-1}\| \le \prod_{j=1}^{\infty} \left(1 + j^{-\beta}\right)$$

EQUIVALENCE iff $\beta > 1$

Research direction initiated in [Hefter and Ritter 2014],
Hilbert spaces setting $p = 2$ and product weights

[Hefter, Ritter and W. 2016]
$p \in \{1, \infty\}$ and general weights,

[Hinrichs and Schneider 2016]
$p \in (1, \infty)$,

[Gnewuch, Hefter, Hinrichs, Ritter, and W. 2016]
more general spaces,

[Kritzer, Pillichshammer, and W. 2017]
sharp lower bounds,

[Hinrichs, Kritzer, Pillichshammer, and W. 2017] most general

GENERALIZATIONS

More General Domain:

Any interval D including $D = \mathbb{R}$

More General Distributions μ on D:

e.g., Exponential, Gaussian

More General Integrals: $\int_{\mathbb{R}^{\mathbb{N}}} f(\boldsymbol{x})\, \mu^{\mathbb{N}}(\mathrm{d}\boldsymbol{x})$

Other Linear Solution Operators: $\mathcal{S}(f) = ???$

e.g., Function Approximation, ODE's, PDE's

General Information about f:

$L_1(f), L_2(f), \ldots, L_n(f), \qquad L_j \in \mathcal{F}^*$

Bayesian Approach:

Endowing $\mathcal{F}$ with
Gaussian probability measure PROB
and studying **average case errors:**

$$\int_{\mathcal{F}} \|\mathcal{S}(f) - Alg(L_1(f), \ldots, L_n(f))\|^p_{\mathcal{S}(\mathcal{F})} \, \mathrm{PROB}(\mathrm{d}f)$$

Similar results in [W. 2014]

Comments to Houman Owhadi's 1st talk:

[Traub, W., and Woźniakowski 1988] has a number of chapters devoted to the average, randomized and probabilistic settings for infinitely dimensional Hilbert and Banach spaces. They are based on a number of earlier papers. Currently there are 100's of IBC such papers, see e.g. 3 Volumes monograph: [E. Novak and H. Woźniakowski 2008-10]

On page 16, the **IBC Probabilistic Setting**
was attributed to H. Woźniakowski's paper.
However, as acknowledged in that paper, the results
were based on some of the results of my paper:
Optimal algorithms for linear problems with Gaussian measures,
Rocky Mountains J. of Math. 1986,
where **IBC Probabilistic Setting**
has been introduced for the first time.

Comment to Ilias Bilionis' talk:

Research that seem to be related:
IBC approach to PDE's with random coefficients
by Ch. Shwab and his collaborators,
e.g., F.Y.Kuo, D. Nuyens, I. H. Sloan

THANK YOU FOR THE ATTENTION