QMC beyond the cube

Art B. Owen, Stanford University
based on joint work with
Kinjal Basu, LinkedIn
Kinjal Basu

PhD 2016, Stanford Statistics
Estimate \( \mu = \int_{[0,1]^d} f(x) \, dx \) by \( \hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} f(x_i) \)

Koksma-Hlawka

\[ |\hat{\mu} - \mu| \leq D_n^*(x_1, \ldots, x_n) \times \|f\|_{\text{HK}} \]

Discrepancy is with respect to axis-oriented boxes \([0, a]\) or \([a, b]\)

Variation is based on axis-oriented differences of differences.
Non-cubic domains

\[ \mu = \int_{\Omega} f(x) \, dx \]

Triangle  Simplex  Cylinder
Disk  Sphere  Ball  Spherical triangle

What axes?

For discrepancy and variation

Cartesian products

\[ \Omega = \prod_{j=1}^{s} \Omega_j, \quad \Omega_j \subset \mathbb{R}^{d_j} \]

Disk \times Sphere \times Sphere \times Interval \times \cdots \times Spherical triangle
“You’ll never get out of the cube.”

The Cube (Jim Henson, 1969) is a surreal film about being stuck in a cube.

Image from wikipedia
General measures

\[ D_n^*(\cdot; \mu) = D_n^*(x_1, \ldots, x_n; \mu) \] is star discrepancy wrt measure \( \mu \)

**Theorem from ‘Gates of Hell’ paper**

Aistleitner, Bilyk & Nikolov (2016), For any normalized measure \( \mu \) on \( \mathbb{R}^d \) there exist points with

\[ D_n^*(\cdot; \mu) \leq \log(n)^{d-1/2}/n \]

**Refs from GoH paper**

- Aistleitner & Dick (2015) 
  discrepancy and Koksma-Hlawka for general signed measures.

- Aistleitner & Dick (2014) For any normalized measure \( \mu \) on \([0, 1]^d\),

\[ D_n^*(\cdot; \mu) \leq 63\sqrt{d}(2 + \log_2(n)^{(3d+1)/2})/n. \]

- Beck (1984) had \( \log(n)^{2d} \).

QMC sampling

We emphasize constructions

1) Measure preserving maps from $[0, 1]^d$ onto $\Omega$, and

2) Direct constructions, e.g., by recursively partitioning $\Omega$. 
Existence proofs

For users, they are frustrating.

• Constructions say how to do something.
  
  Yes! You can do this.
Existence proofs

For users, they are frustrating.

- Constructions say how to do something.
  Yes! You can do this.

- Non-existence results show that constructions don’t exist.
  No! You can’t do that.
Existence proofs

For users, they are frustrating.

- Constructions say how to do something.
  Yes! You can do this.

- Non-existence results show that constructions don’t exist.
  No! You can’t do that.

- Existence proofs show that non-existence proofs don’t exist.
  Maybe! Keep looking.

However

They can be interesting, elegant or deep.
(And may hint at constructions.)
Non-cubic domains

We want

\[ \mu = \int_{\Omega} f(\mathbf{x}) \, d\mathbf{x}, \quad \text{bounded } \Omega \subset \mathbb{R}^d, \quad \text{vol}(\Omega) = 1 \]

Transformations

For measure preserving \( \tau : [0, 1]^s \rightarrow \Omega \)

\[ \hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} (f \circ \tau)(\mathbf{x}_i), \quad \mathbf{x}_i \in [0, 1]^s \]

But \( f \circ \tau \) might not be well behaved. No problem for MC; challenge for QMC.

Choices for \( \tau \)

The triangle

Brandolini, Colzani, Gigante & Travaglini (2013)

- define a ‘trapezoid discrepancy’ in the simplex and a variation
- prove a Koksma-Hlawka inequality

but gave no constructions of points with vanishing discrepancy.

Pillards & Cools (2005)

- lots of measure preserving mappings
- get variation & discrepancy & Koksma-Hlawka

but gave no conditions for vanishing discrepancy of transformed points.

Chen & Travaglini (2013)

prove existence of point sets with vanishing trapezoid discrepancy for the triangle
Trapezoid discrepancy

Brandolini et al. (2013,2014)

\[ \Omega = \triangle (A, B, C) \]

Discrepancy for \( \mathcal{T}_{a,b,C} \cap \Omega \)

sup over trapezoids

Corresponding variation

Elegant argument \( \cdots \)

\( \cdots \) extends to simplices
Triangular van der Corput

For \( i \)'th point in \( T = \triangle(A, B, C) \), write

\[
i = \sum_{k=1}^{K_i} d_{k,i} 4^{k-1}, \quad d_{k,i} \in \{0, 1, 2, 3\}
\]

Split \( T \) into 4 congruent sub-triangles, \( T(0), T(1), T(2), T(3) \)
Place \( x_i \) in \( T(d_{1,i}) \)

Recurse

Basu & O (2015)
Construction continued

Corners of the subtriangle

\[ T(d) = \begin{cases} \triangle \left( \frac{B+C}{2}, \frac{A+C}{2}, \frac{A+B}{2} \right), & d = 0, \\ \triangle \left( A, \frac{A+B}{2}, \frac{A+C}{2} \right), & d = 1, \\ \triangle \left( \frac{A+B}{2}, B, \frac{B+C}{2} \right), & d = 2, \\ \triangle \left( \frac{A+C}{2}, \frac{B+C}{2}, C \right), & d = 3. \end{cases} \]
For $n = 4^k$

- $n$ subtriangles, 1 point each
- all discrepancy from within shaded triangles
- enumerate all possibilities
- upright vs inverted are different
Results

Let $D^P_n$ be (anchored) parallelogram discrepancy.

First $n = 4^k$ points

$$D^P_n = \begin{cases} 
\frac{7}{9}, & n = 1 \\
\frac{2}{3\sqrt{n}} - \frac{1}{9n}, & \text{else}
\end{cases}$$

Any consecutive $n = 4^k$ points

$$D^P_n \leq \frac{2}{\sqrt{n}} - \frac{1}{n}$$

First $n$ points

$$D^P_n \leq \frac{12}{\sqrt{n}}$$

Basu & O (2015)
Kronecker lattice in the triangle

Basu & O (2015)

1) Place a square grid in $\mathbb{R}^2$

2) Rotate it $\alpha$ radians

3) Intersect with right triangle

4) Linear map to desired $\triangle$

Critical: choose good $\alpha$
Kronecker continued

$\theta \in \mathbb{R}$ is **badly approximable** if there exists $c > 0$ with

$$\text{dist}(n\theta, \mathbb{Z}) > c/n, \quad \forall n \in \mathbb{N}$$

**Quadratic irrationals** $\theta = (a + b\sqrt{c})/d$ are badly approximable.

Here $a, b, c, d \in \mathbb{Z}$, $b, d \neq 0$, square free $c > 1$

Chen & Travaglini (2007) There exist points with

Polygon discrepancy $= O(\log(n)/n)$

Basu & O (2015) Here they are for trapezoids: rotate a grid by $\alpha$ radians where $\tan(\alpha)$ is a quadratic irrational.

E.g., for $\alpha = 3\pi/8$, $\tan(\alpha) = 1 + \sqrt{2}$
Triangular Kronecker

Triangular lattice points

A grid with a ‘Kronecker rotation’ gets \( D_n^P = O(\log(n)/n) \). Basu & O (2015)

This is the best possible rate. Chen & Travaglini (2013)

Generalization

Hexagon = six triangles, et cetera

Very unlikely to generalize to higher dimensional simplices or Cartesian products of simplices. (D. Bilyk personal communication)
Geometric van der Corput

Map $i = 1, 2, 3, \ldots$ into $x_i \in \Omega$.

- replace triangle by more general set $\Omega$
- split $\Omega$ into $b$ equal volumes
- recursively
The triangle can be recursively split 2-fold, 3-fold or 4-fold.

This allows digital constructions in those bases.
Not all splits work well

The base 3 split leads to very unfavorable aspect ratios. The regions do not ‘converge nicely’ to a point.

E.g., Stromberg (1994) defines ‘converge nicely’ (Bounded aspect ratios.)
Splits don’t have to be congruent

- Mix ‘arc splits’ and ‘radial splits’ to keep aspect ratio bounded
- Not a global alternation; different cells get different splits

Basu & O (2015)
See Beckers & Beckers (2012) for non-recursive splits
Tetrahedron

- chop off 4 tetrahedral corners
- remaining volume makes 4 more

- but they’re not congruent to first 4
- binary splits may be better (split a longest edge)

Image: By Tomruen - Own work, CC BY-SA 3.0, Wikipedia
Spherical triangles

- 4 way split at arc midpoints \( \cdots \) not equal area

- 4 way equal area split of Song, Kimerling, Sahr (2002) uses ‘small circle’ boundaries, not great circles

- binary splits may be better \( \cdots \) use first step in Arvo (1995)

More about Arvo

Arvo shows how to pick \( D \) so

\[
\frac{\text{vol}(ABD)}{\text{vol}(ABC')} = u
\]

We can use \( u = 1/2 \).

Image by Peter Mercator - Own work, CC BY-SA 3.0, Wikipedia
Geometric nets

We want points in $\Omega^s$ for $\Omega \subset \mathbb{R}^d$

E.g., light path

$$\text{camera} \rightarrow \triangle \rightarrow \triangle \rightarrow \triangle \rightarrow \cdots \rightarrow \triangle \rightarrow \text{light source}$$

Use digital nets

A $(t, m, s)$-net, $b = 4$ or $b = 2$, puts $x_i \in \triangle^s$ (componentwise)

Use other partitions

Other $b$-fold equal area recursive partitions can be used for $\Omega \neq \triangle$

Scramble the nets

Unbiasedness and error cancellation benefits under smoothness.
More generally

\[ \Omega = \prod_{j=1}^{s} \Omega_j, \quad \Omega_j \subset \mathbb{R}^{d_j} \]

\[ \tau_j : [0, 1] \rightarrow \Omega_j \quad \text{digital map, base } b \]

Take \( \mathbf{u}_i = (u_{i1}, \ldots, u_{is}) \in [0, 1]^s \),

\((t, m, s)\)-net or \((t, s)\)-sequence in base \( b \).

Componentwise map: \( \mathbf{x}_i = \tau(\mathbf{u}_i) \)

\[ \mathbf{x}_i = (x_{i1}, \ldots, x_{is}) \]

\[ x_{ij} = \tau_j(u_{ij}) \]
Scrambled geometric nets

Take \( \operatorname{vol}(\Omega_j) = 1 \) and \( \Omega = \prod_{j=1}^{s} \Omega_j \) and let

\[
\mu = \int_{\Omega} f(\boldsymbol{x}) \, d\boldsymbol{x}, \quad \hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} f(\boldsymbol{x}_i)
\]

where \( \boldsymbol{x}_i \) are scrambled geometric nets.

For \( f \in L^2(\Omega) \)

\[
\mathbb{E}(\hat{\mu}) = \mu \quad \text{Var}(\hat{\mu}) = o\left(\frac{1}{n}\right) \quad \text{Var}(\hat{\mu}) \leq \Gamma \times \frac{\sigma^2}{n}
\]

where \( \sigma^2 = \int_{\Omega} (f(\boldsymbol{x}) - \mu)^2 \, d\boldsymbol{x} \), and \( \Gamma \) is the largest gain coefficient of the \((t, m, s)\)-net

E.g., \( t = 0 \) implies \( \Gamma \leq \exp(1) \approx 2.718 \)
Convergence rates

\[ \mu = \int_{\Omega} f(x) \, dx, \quad \Omega \subset \mathbb{R}^D, \quad D = \sum_{j=1}^{s} d_j, \quad \text{e.g., } D = s \times d. \]

For smooth \( f \), nested uniform scrambled nets and nice partitions

\[ \text{Var}(\hat{\mu}) = O\left(\left(\frac{\log n}{n^{1+2/d}}\right)^{s-1}\right) \]

Basu & O (2015)

Two kinds of smooth

1) \( \partial^{1:D} f \) continuous and all \( \Omega_j \) Sobol’ extensible (defined next)

2) \( f \in C^{D} (\Omega) \) (using the Whitney extension)
**Sobol’ extension**

It begins with the fundamental theorem of calculus (FTC)

\[
f(x) = f(c) + \int_{c}^{x} f'(y) \, dy
\]

Dimension \( D \), e.g., \( D = d \times s \)

\[
f(x) = f(c) \text{ plus } 2^D - 1 \text{ integrated partial derivatives along all 'lower faces'
}\]

\[
f(x) = \sum_{u \subseteq 1:D} \int_{[c_u, x_u]} \partial^u f(c_{-u}:y_u) \, dy_u
\]

Hybrid points

For \( x, y \in \mathbb{R}^D \) and \( u \subset \{1, 2, \ldots, D\} \) the point \( z = x_u : y_{-u} \) has \( z_j = x_j \)

for \( j \in u \) and \( x_j = y_j \) for \( j \not\in u \).
Sobol’ extensible region

For $x, c \in \mathbb{R}^D$ define $\text{rect}[c, x] = \prod_{j=1}^{D} [c_j \wedge x_j, c_j \vee x_j]$

Definition

$\Omega \subset \mathbb{R}^d$ is Sobol’ extensible with anchor $c \in \mathbb{R}^D$ if

$$x \in \Omega \implies \text{rect}[c, x] \subset \Omega$$
Sobol’ extensible
Non-Sobol’ extensible

No place to put the anchor \( c \)
Sobol’ extension

Let $\Omega \subset \mathbb{R}^D$ be Sobol’ extensible with anchor $c$ and let $\partial^{1:D} f$ be continuous. Then the Sobol’ extension of $f$ is

\[
\tilde{f}(x) = \sum_{u \subseteq 1:D} \int_{[c_u, x_u]} \partial^u f(c_{-u} : y_u) 1_{c_{-u} : y \in \Omega} \, dy_u
\]

vs.

\[
f(x) = \sum_{u \subseteq 1:D} \int_{[c_u, x_u]} \partial^u f(c_{-u} : y_u) \, dy_u \quad \text{(FTC)}
\]

Properties

\[
\tilde{f}(x) = f(x) \text{ for } x \in \Omega
\]

Low variation

$\partial^{1:D} \tilde{f}$ not necessarily continuous

but $\tilde{f}$ satisfies the FTC
Conditions

1) \( \Omega \subset \mathbb{R}^d \) bounded and Sobol’ extensible
2) \( x_i \) a geometric net, bounded ‘gain’ coefs
3) nice convergent splits
4) \( f \in L^2(\Omega^s) \)
5) \( \partial^{1:s} f \) continuous

Conclusion

\[
\text{Var}(\hat{\mu}) = O\left(\frac{(\log n)^{s-1}}{n^{1+2/d}}\right)
\]
as \( n = b^m \to \infty \). Basu & O (2015)

Challenge:

showing Haar wavelet coefficients decay
via Sobol’ (or Whitney) extension from \( \Omega \) to \( \mathbb{R}^d \)
Dimension $D = ds$

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

Comparisons

1) Better than MC rate for all $s, d$

2) Rate sensitive to $d$, not very sensitive to $s$

3) Better than QMC rate for BVHK on $[0, 1]^D$ when $d = 2$ (just barely)
$$\left(\log(n)\right)^{(s-1)/2} \text{ vs } (\log(n))^{ds-1}$$
Often $f \circ \tau \not\in \text{BVHK}$

4) (Barely) better than Kronecker $\triangle$ for $d = 2$ and $s = 1$ (was $\log(n)/n$)

Note

$$g(x) = 1_{x \in \text{rect } \Omega} \times f(x), \quad \text{usually not BVHK}$$
Followups to geometric nets

- maybe higher order nets would help Dick, Baldeaux
- geometric Halton sequences
- deterministic nets

Central limit theorem

Transformations

Let \( \tau \) transform \( \mathbb{U}[0, 1]^m \) into \( \mathbb{U}(\Omega) \).

\[
\int_{\Omega} f(x) \, dx = \int_{[0,1]^m} (f \circ \tau)(u) \, du
\]

We want \( f \circ \tau \in \text{BVHK} \) for QMC and mixed partials in \( L^2 \) for RQMC

BVHK compositions

For \( f \circ \tau : \mathbb{R} \to \mathbb{R} \to \mathbb{R} \):

\[
f \in \text{Lipschitz}, \quad \tau \in \text{BV} \quad \Rightarrow \quad f \circ \tau \in \text{BV}.
\]

Josephy (1981)

No such simple rule in higher dimensions.

Variation is bounded via integrated absolute mixed partials.

So we study derivatives of \( f(\tau(u)) \).
Faà di Bruno

Derivatives of composite functions, $\mathbb{R} \to \mathbb{R} \to \mathbb{R}$

Faà di Bruno (1855,1857), Arbogast (1800)

$$h(x) = f(g(x))$$
$$h'(x) = f'(g(x))g'(x)$$
$$h''(x) = f''(g(x))g'(x)^2 + f'(g(x))g''(x)$$
$$h'''(x) = f'''(g(x))g'(x)^3 + 3f''(g(x))g'(x)g''(x) + f'(g(x))g'''(x)$$

Our map is
$$\mathbb{R}^D \to \mathbb{R}^d \to \mathbb{R}$$

which has many more terms

Constantine & Savits (1996) give a general Faà di Bruno theorem

Basu & O (2016) simplify it for
$$\partial^u (f \circ \tau), u \subseteq \{1, \ldots, D\}$$
i.e., differentiate at most once wrt each $x_j$

Allows tests of BVHK.
Some mappings

The following mappings work well for MC, but not QMC

**Triangle** $\mathbb{T}^2 \subset \mathbb{R}^3$

$$\boldsymbol{u} \in [0, 1]^3, \quad x_j = \tau_j(\boldsymbol{u}) = \frac{\log(u_j)}{\sum_{i=1}^3 \log(u_i)} \quad \boldsymbol{x} \sim \mathbf{U}(\mathbb{T}^2)$$

Even $x_j(\boldsymbol{u}) \notin \text{BVHK}([0, 1]^3)$.

**Sphere** $\mathbb{S}^{d-1} \subset \mathbb{R}^d$

$$x_j = \tau_j(\boldsymbol{u}) = \frac{\Phi^{-1}(u_j)}{\sqrt{\sum_{i=1}^d \Phi^{-1}(u_i)^2}}, \quad \boldsymbol{x} \sim \mathbf{U}(\mathbb{S}^{d-1})$$

Again, $x_j(\boldsymbol{u}) \notin \text{BVHK}([0, 1]^d)$. 
BVHK compositions

For \( u \in [0, 1]^D \) and
\[
f(\tau_1(u), \ldots, \tau_d(u))
\]
If these hold

1) \( \partial^v \tau_j(u_v:1_{-v}) \in L^{p_j}([0, 1]^{|v|}), \quad p_j \in [1, \infty] \quad v \subseteq \{1, 2, \ldots, D\} \)

2) \( \sum_{j=1}^d 1/p_j \leq 1 \)

3) \( f \in C^{(d)}(\mathbb{R}^d) \)

Then
\[
f \circ \tau \in \text{BVHK}
\]
RQMC smooth

1) $\partial^v \tau_j \in L^{p_j}([0, 1]^D)$, $p_j \in [2, \infty]$, and
2) $\sum_{j=1}^{d} 1/p_j \leq 1/2$
3) $f \in C^{(d)}(\mathbb{R}^d)$

make $f \circ \tau$ smooth enough for RMSE = $O(n^{-3/2+\epsilon})$ under RQMC.

$f \in C^{(d)}$ can be weakened if $p_j$ are increased
Fang & Wang (1993)

Three mappings to a simplex, one to the sphere, and one to a ball.

Example

\[ A_d = \{(x_1, \ldots, x_d) \mid 0 \leq x_1 \leq x_2 \leq \cdots \leq x_d \leq 1\} \]

Transformation

\[ x_1 = \tau_1(u) = u_1 \]
\[ x_2 = \tau_2(u) = u_1 \times u_2^{1/2} \]
\[ x_3 = \tau_3(u) = u_1 \times u_2^{1/2} \times u_3^{1/3} \]
\[ \vdots \]
\[ x_d = \tau_d(u) = u_1 \times u_2^{1/2} \times u_3^{1/3} \times \cdots \times u_d^{1/d} \]
Results

All five Fang & Wang mappings $\tau$ are in BVHK.

So composing with $f$ has a chance.

None of them yield $\tau$ with mixed partials in $L^2$. 
Smoother mappings

Importance sampling from \([0, 1]^d\) to \(\mathbb{T}^d\) (simplex) can yield RQMC smoothness.

The Jacobian exhibits a ‘dimension’ effect.

Effective sample size decays like \((8/9)^d\).

Basu & O (2016)

Conclusion

The unit cube seems to be a relatively easy space to sample.
Despite GoH conjecture that it is the hardest.
Thanks

- Lecturers: Nicolas Chopin, Mark Huber, Jeffrey Rosenthal
- Guest speakers: Michael Giles, Gareth Roberts
- The London Mathematical Society: Elizabeth Fisher, Iain Stewart
- CRISM & The University of Warwick, Statistics
- Sponsors: Amazon, Google
- Partners: ISBA, MCQMC, BAYSM
- Poster: Talissa Gasser, Hidamari Design
- NSF: DMS-1407397 & DMS-1521145
- Planners: Murray Pollock, Christian Robert, Gareth Roberts
- Support: Paula Matthews, Murray Pollock, Shahin Tavakoli