The rate of convergence of sparse grid quadrature on the torus

Paul Leopardi

Mathematical Sciences Institute, Australian National University.
For presentation at CTAC 2010, UNSW.
Joint work with Markus Hegland, ANU.

November 2010 (revised December 2010)

AUSTRALIAN RESEARCH COUNCIL
Centre of Excellence for Mathematics and Statistics of Complex Systems

ARC

ANU
THE AUSTRALIAN NATIONAL UNIVERSITY

Topics

- Weighted tensor product spaces on products of spheres
- Component-by-component construction
- Weighted tensor product quadrature
- Numerical results for the products of the 2-sphere
- Numerical results for the torus

Reproducing kernel Hilbert space H on M

A Reproducing Kernel Hilbert Space (RKHS) H of real functions on a manifold M is a Hilbert space with inner product $\langle , \rangle$ and a kernel

$$K : M \times M \to \mathbb{R},$$

such that for all $x \in M$, if k_x is defined by

$$k_x(y) := K(x, y) \quad \text{for all } y \in M, \text{ then}$$
$$k_x \in H \quad \text{and} \quad \langle k_x, f \rangle = f(x) \text{ for all } f \in H.$$

KS function space $H_{1,\gamma}^{(s,r)}$ on a single sphere

For $f \in L_2(\mathbb{S}^s)$, $f(x) \sim \sum_{\ell=0}^{\infty} \sum_{k=1}^{N_\ell^{(s+1)}} \hat{f}_{\ell,k} Y_{\ell,k}^{(s+1)}(x)$.

For positive weight γ, define the RKHS

$$H_{s,1,\gamma}^{(s,r)} := \{f : \mathbb{S}^s \to \mathbb{R} \mid \|f\|_{1,\gamma} < \infty\},$$

where $\|f\|_{1,\gamma} := \langle f, f \rangle_\gamma^{1/2}$ and

$$\langle f, g \rangle_{1,\gamma} := \hat{f}_{0,0}\, \hat{g}_{0,0} + \gamma^{-1} \sum_{\ell=1}^{\infty} \sum_{k=1}^{N_\ell^{(s+1)}} \big(\ell(\ell + s - 1)\big)^r \hat{f}_{\ell,k}\, \hat{g}_{\ell,k}.$$

(Kuo and Sloan, 2005)

Reproducing kernel of $H^{(s,r)}_{1,\gamma}$

This is

$$K^{(s,r)}_{1,\gamma}(x, y) := 1 + \gamma A_{s,r}(x \cdot y), \quad \text{where for } z \in [-1, 1],$$
$$A_{s,r}(z) := \sum_{\ell=1}^{\infty} \frac{N^{(s+1)}_{\ell}}{\big(\ell(\ell + s - 1)\big)^{r}} P_{\ell}(z),$$

where P is an ultraspherical polynomial, scaled appropriately.

(Kuo and Sloan, 2005)

The weighted tensor product space $H_{d,\gamma}^{(s,r)}$

For $\gamma := (\gamma_1, \ldots, \gamma_d)$, on $(\mathbb{S}^s)^d$ define the tensor product space

$$H_{d,\gamma}^{(s,r)} := \bigotimes_{j=1}^{d} H_{1,\gamma_j}^{(s,r)} .$$

Reproducing kernel of $H_{d,\gamma}^{(s,r)}$ is

$$K_{d,\gamma}^{(s,r)}(x,y) := \prod_{j=1}^{d} K_{s,1,\gamma_j}^{(s,r)}(x_j, y_j)$$

(Kuo and Sloan, 2005)

Equal weight quadrature error on $H_{d,\gamma}^{(s,r)}$

Worst case error of equal weight m-point quadrature $Q_{m,d,\gamma}$:

$$e^2(Q_{m,d,\gamma}^{(s,r)}) := \sup_{\|f\|_{H_{d,\gamma}^{(s,r)}} \leq 1} \Big((\mathbb{I} - Q_{m,d,\gamma}^{(s,r)}) f \Big)^2$$

$$= -1 + \frac{1}{m^2} \sum_{i=1}^{m} \sum_{h=1}^{m} K_{d,\gamma}^{(s,r)}(x_i, x_h).$$

Expected worst case squared error for m equidistributed points:

$$E\big(e^2(Q_{m,d,\gamma}^{(s,r)})\big) = \frac{1}{m}\Big(-1 + \prod_{j=1}^{d} \big(1 + \gamma_j A_{s,r}(1)\big)\Big)$$

$$\leq \frac{1}{m} \exp\big(A_{s,r}(1) \sum_{j=1}^{d} \gamma_j\big).$$

(Kuo and Sloan, 2005)

Construction using permutations

The idea of Hesse, Kuo and Sloan, 2007 for quadrature on $(\mathbb{S}^2)^d$ is to use a spherical design $z = (z_1, \ldots, z_m)$ of strength t for the first sphere and then successively permute the points of the design to obtain the coordinates for each subsequent sphere.

The algorithm chooses permutations $\Pi_1, \ldots, \Pi_d : 1 \ldots m \to 1 \ldots m$, giving

$$x_i = (z_{\Pi_1(i)}, \ldots, z_{\Pi_d(i)})$$

to ensure that the resulting squared worst case quadrature error is better than the average $E\big(e^2(Q^{(2,r)}_{m,d,\gamma})\big)$.

(Hesse, Kuo and Sloan, 2007)

Weighted Korobov spaces on $\mathbb{T}^d$

Consider $s = 1$. $H^{(1,r)}_{1,\gamma}$ is a RKHS on the unit circle $\mathbb{S}^1 = \mathbb{T}$,

$H^{(1,r)}_{d,\gamma}$ is a RKHS on the d-torus $\mathbb{T}^d$.

This is a weighted Korobov space of periodic functions on $[0, 2\pi)^d$.

The Hesse, Kuo and Sloan construction in these spaces gives a rule with the same 1-dimensional projection properties as a lattice rule: the points are equally spaced.

(Wasilkowski and Woźniakowski, 1999; Hesse, Kuo and Sloan, 2007)

General quadrature weights on $H_{d,\gamma}^{(s,r)}$

For $X := \{x_1, \ldots, x_m\}$, if we define

$$Q_w f := \sum_{k=1}^{m} w_k f(x_k),$$
$$G_{i,j} := \langle k_{x_i}, k_{x_j} \rangle = K_{d,\gamma}^{(s,r)}(x_i, x_j),$$

then the worst case error e_w for Q_w satisfies

$$e_w^2 = \|1 - Q_w\|^2 = \langle 1 - Q_w, 1 - Q_w \rangle$$
$$= 1 - 2 \sum_{k=1}^{m} w_k + w^T G w.$$

Optimal quadrature weights on $H^{(s,r)}_{d,\gamma}$

Since

$$e_w^2 = 1 - 2\sum_{k=1}^{m} w_k + w^T G w,$$

the weights w are optimal when $Gw = [1, \ldots, 1]^T$.

In this case, $e_w^2 = 1 - \sum_{k=1}^{m} w_k$.

The Smolyak construction on $(\mathbb{S}^1)^d = \mathbb{T}^d$

The Smolyak construction and variants have been well studied on unweighted and weighted Korobov spaces.

Smolyak construction (unweighted Korobov space case):
For $H_{1,1}^{(1,r)}$, define $Q_{1,-1} := 0$ and define a sequence of equal weight rules $Q_{1,0}, Q_{1,1}, \ldots$ on $[0, 2\pi)$, exact for trigonometric polynomials of degree $t_0 = 0 < t_1 < \ldots$.

Define $\Delta_q := Q_{1,q} - Q_{1,q-1}$ and for $H_{d,1}^{(1,r)}$, define

$$Q_{d,q} := \sum_{0 \leq a_1 + \ldots + a_d \leq q} \Delta_{a_1} \otimes \ldots \otimes \Delta_{a_d}.$$

(Smolyak, 1963; Wasilkowski and Woźniakowski, 1995; Gerstner and Griebel, 1998)

The WTP variant of Smolyak on $H^{(1,r)}_{d,\gamma}$

The WTP algorithm of Wasilkowski and Woźniakowski (1999) generalizes Smolyak by treating spaces of non-periodic functions, by allowing optimal weights, and by allowing other choices for the index sets a.

For $H^{(1,r)}_{d,\gamma}$, define

$$W_{d,n} := \sum_{a \in P_{n,d}(\gamma)} \Delta_{a_1} \otimes \ldots \otimes \Delta_{a_d},$$

where $P_{1,d}(\gamma) \subset P_{2,d}(\gamma) \subset \mathbb{N}^d, \ |P_{n,d}(\gamma)| = n$.

W and W (1999) suggests to define $P_{n,d}(\gamma)$ by including the n rules $\Delta_{a_1} \otimes \ldots \otimes \Delta_{a_d}$ with largest norm.

(Wasilkowski and Woźniakowski, 1999)

WTP rules using spherical designs

For $H_{d,\gamma}^{(2,r)}$ we can define a WTP rule based on spherical designs. Define a sequence of optimal weight rules $Q_0, Q_1, \ldots$ using unions of spherical designs of increasing strength $t_0 = 0 < t_1 < \ldots$ and cardinality $m_0 = 1 < m_1 < \ldots$.

The WTP construction then proceeds similarly to $\mathbb{S}^1$.

One difference between $\mathbb{S}^1$ and $\mathbb{S}^2$ is that the spherical designs themselves cannot be nested in general.

(Wasilkowski and Woźniakowski, 1999)

Generic WTP algorithm for $\mathbb{S}^2$

1. Begin with a sequence of spherical designs $X_1, X_2, \ldots X_L$, with increasing cardinality, nondecreasing strength.
2. For each h, form the optimal weight rule Q_h from the point set $\bigcup_{i=1}^{h} X_i$, and the difference rule $\Delta_h = Q_h - Q_{h-1}$.
3. Form products of the difference rules and rank them in order of decreasing norm divided by the number of additional points.
4. Form WTP rules by adding product difference rules in rank order.

The Hesse, Kuo and Sloan example space

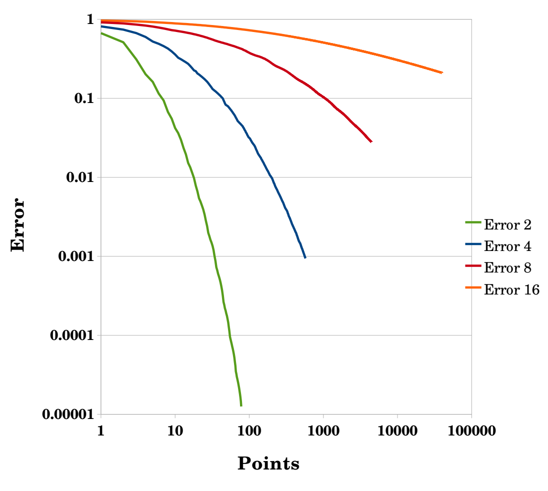
In Hesse, Kuo and Sloan, a numerical example is given with $r = 3$, $\gamma_j = 0.9^j$. In other words,

$$K^{(2,3)}_{d,\gamma}(x,y) := \prod_{j=1}^{d} K^{(2,3)}_{1,0.9^j}(x_j,y_j) = \prod_{j=1}^{d} \left(1 + 0.9^j A_{2,3}(x_j \cdot y_j)\right),$$

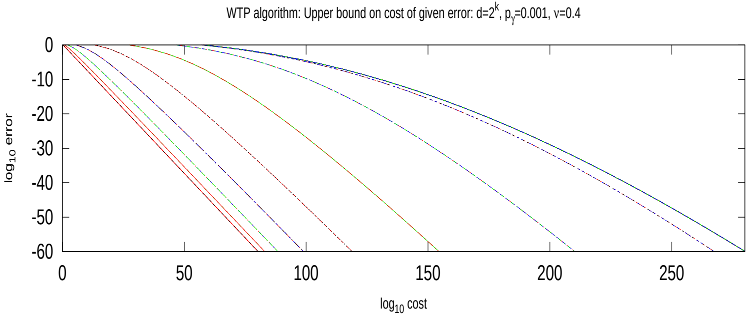
where

$$A_{2,3}(z) = \sum_{\ell=1}^{\infty} \frac{2\ell+1}{\left(\ell(\ell+1)\right)^3} P_\ell(z).$$

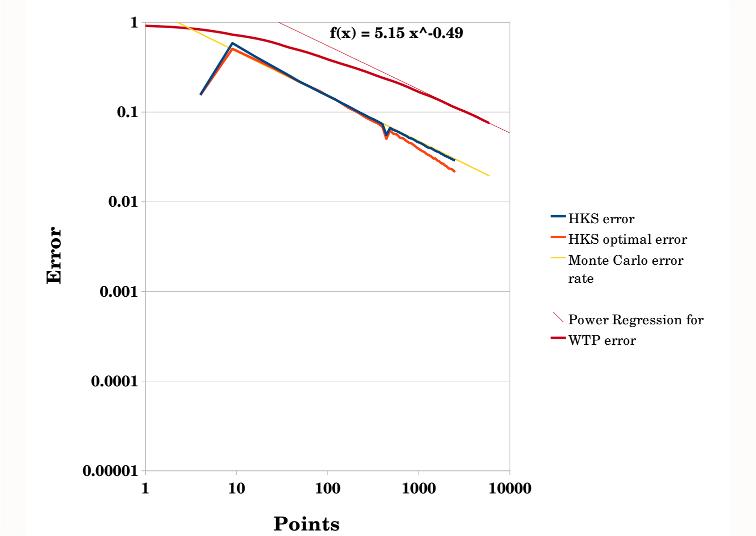
Error of WTP rule for $(\mathbb{S}^2)^d, d = 2, 4, 8, 16$



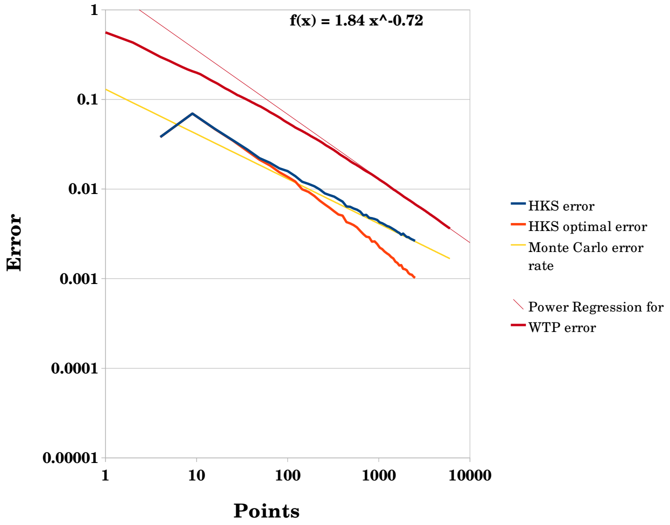
Estimated upper bound of error of WTP rule



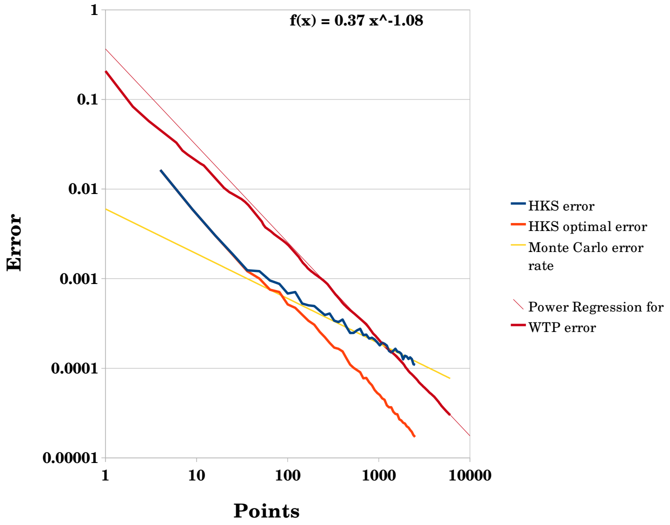
HKS vs WTP: $(\mathbb{S}^2)^8, r = 3, g = 0.9, \gamma = g^j$



HKS vs WTP: $(\mathbb{S}^2)^8, r = 3, g = 0.5, \gamma = g^j$



HKS vs WTP: $(\mathbb{S}^2)^8, r = 3, g = 0.1, \gamma = g^j$



Why does WTP (initially) perform poorly?

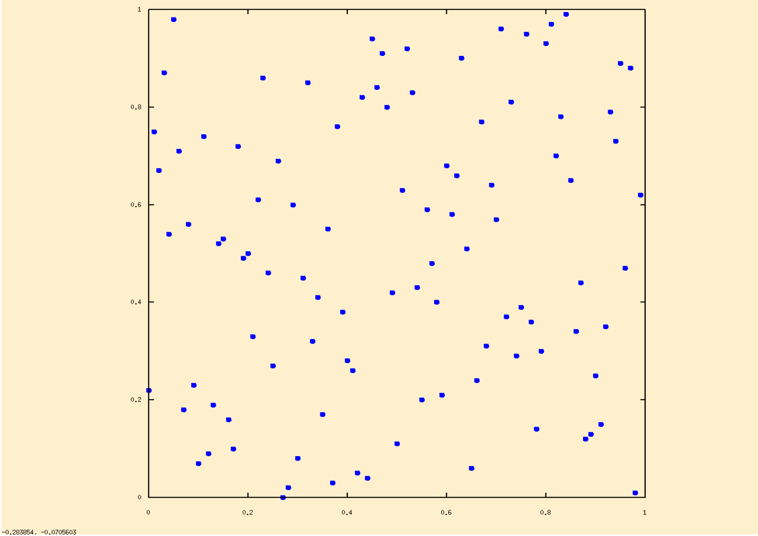
WTP points are too close together.

- Partly [??] because, for one sphere, nesting is forced.
- Mostly [??] because, for higher d, initially only one sphere at a time is changed.

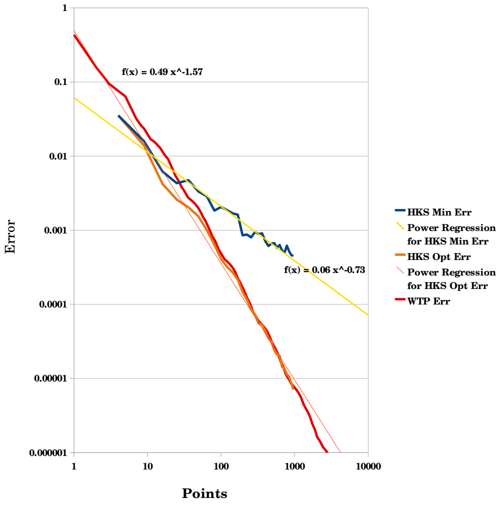
HKS points are better separated. [??]

[Not always: let's look at $(\mathbb{S}^1)^d = \mathbb{T}^d$.]

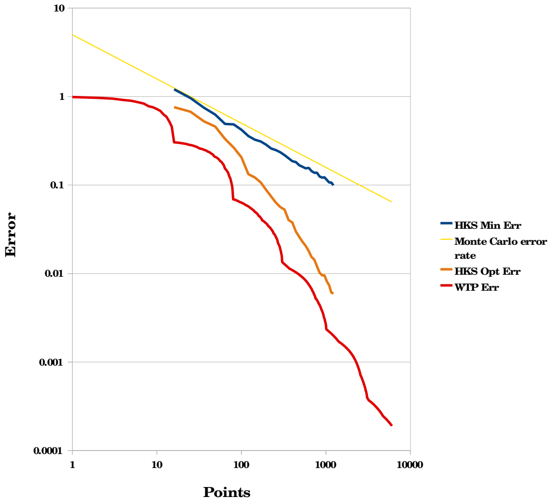
HKS: $\mathbb{T}^2, r = 3, \gamma = 0.9^j$, 100 points



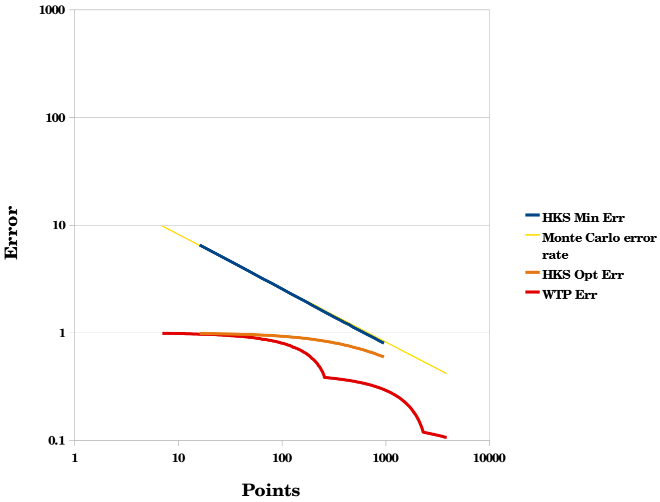
HKS vs WTP: $\mathbb{T}^4, r = 3, g = 0.1, \gamma = g^j$



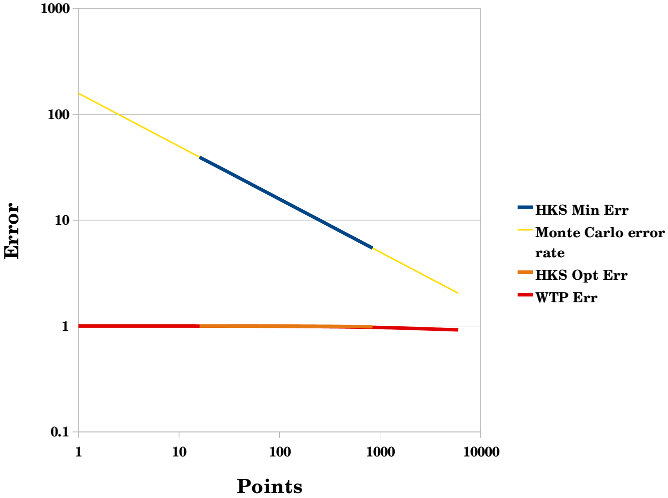
HKS vs WTP: $\mathbb{T}^4, r = 3, g = 0.9, \gamma = g^j$



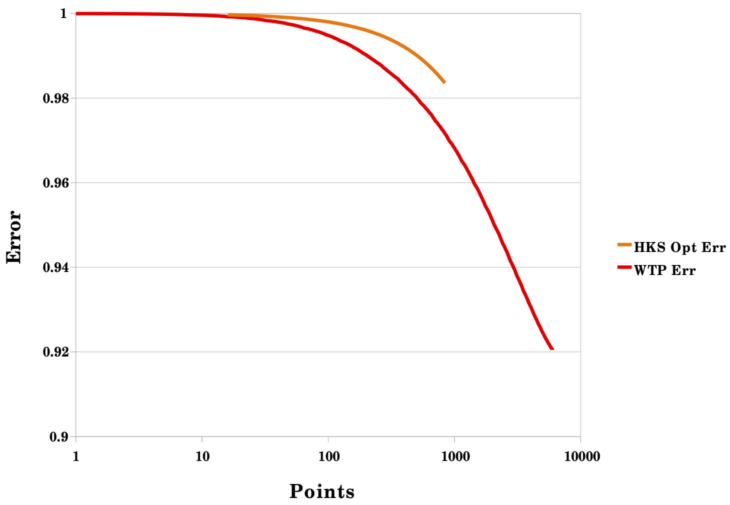
HKS vs WTP: $\mathbb{T}^8, r = 3, g = 0.9, \gamma = g^j$



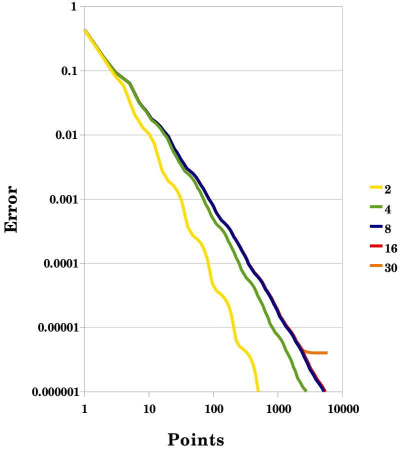
HKS vs WTP: $\mathbb{T}^{16}, r = 3, g = 0.9, \gamma = g^j$



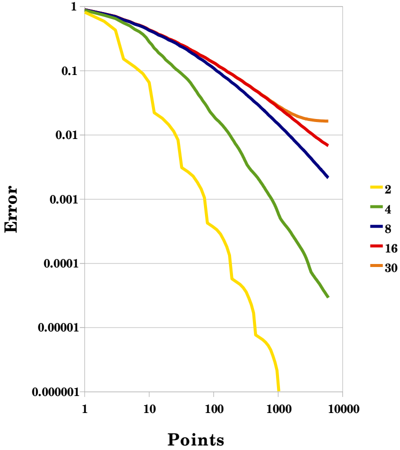
Closeup: $\mathbb{T}^{16}, r = 3, g = 0.9, \gamma = g^j$



Error of WTP: $\mathbb{T}^d, d = 2, 4, 8, 16, 30,\ g = 0.1$



Error of WTP: $\mathbb{T}^d, d = 2, 4, 8, 16, 30,\ g = 0.5$



Error of WTP: $\mathbb{T}^d, d = 2, 4, 8, 16, 30,\ g = 0.9$