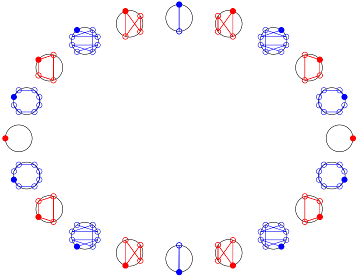
Polynomial copy rules in Walsh spaces

Dirk Nuyens[a,b]

Joint work with Ronald Cools[b]

[a]School of Mathematics & Statistics, University of NSW, Australia

[b]Department of Computer Science, K.U.Leuven



IBC Workshop, Foundations of Computational Mathematics

FoCM'08, City University of Hong Kong, 16–26 June 2008

Outline

Lattice rules

We try to aproximate the s-dimensional integral

$$I(f) := \int_{[0,1)^s} f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}$$

by an n-point quasi-Monte Carlo rule

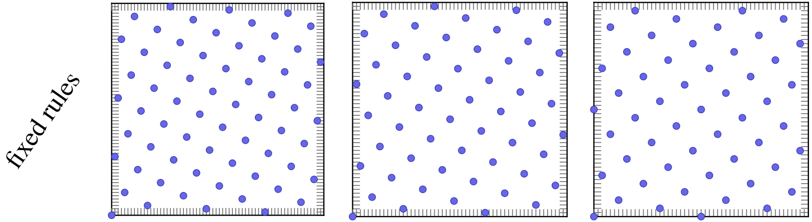
$$Q(f) := \frac{1}{n} \sum_{k=0}^{n-1} f(\boldsymbol{x}^{(k)}).$$

The point set $\{\boldsymbol{x}^{(0)}, \boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n-1)}\}$ could be taken a *lattice rule* where

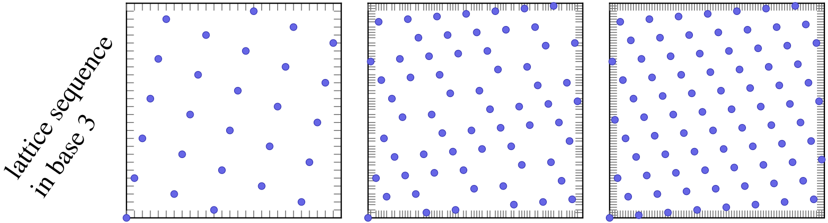
$$\boldsymbol{x}^{(k)} := \frac{k\,\boldsymbol{g} \bmod n}{n},$$

with *generating vector* $\boldsymbol{g} \in (\mathbb{Z}_n^{\times})^s$.

Lattice rules and lattice sequences look like...



(a) rank-1 rule (b) Fibonacci lattice (c) rank-2 copy rule



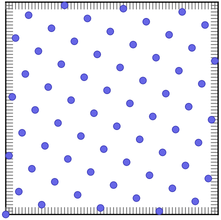
(d) 3^3 sequence points (e) 64 sequence points (f) 3^4 sequence points

In its simplest form: rank-1 rules

The points of a *rank*-1 *lattice* are just multiples of an integer vector $\boldsymbol{g}$:

$$P_n(\boldsymbol{z}) := \left\{ \left\{ \frac{k\,\boldsymbol{g}}{n} \right\} : k \in \mathbb{Z}_n \right\} = \left\{ \frac{k\,\boldsymbol{g} \bmod n}{n} : k \in \mathbb{Z}_n \right\}.$$

Traditional notation $\{x\} := x \bmod 1$, i.e., map the points to the unit cube $[0,1)^s$.



- Strong algebraic structure.
- Unordered set of points.
  Typical use: apply rule with n points, if result not satisfactory, throw away and apply new rule with $\sim 2n$ points.
  $\rightarrow$ **Solution: Lattice sequences.**

Components of $\boldsymbol{g}$ are "well chosen" integers from the set

$$\mathbb{Z}_n^\times := \{v \in \mathbb{Z}_n : \gcd(v,n) = 1\} = U_n.$$

Polynomial lattice rules

Point set of a scalar rank-1 rule:

$$\left\{ \frac{k \boldsymbol{g} \bmod n}{n} : k \in \mathbb{Z}_n \right\}.$$

Now change everything to polynomials over $\mathbb{F}_q[\mathsf{x}]$:

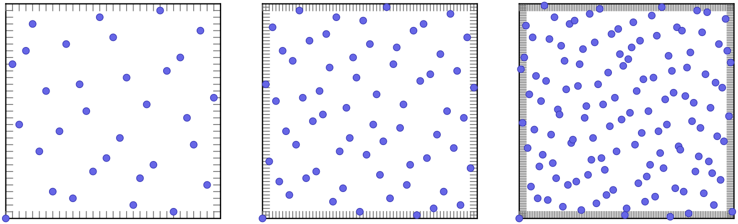
$$\left\{ \frac{k(\mathsf{x})\, \boldsymbol{g}(\mathsf{x}) \bmod f(\mathsf{x})}{f(\mathsf{x})} : k(\mathsf{x}) \in \mathbb{F}_q[\mathsf{x}]/f \right\}.$$

This is a lattice of polynomials in the "unit cube" modulo $f(\mathsf{x})$ with "points"

$$w(\mathsf{x}) = \sum_{i=1}^{\infty} w_i\, \mathsf{x}^{-i} \in \mathbb{F}_q((\mathsf{x}^{-1})).$$

Use a mapping function $[\cdot] : \mathbb{F}_q((\mathsf{x}^{-1})) \to [0,1)^s$ to form a *digital net*.

Polynomial lattice rules look like. . .



(g) $\mathbb{Z}_2/(x^5 + x^2 + 1)$ (h) $\mathbb{Z}_2/(x^6 + x + 1)$ (i) $\mathbb{Z}_2/(x^7 + x + 1)$

Parallel theory to normal lattice rules,
instead of Fourier space using a Walsh space.

The worst-case error for a lattice rule over a quotient ring R

The squared worst-case error for a lattice rule over a quotient ring R and an appropriately shift-invariant space with general weights can be written

$$e^2_{n,s}(\boldsymbol{g}, R) = \frac{1}{n} \sum_{k=0}^{n-1} \sum_{\emptyset \neq \mathfrak{u} \subseteq \mathcal{D}_s} \gamma_{\mathfrak{u}} \prod_{j \in \mathfrak{u}} \sum_{h=1}^{\infty} r(h)\, \psi_h(x_j^{(k)}),$$

where we normally take

R	$\mathbb{Z}$ modulo n	$\mathbb{F}_q[\mathsf{x}]$ modulo $f(\mathsf{x})$
ψ_h	Fourier-basis	Walsh-basis
n	n	$q^{\deg(f)}$
$\boldsymbol{x}^{(k)}$	$\dfrac{k\,\boldsymbol{g} \bmod n}{n}$	$\left[\dfrac{k(\mathsf{x})\,\boldsymbol{g}(\mathsf{x}) \bmod f(\mathsf{x})}{f(\mathsf{x})}\right]_M$
$\boldsymbol{g} \in G_s^\times$	$\boldsymbol{g} \in (\mathbb{Z}_n^\times)^s$	$\boldsymbol{g} \in ((\mathbb{F}_q[\mathsf{x}]/f)^\times)^s$

Why is Fourier- or Walsh-basis the natural choice?

- For lattice rules:

$$f(x) = \sum_{h=-\infty}^{\infty} \hat{f}(h) \exp(2\pi\mathrm{i}\, hx)$$

where $\chi_{[h]}$ is an additive character of $\mathbb{Z}_n$, $[h] = h \bmod n$.

- For a polynomial lattice rule over $\mathbb{F}_q[\mathsf{x}]$:

$$f\left(\left[\frac{k(\mathsf{x})g(\mathsf{x})}{f(\mathsf{x})}\right]_M\right) = \sum_{h=1}^{\infty} \hat{f}_q(h) \exp\left(\frac{2\pi\mathrm{i}}{q}[h(\mathsf{x})]_M \cdot \left[\frac{k(\mathsf{x})g(\mathsf{x})}{f(\mathsf{x})}\right]_M\right)$$

$$= \sum_{h=1}^{\infty} \hat{f}_q(h)\, \chi_{[h]}(k(\mathsf{x})g(\mathsf{x})/f(\mathsf{x}))$$

where for $M = \deg(f)$ (or better: the size of the additive group): $[h]_M = h \bmod q^M = (h_{M-1} \dots h_0)_q \simeq h_0 + h_1\mathsf{x} + \cdots + h_{M-1}\mathsf{x}^{M-1}$.

Why is Fourier- or Walsh-basis the natural choice?

- For lattice rules:

$$f\left(x^{(k)}\right) = \sum_{h=-\infty}^{\infty} \hat{f}(h) \exp\left(2\pi \mathrm{i}\, h x^{(k)}\right)$$

where $\chi_{[h]}$ is an additive character of $\mathbb{Z}_n$, $[h] = h \bmod n$.

- For a polynomial lattice rule over $\mathbb{F}_q[\mathsf{x}]$:

$$f\left(\left[\frac{k(\mathsf{x})g(\mathsf{x})}{f(\mathsf{x})}\right]_M\right) = \sum_{h=1}^{\infty} \hat{f}_q(h) \exp\left(\frac{2\pi \mathrm{i}}{q}[h(\mathsf{x})]_M \cdot \left[\frac{k(\mathsf{x})g(\mathsf{x})}{f(\mathsf{x})}\right]_M\right)$$

$$= \sum_{h=1}^{\infty} \hat{f}_q(h)\, \chi_{[h]}(k(\mathsf{x})g(\mathsf{x})/f(\mathsf{x}))$$

where for $M = \deg(f)$ (or better: the size of the additive group): $[h]_M = h \bmod q^M = (h_{M-1} \dots h_0)_q \simeq h_0 + h_1\mathsf{x} + \cdots + h_{M-1}\mathsf{x}^{M-1}$.

Why is Fourier- or Walsh-basis the natural choice?

- For lattice rules:

$$f\left(\frac{kg \bmod n}{n}\right) = \sum_{h=-\infty}^{\infty} \hat{f}(h) \exp\left(2\pi \mathrm{i}\, h \frac{kg}{n}\right)$$

where $\chi_{[h]}$ is an additive character of $\mathbb{Z}_n$, $[h] = h \bmod n$.

- For a polynomial lattice rule over $\mathbb{F}_q[\mathsf{x}]$:

$$f\left(\left[\frac{k(\mathsf{x})g(\mathsf{x})}{f(\mathsf{x})}\right]_M\right) = \sum_{h=1}^{\infty} \hat{f}_q(h) \exp\left(\frac{2\pi \mathrm{i}}{q}[h(\mathsf{x})]_M \cdot \left[\frac{k(\mathsf{x})g(\mathsf{x})}{f(\mathsf{x})}\right]_M\right)$$

$$= \sum_{h=1}^{\infty} \hat{f}_q(h)\, \chi_{[h]}(k(\mathsf{x})g(\mathsf{x})/f(\mathsf{x}))$$

where for $M = \deg(f)$ (or better: the size of the additive group):
$[h]_M = h \bmod q^M = (h_{M-1} \dots h_0)_q \simeq h_0 + h_1\mathsf{x} + \dots + h_{M-1}\mathsf{x}^{M-1}$.

Why is Fourier- or Walsh-basis the natural choice?

- For lattice rules:

$$f\left(\frac{kg \bmod n}{n}\right) = \sum_{h=-\infty}^{\infty} \hat{f}(h) \exp\left(2\pi \mathrm{i}\, h \frac{kg}{n}\right) = \sum_{h=-\infty}^{\infty} \hat{f}(h)\, \chi_{[h]}(kg/n)$$

where $\chi_{[h]}$ is an additive character of $\mathbb{Z}_n$, $[h] = h \bmod n$.

- For a polynomial lattice rule over $\mathbb{F}_q[\mathsf{x}]$:

$$f\left(\left[\frac{k(\mathsf{x})g(\mathsf{x})}{f(\mathsf{x})}\right]_M\right) = \sum_{h=1}^{\infty} \hat{f}_q(h) \exp\left(\frac{2\pi \mathrm{i}}{q} [h(\mathsf{x})]_M \cdot \left[\frac{k(\mathsf{x})g(\mathsf{x})}{f(\mathsf{x})}\right]_M\right)$$

$$= \sum_{h=1}^{\infty} \hat{f}_q(h)\, \chi_{[h]}(k(\mathsf{x})g(\mathsf{x})/f(\mathsf{x}))$$

where for $M = \deg(f)$ (or better: the size of the additive group): $[h]_M = h \bmod q^M = (h_{M-1} \ldots h_0)_q \simeq h_0 + h_1 \mathsf{x} + \cdots + h_{M-1} \mathsf{x}^{M-1}$.

Why is Fourier- or Walsh-basis the natural choice?

- For lattice rules:

$$f\left(\frac{kg \bmod n}{n}\right) = \sum_{h=-\infty}^{\infty} \hat{f}(h) \exp\left(2\pi \mathrm{i}\, h \frac{kg}{n}\right) = \sum_{h=-\infty}^{\infty} \hat{f}(h)\, \chi_{[h]}(kg/n)$$

where $\chi_{[h]}$ is an additive character of $\mathbb{Z}_n$, $[h] = h \bmod n$.

- For a polynomial lattice rule over $\mathbb{F}_q[\mathsf{x}]$:

$$f\left(\left[\frac{k(\mathsf{x})g(\mathsf{x})}{f(\mathsf{x})}\right]_M\right) = \sum_{h=1}^{\infty} \hat{f}_q(h) \exp\left(\frac{2\pi \mathrm{i}}{q} [h(\mathsf{x})]_M \cdot \left[\frac{k(\mathsf{x})g(\mathsf{x})}{f(\mathsf{x})}\right]_M\right)$$

$$= \sum_{h=1}^{\infty} \hat{f}_q(h)\, \chi_{[h]}(k(\mathsf{x})g(\mathsf{x})/f(\mathsf{x}))$$

where for $M = \deg(f)$ (or better: the size of the additive group):
$[h]_M = h \bmod q^M = (h_{M-1} \ldots h_0)_q \simeq h_0 + h_1 \mathsf{x} + \cdots + h_{M-1} \mathsf{x}^{M-1}$.

Character sums

So for $\chi(\cdot/f)$ an additive character of $\mathbb{F}_q[\mathsf{x}]/f$ and $k(\mathsf{x}) \in \mathbb{F}_q[\mathsf{x}]$

$$\sum_{g(\mathsf{x}) \in \mathbb{F}_q[\mathsf{x}]/f} \chi_{[h]}\left(\frac{k(\mathsf{x}) g(\mathsf{x})}{f(\mathsf{x})}\right) = \begin{cases} q^{\deg(f)} & \text{if } h(\mathsf{x})\, k(\mathsf{x}) \equiv 0 \pmod{f(\mathsf{x})}, \\ 0 & \text{otherwise,} \end{cases}$$

where $[\cdot] : \mathbb{Z} \to \mathbb{F}_q[\mathsf{x}]/f$ as $h = \sum_{i=0}^{\infty} h_i\, q^i \mapsto \sum_{i=0}^{m-1} h_i\, \mathsf{x}^i$.

Set $m := \deg(f)$, there are q^m characters. For $h \in \mathbb{Z}_{q^m}$, we have

$$\chi_h(w(\mathsf{x})) = \exp(2\pi \mathrm{i}\, (h_0 w_1 + h_1 w_2 + \cdots + h_{m-1} w_m)/q)$$

where $h = \sum_{i=0}^{m-1} h_i\, q^i$, and
$w(\mathsf{x}) = v(\mathsf{x})/f(\mathsf{x}) = \sum_{i=1}^{\infty} w_i\, \mathsf{x}^{-i}$, for $v(\mathsf{x}) \in \mathbb{F}_q[\mathsf{x}]/f$.

$\Rightarrow$ Trivial character is χ_0, i.e., $h \equiv 0 \pmod{q^m}$ for $h = 1, \ldots, \infty$.

Smoothness conditions $r(h)$ and their sum:

$$\mu := \sum_{h=1}^{\infty} r(h) = \begin{cases} 2\sum_{h=1}^{\infty} h^{-\alpha} & = 2\zeta(\alpha) & \text{for Korobov space,} \\ \sum_{h=1}^{\infty} q^{-\alpha\lfloor \log_q(h) \rfloor} & = \frac{(q-1)\, q^{\alpha}}{q^{\alpha}-q} & \text{for Walsh space.} \end{cases}$$

Now define

$$\mu_n := 2 \sum_{\substack{h \in \mathbb{Z}^* \\ h \equiv 0 \pmod{n}}} h^{-\alpha} = 2 \sum_{h \in \mathbb{Z}^*} (nh)^{-\alpha} = \mu / n^{\alpha},$$

$$\mu_{z(\mathsf{x})} := \sum_{\substack{h \in \mathbb{Z}^* \\ h(\mathsf{x}) \equiv 0 \pmod{z(\mathsf{x})}}} q^{-\alpha \deg(h)} \qquad (h \mapsto h(\mathsf{x}) = \textstyle\sum_{i=0}^{M-1} h_i\, \mathsf{x}^i \text{ but here } \deg(h) = \lfloor \log_q(h) \rfloor)$$

$$= \sum_{h \in \mathbb{Z}^*} q^{-\alpha(\deg(z)+\deg(h))}$$

$$= q^{-\alpha \deg(z)} \mu = \mu / n^{\alpha} = \mu_n = \mu_{q^{\deg(z)}}, \qquad n = |\mathbb{F}_q[\mathsf{x}]/z| = q^{\deg(z)}.$$

The mean worst-case error

For a general weighted space with $\gamma_\emptyset = \int_0^1 \omega(t)\,\mathrm{d}t = 1$:

$$
\begin{aligned}
M_{n,s}(R,\mathcal{G}) &:= \frac{1}{|\mathcal{G}|^s} \sum_{\boldsymbol{g}\in\mathcal{G}^s} e_{n,s}^2(\boldsymbol{g},R) \\
&= \frac{1}{|\mathcal{G}|^s} \sum_{\boldsymbol{g}\in\mathcal{G}^s} \frac{1}{n} \sum_{k=0}^{n-1} \sum_{\emptyset\neq\mathfrak{u}\subseteq\mathcal{D}_s} \gamma_{\mathfrak{u}} \prod_{j\in\mathfrak{u}} \omega(x_j^{(k)}) \\
&= \frac{1}{n} \sum_{k\in R} \sum_{\emptyset\neq\mathfrak{u}\subseteq\mathcal{D}_s} \gamma_{\mathfrak{u}} \prod_{j\in\mathfrak{u}} \underbrace{\sum_{h=1}^{\infty} r(h) \frac{1}{|\mathcal{G}|} \sum_{g\in\mathcal{G}} \chi_{[h]}(kg)}_{T(k,R,\mathcal{G})} \\
&= \frac{1}{n} \sum_{k\in R} \sum_{\emptyset\neq\mathfrak{u}\subseteq\mathcal{D}_s} \gamma_{\mathfrak{u}}\, T^{|\mathfrak{u}|}(k,R,\mathcal{G}).
\end{aligned}
$$

For convenience we add the subscript n, but here $n = |R|$.

We now have the following double character sum

$$T(k, R, \mathcal{G}) := \sum_{h=1}^{\infty} r(h) \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \chi_{[h]}(kg).$$

This is easy if $\mathcal{G} = R$ since then

$$\sum_{g \in R} \chi_{[h]}(kg) = \begin{cases} |R| & \text{if } k[h] \equiv 0 \text{ in } R, \\ 0 & \text{otherwise.} \end{cases}$$

Likewise, if $\mathcal{G} = R^* = R \setminus \{0\}$ then

$$\sum_{g \in R^*} \chi_{[h]}(kg) = \sum_{g \in R} \chi_{[h]}(kg) - 1 = \begin{cases} |R| - 1 & \text{if } k[h] \equiv 0 \text{ in } R, \\ -1 & \text{otherwise.} \end{cases}$$

We will defer the case $\mathcal{G} = R^\times$ (which uses the Möbius function).

Now using (mapping $h \in \mathbb{Z} \mapsto [h] \simeq h(\mathsf{x}) \in \mathbb{F}_q[\mathsf{x}]$, $\deg(h) < \deg(f)$)

$$\begin{aligned} & h(\mathsf{x})\,k(\mathsf{x}) \equiv 0 \pmod{f(\mathsf{x})} \\ \Leftrightarrow \quad & h(\mathsf{x}) \equiv 0 \pmod{z(\mathsf{x})}, \quad \text{where } z(\mathsf{x}) = f(\mathsf{x}) \,/\, \gcd(f(\mathsf{x}), k(\mathsf{x})), \end{aligned}$$

(which works for polynomials as well as integers), we get

$$\begin{aligned} T(k,R,R) = \sum_{\substack{h=1 \\ [h]k\equiv 0}}^{\infty} r(h) &= \sum_{\substack{h=1 \\ [h]\equiv 0 \pmod{z}}}^{\infty} r(h) \\ &= \mu_z, \qquad \text{with } z \text{ defined as above.} \end{aligned}$$

If R is a field, i.e., n prime or $f(\mathsf{x})$ irreducible, it follows that we have

$$\begin{aligned} T(k \neq 0, \mathbb{Z}_n, \mathbb{Z}_n) &= \mu_n &= \mu/n^\alpha \quad &\text{Korobov space,} \\ T(k \neq 0, \mathbb{F}_q[\mathsf{x}]/f, \mathbb{F}_q[\mathsf{x}]/f) &= \mu_{f(\mathsf{x})} &= \mu/n^\alpha \quad &\text{Walsh space.} \end{aligned}$$

For n and $f(\mathsf{x})$ arbitrary we get

$$M_{n,s}(\mathbb{Z}_n, \mathbb{Z}_n) = \frac{1}{n} \sum_{d|n} \sum_{k \in d\,\mathbb{Z}_{n/d}^\times} \sum_{\emptyset \neq \mathfrak{u} \subseteq \mathcal{D}_s} \gamma_{\mathfrak{u}}\, \mu_{n/d}^{|\mathfrak{u}|}$$

$$= \frac{1}{n} \sum_{d|n} \varphi(d) \sum_{\emptyset \neq \mathfrak{u} \subseteq \mathcal{D}_s} \gamma_{\mathfrak{u}} \left(\frac{\mu}{d^\alpha}\right)^{|\mathfrak{u}|},$$

and likewise

$$M_{n,s}(\mathbb{F}_q[\mathsf{x}]/f, \mathbb{F}_q[\mathsf{x}]/f) = \frac{1}{n} \sum_{\substack{z|f \\ L(z)=1}} \sum_{k \in z(\mathbb{F}_q[\mathsf{x}]/f)^\times} \sum_{\emptyset \neq \mathfrak{u} \subseteq \mathcal{D}_s} \gamma_{\mathfrak{u}}\, \mu_{f/z}^{|\mathfrak{u}|}$$

$$= \frac{1}{n} \sum_{\substack{z|f \\ L(z)=1}} \Phi_q(z) \sum_{\emptyset \neq \mathfrak{u} \subseteq \mathcal{D}_s} \gamma_{\mathfrak{u}} \left(\frac{\mu}{q^{\alpha \deg(z)}}\right)^{|\mathfrak{u}|}.$$

Note: if $f(\mathsf{x})$ is irreducible then only $\deg(f)$ comes into play, ...

If we use $\mathcal{G} = R^*$ and with $z(\mathsf{x}) = f(\mathsf{x}) \,/\, \gcd(f(\mathsf{x}), k(\mathsf{x}))$ or $z = n \,/\, \gcd(n, k)$, whichever is appropriate, we find:

$$
\begin{aligned}
T(k, R, R^*) &= \sum_{\substack{h=1 \\ [h] \equiv 0 \pmod{z}}}^{\infty} r(h) - \frac{1}{n-1} \sum_{\substack{h=1 \\ [h] \not\equiv 0 \pmod{z}}}^{\infty} r(h) \\
&= \mu_z - \frac{1}{n-1} \left(\sum_{h=1}^{\infty} r(h) - \sum_{\substack{h=1 \\ [h] \equiv 0 \pmod{z}}}^{\infty} r(h) \right) \\
&= \frac{n\,\mu_z - \mu}{n-1} \\
&= \frac{n\,\delta^{-\alpha} - 1}{n-1} \mu, \quad \delta \mid n \text{ and } \alpha > 1 \text{ thus mostly negative,}
\end{aligned}
$$

where $\delta = z$ for $R = \mathbb{Z}_n$ and $z \mid n \quad \Rightarrow \delta \mid n$;
or $\delta = q^{\deg(z)}$ for $R = \mathbb{F}_q[\mathsf{x}]/f$ and $z(\mathsf{x}) \mid f(\mathsf{x}) \quad \Rightarrow \delta \mid n$.

Now for $\mathcal{G} = R^*$ and for n and $f(\mathsf{x})$ arbitrary we get

$$M_{n,s}(\mathbb{Z}_n, \mathbb{Z}_n^*) = \frac{1}{n} \sum_{d|n} \varphi(d) \sum_{\emptyset \neq \mathfrak{u} \subseteq \mathcal{D}_s} \gamma_{\mathfrak{u}} \left(\frac{n\,\mu_d - \mu}{n-1} \right)^{|\mathfrak{u}|},$$

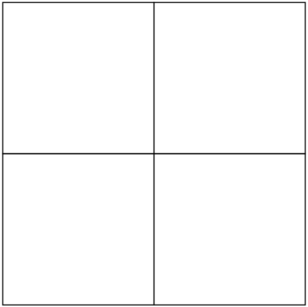
and likewise

$$M_{n,s}(\mathbb{F}_q[\mathsf{x}]/f, (\mathbb{F}_q[\mathsf{x}]/f)^*) = \frac{1}{n} \sum_{\substack{z|f \\ L(z)=1}} \Phi_q(z) \sum_{\emptyset \neq \mathfrak{u} \subseteq \mathcal{D}_s} \gamma_{\mathfrak{u}} \left(\frac{n\,\mu_z - \mu}{n-1} \right)^{|\mathfrak{u}|}.$$

Same note: if $f(\mathsf{x})$ is irreducible then only $\deg(f)$ comes into play.
Note: $\Phi_q(z)$ is easier to work with than $\varphi(d)$…

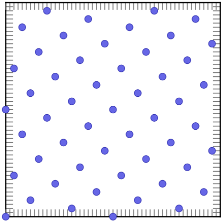
Copy rules

What is a copy rule?



Scale down the rule m^{-1} and "copy" m^s times.
$\Rightarrow$ Intractable by construction for s high.
Tractable approach: copy only in first r dimensions.

Classical definition

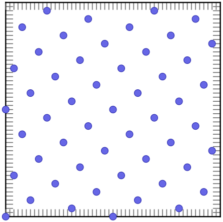


Given an n point rank-1 rule with generating vector $\boldsymbol{g}$ the m^r-copy rule is given by

$$Q(f) = \frac{1}{m^r n} \sum_{k \in \mathbb{Z}_n} \sum_{\boldsymbol{v} \in \mathbb{Z}_n^r} f\left(\left\{ \frac{\left\{ \frac{k \boldsymbol{g}}{n} \right\}}{m} + \frac{(v_1, \ldots, v_r, 0, \ldots, 0)}{m} \right\} \right).$$

Worst-case error can be calculated as that of the rank-1 rule.

Classical definition



Given an n point rank-1 rule with generating vector $\boldsymbol{g}$ the m^r-copy rule is given by (for $\gcd(n, m) = 1$)

$$Q(f) = \frac{1}{m^r n} \sum_{k \in \mathbb{Z}_n} \sum_{\boldsymbol{v} \in \mathbb{Z}_n^r} f\left(\left\{ \frac{k\,\boldsymbol{g}}{n} + \frac{(v_1, \ldots, v_r, 0, \ldots, 0)}{m} \right\} \right).$$

Worst-case error can be calculated as that of the rank-1 rule.

Polynomial copy rules

For $f(\mathsf{x}), c(\mathsf{x}) \in (\mathbb{F}_q[\mathsf{x}])^*$, $\gcd(f(\mathsf{x}), c(\mathsf{x})) = 1$, define

$$Q(f) = \frac{1}{m^r n} \sum_{k(\mathsf{x}) \in \mathbb{F}_q[\mathsf{x}]/f} \quad \sum_{\boldsymbol{v}(\mathsf{x}) \in (\mathbb{F}_q[\mathsf{x}]/c)^r} f\left(\left[\frac{k(\mathsf{x}) g(\mathsf{x})}{f(\mathsf{x})} + \frac{(v_1(\mathsf{x}), \ldots, v_r(\mathsf{x}), 0, \ldots, 0)}{c(\mathsf{x})} \right]_M \right).$$

Same equivalence as for scalars:

$$\left\{ \frac{k(\mathsf{x})}{f(\mathsf{x})\, c(\mathsf{x})} + \frac{v(\mathsf{x})}{c(\mathsf{x})} : k(\mathsf{x}) \in \mathbb{F}_q[\mathsf{x}]/f, v(\mathsf{x}) \in \mathbb{F}_q[\mathsf{x}]/c \right\}$$
$$= \left\{ \frac{k(\mathsf{x})}{f(\mathsf{x})} + \frac{v(\mathsf{x})}{c(\mathsf{x})} : k(\mathsf{x}) \in \mathbb{F}_q[\mathsf{x}]/f, v(\mathsf{x}) \in \mathbb{F}_q[\mathsf{x}]/c \right\}.$$

The additive group for arbitrary resolution M

For any $k(\mathsf{x}) \in \mathbb{F}_q[\mathsf{x}]/f$, thus $\deg(k) < \deg(f)$, do a "real" division

$$w(\mathsf{x}) = \frac{k(\mathsf{x})}{f(\mathsf{x})} = \sum_{i=1}^{\infty} w_i \, \mathsf{x}^{-i}.$$

The set of all such polynomial fractions L is an abelian group under polynomial addition.

Lemma

The set

$$L_M := \left\{ \left[\frac{k(\mathsf{x})}{f(\mathsf{x})} \right]_M : k \in \mathbb{F}_q[\mathsf{x}]/f \right\}$$

is a finite group under polynomial addition for any non-negative M, where $[\cdot]_M$ menas to truncate the Laurant series after x^{-M}. The order of this group is $q^{\min(M,m)}$ for $m = \deg(f)$. For $M = \infty$ there is no truncation and we set $L_\infty = L$.

The worst-case error for polynomial copy rules

With $N = q^{\deg(f)+|\mathcal{C}|\deg(c)} = nm^{|\mathcal{C}|}$, $M \geq \deg(f) + \deg(c)$, $\mathcal{C} \subseteq \mathcal{D}_s$:

$$e^2_{N,s}(\boldsymbol{g}, \mathbb{F}_q[\mathsf{x}]/f, \mathrm{copy}^{\mathcal{C}}_{c(\mathsf{x})}) = \frac{1}{n}\sum_{k=0}^{n-1}\frac{1}{m^{|\mathcal{C}|}}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_m^{\mathcal{C}}}\sum_{\emptyset\neq\mathfrak{u}\subseteq\mathcal{D}_s}\gamma_{\mathfrak{u}}\prod_{j\in\mathfrak{u}}\sum_{h=1}^{\infty} r(h)\,\chi_{[h]}(x_j^{(k,\boldsymbol{\ell})}).$$

Theorem (Cost as if plain rank-1 with n points instead of N)

The worst-case error when copying a polynomial rank-1 rule with points $\{\boldsymbol{x}^{(k)}(\mathsf{x})\}_{k=0}^{n-1}$ evaluated up to $M \geq \deg(f) + \deg(c)$ digits can also be calculated as a sum with only n terms instead of $N = nm^{|\mathcal{C}|}$ as

$$\frac{1}{n}\sum_{k=0}^{n-1}\sum_{\emptyset\neq\mathfrak{u}\subseteq\mathcal{D}_s}\gamma_{\mathfrak{u}}\left[\prod_{\substack{j\in\mathfrak{u}\\ j\in\mathcal{C}}}\sum_{\substack{h=1\\ [h]\equiv 0 \pmod{c}}}^{\infty} r(h)\,\chi_{[h]}\!\left(\frac{x_j^{(k)}}{c}\right)\right]\left[\prod_{\substack{j\in\mathfrak{u}\\ j\notin\mathcal{C}}}\sum_{h=1}^{\infty} r(h)\,\chi_{[h]}(x_j^{(k)})\right]$$

We use the following easy verifiable properties.
For χ_h an additive character of a group $L \subset \mathbb{F}_q((x^{-1}))$, $|L| = q^m$, $M \in \mathbb{Z}$ and $a, b \in \mathbb{F}_q((x))$:
the following is a trivial consequence of polynomial addition

$$[a + b]_M = [a]_M + [b]_M$$

the following is a property of characters

$$\chi_h(a + b) = \chi_h(a)\,\chi_h(b)$$

and for $M' \geq M$

$$\chi_h([a + b]_{M'}) = \chi_h(a + b)$$
$$\chi_h([a + b]_{M'}) = \chi_h([a]_{M'})\,\chi_h([b]_{M'})$$

Existence of good polynomial copy rules

For a copy rule with $N = nm^{|\mathcal{C}|}$ points, define

$$M_{N,s}(R, \mathcal{G}, \mathrm{copy}_c^{\mathcal{C}}) := \frac{1}{|\mathcal{G}|^s} \sum_{\boldsymbol{g} \in \mathcal{G}^s} e_{N,s}^2(\boldsymbol{g}, R, \mathrm{copy}_c^{\mathcal{C}}),$$

then, assuming $(n, m) = 1$ and $(f(\mathsf{x}), c(\mathsf{x})) = 1$ for simplicity,

$$M_{N,s}(\mathbb{Z}_n, \mathbb{Z}_n, \mathrm{copy}_m^{\mathcal{C}}) = \frac{1}{n} \sum_{d|n} \varphi(d) \sum_{\emptyset \neq \mathfrak{u} \subseteq \mathcal{D}_s} \gamma_{\mathfrak{u}} \, \mu_{md}^{|\mathfrak{u} \cap \mathcal{C}|} \mu_d^{|\mathfrak{u} \setminus \mathcal{C}|},$$

$$M_{N,s}(\mathbb{F}_q[\mathsf{x}]/f, \mathbb{F}_q[\mathsf{x}]/f, \mathrm{copy}_c^{\mathcal{C}}) = \frac{1}{n} \sum_{\substack{z|f \\ L(z)=1}} \Phi_q(z) \sum_{\emptyset \neq \mathfrak{u} \subseteq \mathcal{D}_s} \gamma_{\mathfrak{u}} \, \mu_{cz}^{|\mathfrak{u} \cap \mathcal{C}|} \mu_z^{|\mathfrak{u} \setminus \mathcal{C}|}.$$

Note: independent of the form of $c(\mathsf{x})$, only $\deg(c)$ matters.

The same for generating elements different from 0, i.e., from R^*.
This is now becoming a tedious trivial exercise.
Assuming $(n, m) = 1$ and $(f(\mathtt{x}), c(\mathtt{x})) = 1$ for simplicity,

$$M_{N,s}(\mathbb{Z}_n, \mathbb{Z}_n^*, \mathrm{copy}_m^{\mathcal{C}}) = \frac{1}{n} \sum_{d|n} \varphi(d) \sum_{\emptyset \neq \mathfrak{u} \subseteq \mathcal{D}_s} \gamma_{\mathfrak{u}} \left(\frac{n\, \mu_{md} - \mu_m}{n-1} \right)^{|\mathfrak{u} \cap \mathcal{C}|} \left(\frac{n\, \mu_d - \mu}{n-1} \right)^{|\mathfrak{u} \setminus \mathcal{C}|},$$

$$M_{N,s}(\mathbb{F}_q[\mathtt{x}]/f, (\mathbb{F}_q[\mathtt{x}]/f)^*, \mathrm{copy}_c^{\mathcal{C}}) = \frac{1}{n} \sum_{\substack{z|f \\ L(z)=1}} \Phi_q(z) \sum_{\emptyset \neq \mathfrak{u} \subseteq \mathcal{D}_s} \gamma_{\mathfrak{u}} \left(\frac{n\, \mu_{cz} - \mu_c}{n-1} \right)^{|\mathfrak{u} \cap \mathcal{C}|} \left(\frac{n\, \mu_z - \mu}{n-1} \right)^{|\mathfrak{u} \setminus \mathcal{C}|}.$$

To copy or not to copy

Similar to the classical copy rule analysis define (with $\gcd(f, c) = 1$)

$$\Omega := \frac{M_{N,s}(\mathbb{F}_q[\mathsf{x}]/f, (\mathbb{F}_q[\mathsf{x}]/f)^*, \mathrm{copy}^{\mathcal{C}}_{c(\mathsf{x})})}{M_{N,s}(\mathbb{F}_q[\mathsf{x}]/F, (\mathbb{F}_q[\mathsf{x}]/F)^*)}$$

$$= \frac{\displaystyle \frac{1}{n} \sum_{\substack{z|f \\ L(z)=1}} \Phi_q(z) \sum_{\emptyset \neq \mathfrak{u} \subseteq \mathcal{D}_s} \gamma_{\mathfrak{u}} \left(\frac{n\,\mu_{cz} - \mu_c}{n-1} \right)^{|\mathfrak{u} \cap \mathcal{C}|} \left(\frac{n\,\mu_z - \mu}{n-1} \right)^{|\mathfrak{u} \setminus \mathcal{C}|}}{\displaystyle \frac{1}{N} \sum_{\substack{Z|F \\ L(Z)=1}} \Phi_q(Z) \sum_{\emptyset \neq \mathfrak{u} \subseteq \mathcal{D}_s} \gamma_{\mathfrak{u}} \left(\frac{N\,\mu_Z - \mu}{N-1} \right)^{|\mathfrak{u}|}},$$

where $N = q^{\deg(f) + |\mathcal{C}| \deg(c)}$ and we take a $F(\mathsf{x}) \in \mathbb{F}_q[\mathsf{x}]$ with $\deg(F) = \deg(f) + |\mathcal{C}| \deg(c)$. Note: we can take F irreducible!

The question: when is $\Omega \leq 1$?

In the classical case only n can be taken prime, here we can take both f and F as irreducible polynomials over $\mathbb{F}_q$ and simplify the expressions for their average while still keeping the correct formula.

We then obtain:

$$\Omega = \frac{\sum_{\emptyset \neq \mathfrak{u} \subseteq \mathcal{D}_s} \gamma_{\mathfrak{u}} \frac{1}{n} \left[\mu_c^{|\mathfrak{u} \cap \mathcal{C}|} \mu^{|\mathfrak{u} \setminus \mathcal{C}|} + (n-1) \left(\frac{n\,\mu_{cf} - \mu_c}{n-1} \right)^{|\mathfrak{u} \cap \mathcal{C}|} \left(\frac{n\,\mu_f - \mu}{n-1} \right)^{|\mathfrak{u} \setminus \mathcal{C}|} \right]}{\sum_{\emptyset \neq \mathfrak{u} \subseteq \mathcal{D}_s} \gamma_{\mathfrak{u}} \frac{1}{N} \left[\mu^{|\mathfrak{u}|} + (N-1) \left(\frac{N\,\mu_F - \mu}{N-1} \right)^{|\mathfrak{u}|} \right]},$$

where for $m = q^{\deg(c)}$, $n = q^{\deg(f)}$ and $N = q^{\deg(f) + |\mathcal{C}| \deg(c)}$:

$$\mu_c = \frac{\mu}{m^\alpha}, \qquad \mu_f = \frac{\mu}{n^\alpha},$$
$$\mu_{cf} = \frac{\mu}{(mn)^\alpha}, \qquad \mu_F = \frac{\mu}{(m^{|\mathcal{C}|} n)^\alpha}.$$

Since $f(\mathsf{x})$ and $F(\mathsf{x})$ are both irreducible, the expression only depends on their degrees.

Equal weights and full copy rule

Setting all $\gamma_{\mathfrak{u}} = 1$ and copying in all dimensions $\mathcal{C} = \mathcal{D}_s$ we get the familiar form from Sloan & Joe:

$$\Omega = \frac{\left(m + \frac{\mu}{m^{\alpha-1}}\right)^s + m^s(n-1)\left(1 + \frac{n^{1-\alpha}-1}{m^{\alpha}(n-1)}\mu\right)^s - m^s n}{(1+\mu)^s + (m^s n - 1)\left(1 + \frac{(m^s n)^{1-\alpha}-1}{m^s n - 1}\mu\right)^s - m^s n}$$
$$= \frac{t_1 + t_2 - c}{b_1 + b_2 - c}$$

but here the expression is more honest since $F(\mathsf{x})$ is irreducible.

The classical proof is trying to find conditions such that

$$\frac{t_1 + t_2 - c}{b_1 + b_2 - c} < \frac{t_1}{b_1}.$$

Now

$$\frac{t_1 + t_2 - c}{b_1 + b_2 - c} < \frac{t_1}{b_1} \quad \Leftrightarrow \quad b_1(1 - t_2/c) - t_1(1 - b_2/c) > 0,$$

since $b_1 > 0$, $c > 0$ and $b_1 + b_2 - c > 0$ (which is always true since F is taken irreducible).

Now some analysis is needed, which we skip (since the conditions aren't that nice). We show some better results in pictures later.

First see what happens if indeed t_1/b_1 is the determining quantity for the ratio.

How many times do we copy?

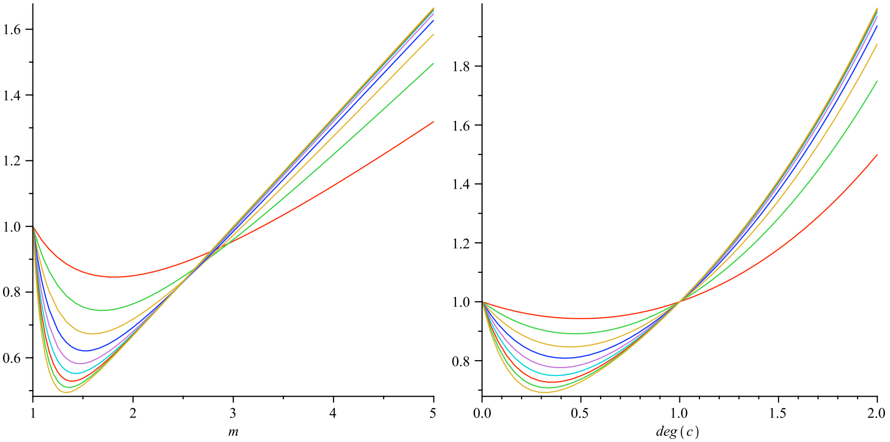
So, if

$$\Omega := \frac{M_{N,s}(\mathbb{F}_q[\mathsf{x}]/f, (\mathbb{F}_q[\mathsf{x}]/f)^*, \mathrm{copy}^{\mathcal{C}}_{c(\mathsf{x})})}{M_{N,s}(\mathbb{F}_q[\mathsf{x}]/F, (\mathbb{F}_q[\mathsf{x}]/F)^*)} < \frac{t_1}{b_1} = \left(\frac{m + m^{1-\alpha}\mu}{1+\mu}\right)^s = \rho^s$$

then what value of $m = q^{\deg(c)}$ should we pick?
(Note that μ is a function of q and α.)

We will first plot the function $\rho(\deg(c), q, \alpha)$ to get an idea to when it is smaller than 1. $\rightarrow$ In such a case then copying certainly works.

Classical copying versus polynomial copying



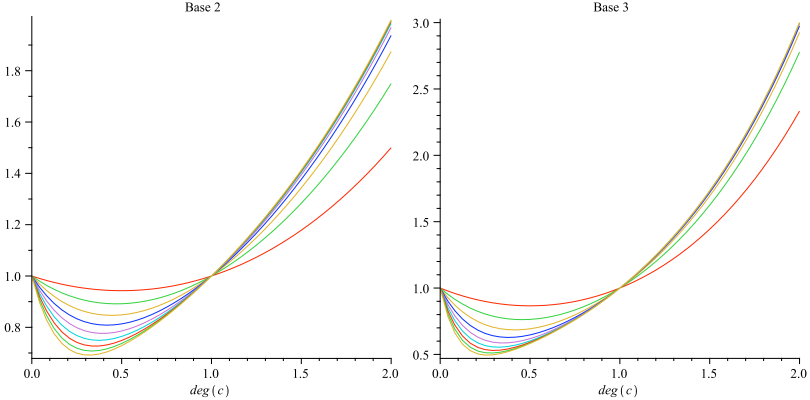
Left: classical copying in Korobov space: $m = 2$ and $m = 3$ are okay.
Right: polynomial copying in Walsh space ($q = 2$): only $m = 2^1$ is okay!

Polynomial copying

In fact, for a copying polynomial with degree 1:

$$\rho(1, q, \alpha) = 1,$$

independent of α or q; after that it only increases. → One choice!

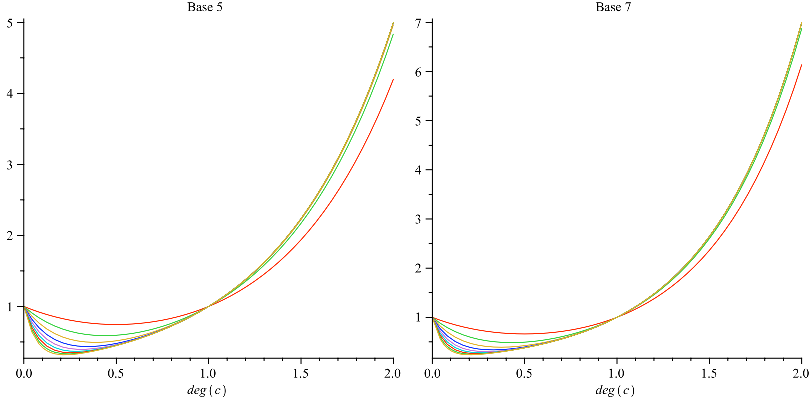


Polynomial copying

In fact, for a copying polynomial with degree 1:

$$\rho(1, q, \alpha) = 1,$$

independent of α or q; after that it only increases. $\rightarrow$ One choice!

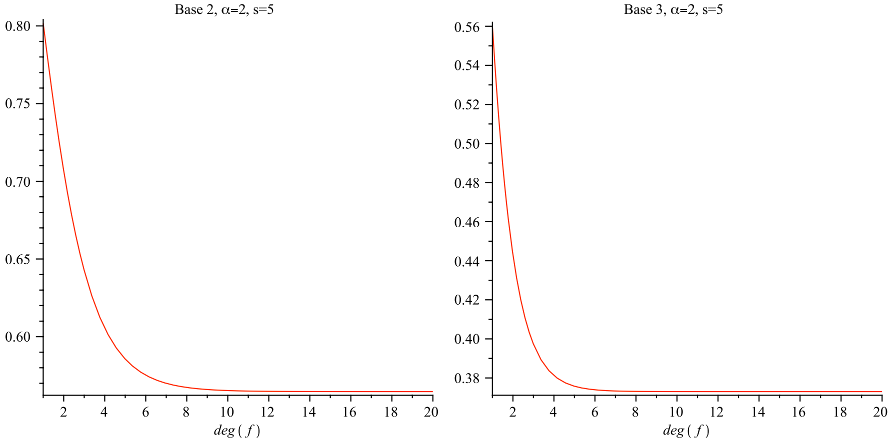


Is this a negative result?

The preceding graphs suggest that only copy polynomials of degree 1 should be considered, and, in that case we might even not do better at all (since $\rho = 1$).

Let's forget the $\rho^s = t_1/b_1$ part for a while and look at the ratio Ω directly in some plots… The situation is much better than it seems.

An example in base 2 (left) and base 3 (right) in 5 dimensions, $\alpha = 2$:



$\rightarrow$ Copying with a degree 1 copy polynomial can result in a dramatic decrease of the mean squared worst-case error compared to a rank-1 rule with the same amount of points! However: $\Omega \rightarrow 1$ as $s \rightarrow \infty$.

Using product weights

To tackle the problem of $s \to \infty$ we introduce product weights:

$$\Omega = \frac{t_1 + t_2 - c}{b_1 + b_2 - c}$$

$$t_1 = \prod_{j=1}^{r} \left(m + \gamma_j \frac{\mu}{m^{\alpha-1}} \right) \prod_{j=r+1}^{s} (1 + \gamma_j \, \mu)$$

$$b_1 = \prod_{j=1}^{s} (1 + \gamma_j \, \mu)$$

$$t_2 = m^r (n-1) \prod_{j=1}^{r} \left(1 + \gamma_j \frac{n^{1-\alpha} - 1}{m^\alpha (n-1)} \mu \right) \prod_{j=1}^{r} \left(1 + \gamma_j \frac{n^{1-\alpha} - 1}{n-1} \right)$$

$$b_2 = (m^r n - 1) \prod_{j=1}^{s} \left(1 + \gamma_j \frac{(m^r n)^{1-\alpha} - 1}{m^r n - 1} \right)$$

Using product weights

To tackle the problem of $s \to \infty$ we introduce product weights:

$$\Omega = \frac{t_1 + t_2 - c}{b_1 + b_2 - c}$$

$$t_1 = \prod_{j=1}^{r} \left(m + \gamma_j \frac{\mu}{m^{\alpha-1}}\right) \prod_{j=r+1}^{s} (1 + \gamma_j \, \mu)$$

$$b_1 = \prod_{j=1}^{r} (1 + \gamma_j \, \mu) \prod_{j=r+1}^{s} (1 + \gamma_j \, \mu)$$

$$t_2 = m^r (n-1) \prod_{j=1}^{r} \left(1 + \gamma_j \frac{n^{1-\alpha} - 1}{m^\alpha (n-1)} \mu\right) \prod_{j=1}^{r} \left(1 + \gamma_j \frac{n^{1-\alpha} - 1}{n-1}\right)$$

$$b_2 = (m^r n - 1) \prod_{j=1}^{s} \left(1 + \gamma_j \frac{(m^r n)^{1-\alpha} - 1}{m^r n - 1}\right)$$

Using the weights to get $t_1/b_1 < 1$

We again look at the ratio of the first terms

$$\frac{t_1}{b_1} = \frac{\prod_{j=1}^{r} \left(m + \gamma_j \frac{\mu}{m^{\alpha-1}}\right)}{\prod_{j=1}^{r} (1 + \gamma_j \, \mu)}$$

and ask this to be < 1.
A sufficient condition is to have each term < 1 for $j = 1, \ldots, r$:

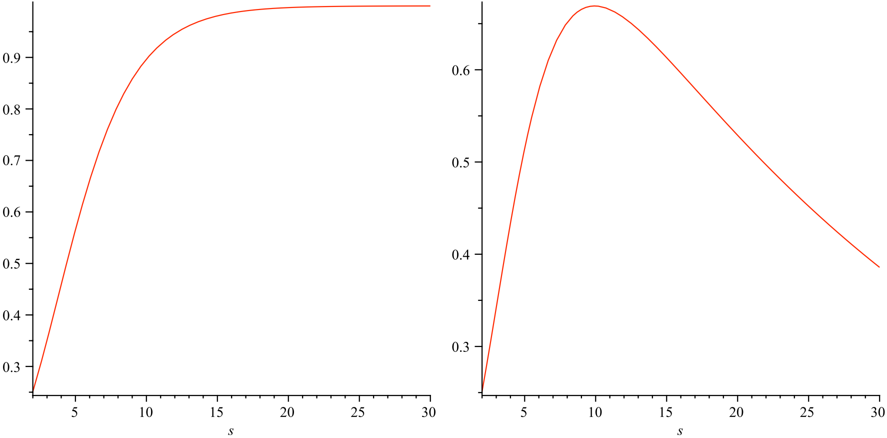
$$\frac{m + \gamma_j \frac{\mu}{m^{\alpha-1}}}{1 + \gamma_j \, \mu} < 1$$

inspired by the unweighted case we set $\deg(c) = 1$

$$\frac{q(q^\alpha - q) + q(q-1)\gamma_j}{q^\alpha - q + (q^{\alpha+1} - q^\alpha)\gamma_j} < 1$$

$$\Leftrightarrow \qquad \gamma_j > 1.$$

E.g., for $q = 2$, $\alpha = 2$, $\deg(c) = 1$, $\deg(f) = 10$,
$\gamma_j = 1$ (left) and $\gamma_j = 1.1$ (right):



With the correct weights everything is fine.
(Kuo & Joe have conditions on product weights for classical rules.)

When to copy for product weights in Walsh space

For product weights copying of a polynomial lattice rules over $\mathbb{F}_q[\mathtt{x}]$ works if:

- $\deg(c) = 1$
- $\gamma_j > 1$ for the copying dimensions
- $s > 2$ (?)

(To have tractability the weights have to satisfy the usual conditions.)

Construction of polynomial copy rules

The fast component-by-component construction algorithm can be used to construct polynomial copy rules, for reasonable weights. (I.e., not 2^s different weights.)

For the standard product weighted Walsh space we can use the same implementation of the algorithm as for a normal rank-1 rule by adjusting the weights:

$$\tilde{\gamma}_j = \frac{\gamma_j}{m^\alpha}$$

and multiplying the generating components by the multiplicative inverse of $c(\mathsf{x})$ modulo $f(\mathsf{x})$ ($\gcd(c,f) = 1$)

$$z_j(\mathsf{x}) \equiv c(\mathsf{x})^{-1}\,\tilde{z}_j(\mathsf{x}) \pmod{f(\mathsf{x})}.$$

Conclusion

- Both for classical lattice rules and polynomial lattice rules: copying works if $\gamma\mu$ is large enough.
- For both of them the shift-invariant Sobolev space has a kernel which is much smaller than that of the Korobov or Walsh space. I.e., a factor $2\pi^2$ and 12, and so copying does not work there!
- If copying works, what does it tell us about the function space???

- Not surprisingly: fast construction for polynomial copy rules works in exactly the same way as for classical lattice rules.