On the Step-by-Step Construction of Polynomial Lattices for Numerical Integration in Weighted Sobolev Spaces

Peter Kritzer
(joint work with F. Pillichshammer)

Department of Financial Mathematics
Johannes Kepler University Linz, Austria
peter.kritzer@jku.at

Eighth IMACS Seminar on Monte Carlo Methods
Borovets, August 2011

Outline

Polynomial Lattice Points

In many applications (e.g., financial engineering): need to numerically approximate the value of an integral,

$$I_s(F) := \int_{[0,1]^s} F(\boldsymbol{x}) d\boldsymbol{x},$$

$F : [0, 1]^s \to \mathbb{R}$.

In many applications (e.g., financial engineering): need to numerically approximate the value of an integral,

$$I_s(F) := \int_{[0,1]^s} F(\boldsymbol{x}) d\boldsymbol{x},$$

$F : [0,1]^s \to \mathbb{R}$.

Approximate $I_s(F)$ by

$$Q_{N,s}(F) := \frac{1}{N} \sum_{n=0}^{N-1} F(\boldsymbol{x}_n),$$

where $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}$ are points in $[0,1)^s$.

How should we choose the points?

How should we choose the points?

In this talk:

How should we choose the points?

In this talk:

Polynomial lattice points (polynomial analogue of lattice points).

Let $\mathbb{F}_p((x^{-1}))$ be the field of formal Laurent series over $\mathbb{F}_p$ (p a prime), with elements of the form

$$L = \sum_{l=w}^{\infty} t_l x^{-l},$$

$w \in \mathbb{Z}$ and all $t_l \in \mathbb{F}_p$.

Let $\mathbb{F}_p((x^{-1}))$ be the field of formal Laurent series over $\mathbb{F}_p$ (p a prime), with elements of the form

$$L = \sum_{l=w}^{\infty} t_l x^{-l},$$

$w \in \mathbb{Z}$ and all $t_l \in \mathbb{F}_p$.

For $m \in \mathbb{N}$, define $\nu_m : \mathbb{F}_p((x^{-1})) \to [0, 1)$,

$$\nu_m\left(\sum_{l=w}^{\infty} t_l x^{-l}\right) = \sum_{l=\max(1,w)}^{m} t_l p^{-l}.$$

For $0 \le n < p^m$ let $n = n_0 + n_1 p + \cdots + n_{m-1} p^{m-1}$ be the p-adic expansion of n. With each such n we associate the polynomial

$$n(x) = \sum_{l=0}^{m-1} n_l x^l \in \mathbb{F}_p[x].$$

Definition

Choose a polynomial $f \in \mathbb{F}_p[x]$, $\deg f = m$,

Definition

Choose a polynomial $f \in \mathbb{F}_p[x]$, $\deg f = m$, *and* $\boldsymbol{g} = (g_1, \ldots, g_s) \in (\mathbb{F}_p[x])^s$.

Definition

Choose a polynomial $f \in \mathbb{F}_p[x]$, $\deg f = m$, *and* $\boldsymbol{g} = (g_1, \ldots, g_s) \in (\mathbb{F}_p[x])^s$.
For each $n \in \{0, 1, \ldots, p^m - 1\}$, *construct a point*

$$\boldsymbol{x}_n := \nu_m\left(\frac{n(x)\boldsymbol{g}(x)}{f(x)}\right) = \left(\nu_m\left(\frac{n(x)g_1(x)}{f(x)}\right), \ldots, \nu_m\left(\frac{n(x)g_s(x)}{f(x)}\right)\right).$$

Definition

Choose a polynomial $f \in \mathbb{F}_p[x]$, $\deg f = m$, *and* $\boldsymbol{g} = (g_1, \ldots, g_s) \in (\mathbb{F}_p[x])^s$.
For each $n \in \{0, 1, \ldots, p^m - 1\}$, *construct a point*

$$\boldsymbol{x}_n := \nu_m\left(\frac{n(x)\boldsymbol{g}(x)}{f(x)}\right) = \left(\nu_m\left(\frac{n(x)g_1(x)}{f(x)}\right), \ldots, \nu_m\left(\frac{n(x)g_s(x)}{f(x)}\right)\right).$$

This yields a point set $\mathcal{P}(\boldsymbol{g}, f)$ *of* $N = p^m$ *points, which is called polynomial lattice point set.*

Definition

Choose a polynomial $f \in \mathbb{F}_p[x]$, $\deg f = m$, *and* $\boldsymbol{g} = (g_1, \ldots, g_s) \in (\mathbb{F}_p[x])^s$.
For each $n \in \{0, 1, \ldots, p^m - 1\}$, *construct a point*

$$\boldsymbol{x}_n := \nu_m\left(\frac{n(x)\boldsymbol{g}(x)}{f(x)}\right) = \left(\nu_m\left(\frac{n(x)g_1(x)}{f(x)}\right), \ldots, \nu_m\left(\frac{n(x)g_s(x)}{f(x)}\right)\right).$$

This yields a point set $\mathcal{P}(\boldsymbol{g}, f)$ *of* $N = p^m$ *points, which is called polynomial lattice point set. A QMC rule using* $\mathcal{P}(\boldsymbol{g}, f)$ *as node set is called a polynomial lattice rule.*

We call

- f the *modulus*,
- $\boldsymbol{g}$ the *generating vector* of $\mathcal{P}(\boldsymbol{g}, f)$.

Polynomial lattice point sets $\subseteq$

Polynomial lattice point sets $\subseteq$

$\subseteq$ Hyperplane nets (Dick, Pillichshammer, Pirsic) $\subseteq$

Polynomial lattice point sets $\subseteq$

$\subseteq$ Hyperplane nets (Dick, Pillichshammer, Pirsic) $\subseteq$

$\subseteq$ (t, m, s)-nets (Faure, Sobol', Niederreiter)

Polynomial Lattice Rules for Numerical Integration in Weighted Sobolev Spaces

Consider the problem of approximating

$$I_s(F) := \int_{[0,1]^s} F(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x},$$

for functions F in a reproducing kernel Hilbert space $\mathcal{H}$.

The spaces we study here are weighted Sobolev spaces.

The spaces we study here are weighted Sobolev spaces.

Unanchored Sobolev space: Kernel function

$$K(\boldsymbol{x}, \boldsymbol{y}) := \prod_{j=1}^{s} \left(1 + \gamma_j \left(\frac{1}{2} B_2 \left(\{x_j - y_j\}\right) + \left(x_j - \tfrac{1}{2}\right)\left(y_j - \tfrac{1}{2}\right)\right)\right),$$

where B_2 is the second Bernoulli polynomial and $\boldsymbol{\gamma} = (\gamma_j)_{j \geq 1}$ is a sequence of weights.

The spaces we study here are weighted Sobolev spaces.

Unanchored Sobolev space: Kernel function

$$K(\boldsymbol{x}, \boldsymbol{y}) \quad := \quad \prod_{j=1}^{s} \left(1 + \gamma_j \left(\frac{1}{2} B_2 \left(\{x_j - y_j\}\right) + \left(x_j - \tfrac{1}{2}\right)\left(y_j - \tfrac{1}{2}\right)\right)\right),$$

where B_2 is the second Bernoulli polynomial and $\boldsymbol{\gamma} = (\gamma_j)_{j \geq 1}$ is a sequence of weights.

Here we assume that $1 \geq \gamma_1 \geq \gamma_2 \geq \cdots > 0$.

The spaces we study here are weighted Sobolev spaces.

Unanchored Sobolev space: Kernel function

$$K(\boldsymbol{x},\boldsymbol{y}) := \prod_{j=1}^{s}\left(1+\gamma_j\left(\frac{1}{2}B_2\left(\{x_j-y_j\}\right)+\left(x_j-\tfrac{1}{2}\right)\left(y_j-\tfrac{1}{2}\right)\right)\right),$$

where B_2 is the second Bernoulli polynomial and $\boldsymbol{\gamma}=(\gamma_j)_{j\geq 1}$ is a sequence of weights.

Here we assume that $1\geq\gamma_1\geq\gamma_2\geq\cdots>0$.

Idea of weighting introduced by Sloan and Woźniakowski.

Anchored Sobolev space: Kernel function

$$\widetilde{K}(\boldsymbol{x},\boldsymbol{y}) := \prod_{j=1}^{s} \left(1+\gamma_j \min(1-x_j, 1-y_j)\right),$$

with weights as above.

One frequently considers the worst case integration error,

$$e(\mathcal{P}(\boldsymbol{g}, f), \mathcal{H}) := \sup_{\substack{F \in \mathcal{H} \\ \|F\| \leq 1}} \left| I_s(F) - Q_{N,s}(F) \right| ,$$

Digital shifts: convenient way of randomizing polynomial lattices.

Digital shifts: convenient way of randomizing polynomial lattices.
For each component $j \in \{1, 2, \ldots, s\}$, choose

$$\boldsymbol{\sigma}^{(j)} = \sum_{k=1}^{\infty} \frac{\sigma_k^{(j)}}{p^k}.$$

Digital shifts: convenient way of randomizing polynomial lattices.
For each component $j \in \{1, 2, \ldots, s\}$, choose

$$\boldsymbol{\sigma}^{(j)} = \sum_{k=1}^{\infty} \frac{\sigma_k^{(j)}}{p^k}.$$

Modify j-th component of point $\boldsymbol{x}_n \in \mathcal{P}(\boldsymbol{g}, f)$,

$$x_n^{(j)} = \sum_{k=1}^{\infty} \frac{x_{n,k}^{(j)}}{p^k}$$

Digital shifts: convenient way of randomizing polynomial lattices.
For each component $j \in \{1, 2, \ldots, s\}$, choose

$$\boldsymbol{\sigma}^{(j)} = \sum_{k=1}^{\infty} \frac{\sigma_k^{(j)}}{p^k}.$$

Modify j-th component of point $\boldsymbol{x}_n \in \mathcal{P}(\boldsymbol{g}, f)$,

$$x_n^{(j)} = \sum_{k=1}^{\infty} \frac{x_{n,k}^{(j)}}{p^k}$$

into

$$x_n^{(j)} \oplus \boldsymbol{\sigma}^{(j)} := \sum_{k=1}^{\infty} \frac{x_{n,k}^{(j)} \oplus \sigma_k^{(j)}}{p^k}.$$

One sometimes studies the *mean square* worst case integration error of polynomial lattices $\mathcal{P}(\boldsymbol{g}, f)$, defined by

$$\widehat{e}^2(\mathcal{P}(\boldsymbol{g}, f), \mathcal{H}) := \int_{[0,1]^s} e^2(\mathcal{P}(\boldsymbol{g}, f) \oplus \boldsymbol{\sigma}, \mathcal{H}) \, \mathrm{d}\boldsymbol{\sigma},$$

i.e., the expectation of the squared worst case integration error with respect to a randomly chosen digital shift.

Dick and Pillichshammer (2005) showed the following result for the unanchored Sobolev space.

Dick and Pillichshammer (2005) showed the following result for the unanchored Sobolev space.

Theorem

For any irreducible polynomial $f \in \mathbb{F}_p[x]$ *with* $\deg(f) = m$ *one can construct, component by component, a generating vector* $\boldsymbol{g}$ *such that*

$$\widehat{e}^2(\mathcal{P}(\boldsymbol{g}, f), \mathcal{H}) \leq c_{s,p,\gamma,\lambda} p^{-m/\lambda}$$

for any $\lambda \in (1/2, 1]$, *with an explicitly known positive constant* $c_{s,p,\gamma,\lambda}$.

Dick and Pillichshammer (2005) showed the following result for the unanchored Sobolev space.

Theorem

For any irreducible polynomial $f \in \mathbb{F}_p[x]$ with $\deg(f) = m$ one can construct, component by component, a generating vector $\boldsymbol{g}$ such that

$$\widehat{e}^2(\mathcal{P}(\boldsymbol{g}, f), \mathcal{H}) \leq c_{s,p,\gamma,\lambda} p^{-m/\lambda}$$

for any $\lambda \in (1/2, 1]$, with an explicitly known positive constant $c_{s,p,\gamma,\lambda}$.

Similar result for anchored Sobolev space: Dick, Kuo, Pillichshammer, Sloan, 2005.

Dick and Pillichshammer (2005) showed the following result for the unanchored Sobolev space.

Theorem

For any irreducible polynomial $f \in \mathbb{F}_p[x]$ *with* $\deg(f) = m$ *one can construct, component by component, a generating vector* $\boldsymbol{g}$ *such that*

$$\widehat{e}^2(\mathcal{P}(\boldsymbol{g}, f), \mathcal{H}) \leq c_{s,p,\gamma,\lambda} p^{-m/\lambda}$$

for any $\lambda \in (1/2, 1]$*, with an explicitly known positive constant* $c_{s,p,\gamma,\lambda}$*.*

Similar result for anchored Sobolev space: Dick, Kuo, Pillichshammer, Sloan, 2005.
Similar result for general f: Pillichshammer, K., 2007.

These results are interesting, but ...

These results are interesting, but ...

... they are only valid for the mean-square worst case error with respect to all possible digital shifts.

These results are interesting, but . . .

. . . they are only valid for the mean-square worst case error with respect to all possible digital shifts.

Explicit construction of good digital shifts remained an open problem in the above-mentioned papers.

These results are interesting, but . . .

. . . they are only valid for the mean-square worst case error with respect to all possible digital shifts.

Explicit construction of good digital shifts remained an open problem in the above-mentioned papers.

Can we make the result “more explicit”?

Theorem (Pillichshammer, K., 2011)

Let $\mathcal{H}$ be one of the spaces defined above. Let $m, s \in \mathbb{N}$ be given, and set $N := p^m$. Let $f \in \mathbb{F}_p[x]$ (p a prime) be irreducible with $\deg(f) = m$. Then we can construct a polynomial lattice rule $\mathcal{P}(\boldsymbol{g}, f)$ and a vector $\boldsymbol{\sigma}$ such that

$$e^2\left(\mathcal{P}(\boldsymbol{g}, f) \oplus \boldsymbol{\sigma}, \mathcal{H}\right) \leq \frac{1}{N} \prod_{j=1}^{s} \left(1 + \gamma_j c(\mathcal{H})\right),$$

where $c(\mathcal{H}) \leq (p+1)/3$.

- **Proof:** Component by component construction.

- **Proof:** Component by component construction.
- Squared worst case errors in both spaces can be conveniently expressed in terms of the respective kernel functions.

- **Proof:** Component by component construction.
- Squared worst case errors in both spaces can be conveniently expressed in terms of the respective kernel functions.
- Construct components of $\boldsymbol{g}$ and of $\boldsymbol{\sigma}$ in parallel:

- **Proof:** Component by component construction.
- Squared worst case errors in both spaces can be conveniently expressed in terms of the respective kernel functions.
- Construct components of $\boldsymbol{g}$ and of $\boldsymbol{\sigma}$ in parallel:
- Construct one component g_d of $\boldsymbol{g}$.

- **Proof:** Component by component construction.
- Squared worst case errors in both spaces can be conveniently expressed in terms of the respective kernel functions.
- Construct components of $\boldsymbol{g}$ and of $\boldsymbol{\sigma}$ in parallel:
- Construct one component g_d of $\boldsymbol{g}$.
- Then construct corresponding component $\sigma^{(d)}$ of $\boldsymbol{\sigma}$.

- **Proof:** Component by component construction.
- Squared worst case errors in both spaces can be conveniently expressed in terms of the respective kernel functions.
- Construct components of $\boldsymbol{g}$ and of $\boldsymbol{\sigma}$ in parallel:
- Construct one component g_d of $\boldsymbol{g}$.
- Then construct corresponding component $\sigma^{(d)}$ of $\boldsymbol{\sigma}$.
- Increase d by 1 and repeat.

- Similar to what was done for lattice points by Sloan, Kuo, Joe (2002).

- Similar to what was done for lattice points by Sloan, Kuo, Joe (2002).
- **Trade-off:** Weaker bound on the worst case error, but explicit construction.

Define the initial error of multivariate integration (i.e., error without sampling a function) in $\mathcal{H}$ by

$$e_{0,s}(\mathcal{H}) := \sup_{\substack{F \in \mathcal{H} \\ \|F\| \leq 1}} |I_s(F)| .$$

Define the initial error of multivariate integration (i.e., error without sampling a function) in $\mathcal{H}$ by

$$e_{0,s}(\mathcal{H}) := \sup_{\substack{F \in \mathcal{H} \\ \|F\| \leq 1}} |I_s(F)| .$$

In both spaces, $e_{0,s}(\mathcal{H}) = 1$.

Task: Reduce $e_{0,s}(\mathcal{H})$ by a factor of $\varepsilon \in (0,1)$.

Task: Reduce $e_{0,s}(\mathcal{H})$ by a factor of $\varepsilon \in (0,1)$.

Define

$$N_{\min}(\varepsilon, s, \mathcal{H}) := \min\{N \in \mathbb{N} : \ \exists \ P_{N,s} : \ e(P_{N,s}, \mathcal{H}) \leq \varepsilon\}.$$

Task: Reduce $e_{0,s}(\mathcal{H})$ by a factor of $\varepsilon \in (0,1)$.

Define

$$N_{\min}(\varepsilon, s, \mathcal{H}) := \min\{N \in \mathbb{N} : \exists\, P_{N,s} : \; e(P_{N,s}, \mathcal{H}) \le \varepsilon\}.$$

Integration in $\mathcal{H}$ is (polynomially) QMC-tractable if there exist non-negative integers c, r, q:

$$N_{\min}(\varepsilon, s, \mathcal{H}) \le c s^q \varepsilon^{-r}$$

holds for all $s \in \mathbb{N}$ and all $\varepsilon \in (0,1)$.

Task: Reduce $e_{0,s}(\mathcal{H})$ by a factor of $\varepsilon \in (0,1)$.

Define

$$N_{\min}(\varepsilon, s, \mathcal{H}) := \min\{N \in \mathbb{N} : \exists\, P_{N,s} : \, e(P_{N,s}, \mathcal{H}) \leq \varepsilon\}.$$

Integration in $\mathcal{H}$ is (polynomially) QMC-tractable if there exist non-negative integers c, r, q:

$$N_{\min}(\varepsilon, s, \mathcal{H}) \leq c s^q \varepsilon^{-r}$$

holds for all $s \in \mathbb{N}$ and all $\varepsilon \in (0,1)$.

Integration in $\mathcal{H}$ is strongly (polynomially) QMC-tractable, if the above inequality holds with $q = 0$.

(Strong) tractability if weights decay sufficiently fast.

(Strong) tractability if weights decay sufficiently fast.

- Sufficient for (polynomial) tractability:

$$\limsup_{s\to\infty} \frac{\sum_{j=1}^{s} \gamma_j}{s} < \infty.$$

- Sufficient for strong (polynomial) tractability:

$$\sum_{j=1}^{\infty} \gamma_j < \infty.$$

Conclusion

- We can use polynomial lattice point sets for integration in weighted Sobolev spaces.

- We can use polynomial lattice point sets for integration in weighted Sobolev spaces.
- Explicit constructions of digitally shifted polynomial lattices yielding good results has remained an open problem before.

- We can use polynomial lattice point sets for integration in weighted Sobolev spaces.
- Explicit constructions of digitally shifted polynomial lattices yielding good results has remained an open problem before.
- We now have explicit constructions, but we have to pay a price for this feature (worse rate of convergence).

Thank you.