Supplement to "Multivariate frequency polygon for stationary random fields"

by

Michel Carbon and Thierry Duchesne*

Department of Mathematics and Statistics

Université Laval

7 septembre 2023

*Corresponding author: thierry.duchesne@mat.ulaval.ca

1 Properties of the coefficients $c_{j_1,\cdots,j_d}(x)$

We now demonstrate some of the properties of the $c_{j_1,\cdots,j_d}(x)$ coefficients.

1.1 Sum of the coefficients : $\sum_{j_1,\cdots,j_d\in\{0,1\}^d} c_{j_1,\cdots,j_d}(x) = 1$

Clearly, the coefficients $c_{j_1,\cdots,j_d}(x) \in [0,1]$. Let $Y_1, \cdots, Y_d$ be d independent random variables such that every Y_s is $\mathcal{B}(1, u_s(x))$ distributed, for $s = 1, \ldots, d$; recall that $u_s(x) \in [0,1]$. Under this assumption, for $j_1, \cdots, j_d \in \{0,1\}^d$, the joint probability of $Y_1, \ldots, Y_d$ is

$$P\{(Y_1, \ldots, Y_d) = (j_1, \ldots, j_d)\} = \prod_{s=1}^{d} P(Y_s = j_s) = c_{j_1,\cdots,j_d}(x).$$

Then, immediately

$$\sum_{j_1,\cdots,j_d\in\{0,1\}^d} c_{j_1,\cdots,j_d}(x) = 1. \qquad (1)$$

An immediate consequence of this property is that $\hat{f}_{\mathbf{n}}(x)$ is a density in $\mathbb{R}^d$.

1.2 For all fixed $s = 1, \ldots, d$, $\sum_{j_1,\cdots,j_d\in\{0,1\}^d} (-1)^{j_s+1} c_{j_1,\cdots,j_d}(x) = 2u_s - 1 = 2\,\frac{x_s - a_s}{b_s}$

There is no loss of generality in taking $s = 1$. Then the summation is

$$\sum_{j_1,\cdots,j_d\in\{0,1\}^d} (-1)^{j_1+1} c_{j_1,\cdots,j_d}(x) = -\sum_{j_2,\cdots,j_d\in\{0,1\}^{d-1}} c_{0,j_2,\cdots,j_d}(x) + \sum_{j_2\cdots,j_d\in\{0,1\}^{d-1}} c_{1,j_2,\cdots,j_d}(x).$$

Using the definition of $c_{j_1,\cdots,j_d}(x) = \prod_{s=1}^{d} u_s^{j_s}(1-u_s)^{1-j_s}$, we obtain

$$\begin{aligned}
\sum_{j_1,\cdots,j_d\in\{0,1\}^d} (-1)^{j_1+1} c_{j_1,\cdots,j_d}(x) &= -(1-u_1) \sum_{j_2,\cdots,j_d\in\{0,1\}^{d-1}} c_{j_2,\cdots,j_d}(x) + u_1 \sum_{j_2\cdots,j_d\in\{0,1\}^{d-1}} c_{j_2,\cdots,j_d}(x) \\
&= (2u_1 - 1) \sum_{j_2\cdots,j_d\in\{0,1\}^{d-1}} c_{j_2,\cdots,j_d}(x) = 2u_1 - 1 = 2\,\frac{x_1 - a_1}{b_1},
\end{aligned}$$

by using (1).

1.3 For all fixed s and t, $\sum_{j_1,\cdots,j_d\in\{0,1\}^d}(-1)^{j_s+j_t}c_{j_1,\cdots,j_d}(x)=(2u_s-1)(2u_t-1)=4\,\frac{x_s-a_s}{b_s}\frac{x_t-a_t}{b_t}$

There is no loss of generality in taking $s=1$ and $t=2$. Then, letting $j=j_1,\ldots,j_d$, the summation is

$$\begin{aligned}\sum_{j_1,\cdots,j_d\in\{0,1\}^d}(-1)^{j_1+j_2}c_j(x) &= \sum_{j_3,\cdots,j_d\in\{0,1\}^{d-2}}(-1)^{0+0}c_{0,0,j_3,\cdots,j_d}(x)+\sum_{j_3,\cdots,j_d\in\{0,1\}^{d-2}}(-1)^{0+1}c_{0,1,j_3,\cdots,j_d}(x)\\ &+ \sum_{j_3,\cdots,j_d\in\{0,1\}^{d-2}}(-1)^{1+0}c_{1,0,j_3,\cdots,j_d}(x)+\sum_{j_3,\cdots,j_d\in\{0,1\}^{d-2}}(-1)^{1+1}c_{1,1,j_3,\cdots,j_d}(x)\\ &= \sum_{j_3,\cdots,j_d\in\{0,1\}^{d-2}}(1-u_1)(1-u_2)c_{j_3,\cdots,j_d}(x)-\sum_{j_3,\cdots,j_d\in\{0,1\}^{d-2}}(1-u_1)u_2c_{j_3,\cdots,j_d}(x)\\ &- \sum_{j_3,\cdots,j_d\in\{0,1\}^{d-2}}u_1(1-u_2)c_{j_3,\cdots,j_d}(x)+\sum_{j_3,\cdots,j_d\in\{0,1\}^{d-2}}u_1u_2c_{j_3,\cdots,j_d}(x)\end{aligned}$$

Note that

$$(1-u_1)(1-u_2)-(1-u_1)u_2-u_1(1-u_2)+u_1u_2=(2u_1-1)(2u_2-1).$$

Finally, using (1), we obtain

$$\sum_{j_1,\cdots,j_d\in\{0,1\}^d}(-1)^{j_1+j_2}c_{j_1,\cdots,j_d}(x)=(2u_1-1)(2u_2-1)=4\,\frac{x_1-a_1}{b_1}\frac{x_2-a_2}{b_2}.$$

1.4 $\sum_{j_1,\cdots,j_d\in\{0,1\}^d}\{c_{j_1,\cdots,j_d}(x)\}^2=\prod_{i=1}^{d}\left[(1-u_i)^2+u_i^2\right]$

Let us proceed by induction. When $d=1$ we get directly

$$\sum_{j_1\in\{0,1\}}\{c_{j_1}(x)\}^2=(1-u_1)^2+u_1^2.$$

Now let us suppose that the formula is verified for a fixed $d \in \mathbb{N}^* - \{1\}$. Using the definition of $c_{j_1,\cdots,j_d}(x)$, we have

$$\begin{aligned}
\sum_{j_1,\cdots,j_{d+1}\in\{0,1\}^{d+1}} \left\{c_{j_1,\cdots,j_{d+1}}(x)\right\}^2 &= \sum_{j_1,\cdots,j_{d+1}\in\{0,1\}^{d+1}} \prod_{s=1}^{d} \left\{u_s^{j_s}(1-u_s)^{1-j_s}\right\}^2 \left\{u_{d+1}^{j_{d+1}}(1-u_{d+1})^{1-j_{d+1}}\right\}^2 \\
&= \sum_{j_1,\cdots,j_d\in\{0,1\}^d} \left\{c_{j_1,\cdots,j_d}(x)\right\}^2 (1-u_{d+1})^2 + \sum_{j_1,\cdots,j_d\in\{0,1\}^d} \left\{c_{j_1,\cdots,j_d}(x)\right\}^2 u_{d+1}^2 \\
&= \sum_{j_1,\cdots,j_d\in\{0,1\}^d} \left\{c_{j_1,\cdots,j_d}(x)\right\}^2 \left\{(1-u_{d+1})^2 + u_{d+1}^2\right\} \\
&= \prod_{i=1}^{d+1} \left[(1-u_i)^2 + u_i^2\right],
\end{aligned}$$

which completes the proof.

1.5 For all fixed s, $\sum_{j_1,\cdots,j_d\in\{0,1\}^d} (-1)^{j_s+1} \left\{c_{j_1,\cdots,j_d}(x)\right\}^2 =$

$\left[u_s^2 - (1-u_s)^2\right] \prod_{i=1, i\neq s}^{d} \left[u_i^2 + (1-u_i)^2\right]$

We can take $s = 1$ without loss of generality. So we easily have

$$\sum_{j_1,\cdots,j_d\in\{0,1\}^d} (-1)^{j_s+1} \left\{c_{j_1,\cdots,j_d}(x)\right\}^2$$

$$= -(1-u_1)^2 \sum_{j_2,\cdots,j_d\in\{0,1\}^{d-1}} \left(c_{j_2,\cdots,j_d}(x)\right)^2 + u_1^2 \sum_{j_2,\cdots,j_d\in\{0,1\}^{d-1}} \left(c_{j_2,\cdots,j_d}(x)\right)^2$$

$$= \left[u_1^2 - (1-u_1)^2\right] \prod_{i=2}^{d} \left[u_i^2 + (1-u_i)^2\right].$$

2 Bias of the LBFP $\hat{f}_{\mathbf{n}}(x)$.

Using a Taylor-Lagrange expansion around the point $a = x_k + \frac{1}{2}b$, there exists ξ_x between a and x such that

$$\begin{aligned} f(x) &= f(a) + \sum_{s=1}^{d} f'_s(a)(x_s - a_s) + \frac{1}{2}\sum_{s,t=1}^{d} f''_{s,t}(a)(x_s - a_s)(x_t - a_t) \\ &\quad + \frac{1}{6}\sum_{s,t,m=1}^{d} f'''_{s,t,m}(\xi_x)(x_s - a_s)(x_t - a_t)(x_m - a_m). \end{aligned} \tag{2}$$

By integrating (2) we get

$$\begin{aligned} \int_{I(j,k)} f(x)\,dx &= b_1\cdots b_d f(a) + \sum_{s=1}^{d} f'_s(a)\int_{I(j,k)}(x_s - a_s)\,dx + \frac{1}{2}\sum_{s,t=1}^{d} f''_{s,t}(a)\int_{I(j,k)}(x_s - a_s)(x_t - a_t)\,dx \\ &\quad + \frac{1}{6}\sum_{s,t,m=1}^{d}\int_{I(j,k)} f'''_{s,t,m}(\xi_x)(x_s - a_s)(x_t - a_t)(x_m - a_m)\,dx. \end{aligned} \tag{3}$$

Let us first calculate

$$\frac{1}{b_1\cdots b_d}\int_{I(j,k)}(x_s - a_s)\,dx = \frac{1}{b_s}\int_{a_s+(j_s-1)b_s}^{a_s+j_s b_s}(x_s - a_s)\,dx_s = \frac{1}{2}b_s(-1+2j_s) = \frac{1}{2}b_s(-1)^{j_s+1}. \tag{4}$$

Then

$$\frac{1}{b_1\cdots b_d}\int_{I(j,k)}(x_s - a_s)^2\,dx = \frac{1}{b_s}\int_{a_s+(j_s-1)b_s}^{a_s+j_s b_s}(x_s - a_s)^2\,dx_s = \frac{b_s^2}{3}(3j_s^2 - 3j_s + 1) = \frac{b_s^2}{3}, \tag{5}$$

and, using (4), we obtain

$$\frac{1}{b_1\cdots b_d}\int_{I(j,k)}(x_s-a_s)(x_t-a_t)\,dx = \frac{1}{b_s b_t}\int_{a_s+(j_s-1)b_s}^{a_s+j_s b_s}(x_s-a_s)\,dx_s\int_{a_t+(j_t-1)b_t}^{a_t+j_t b_t}(x_t-a_t)\,dx_t = \frac{1}{4}b_s b_t(-1)^{j_s+j_t}. \tag{6}$$

For the third term in Equation (3), we will use Lemma 1 below which is a multivariate version of what Scott (1985) calls the generalized mean-value theorem.

Lemma 1 *Let f be a non negative and continuous function defined on a cell $[a,b] = \prod_{s=1}^{d}[a_s, b_s]$. If ϕ is an other continuous function on that same cell, then*

$$\int_{[a,b]} \phi(x) f(x)\,dx = \phi(\tilde{x})\int_{[a,b]} f(x)\,dx \tag{7}$$

for some $\tilde{x} \in [a,b]$.

Proof. First of all, (7) is trivially true if $\int_{[a,b]} f(x)\,dx = 0$. So, let us assume that this last integral is not zero and let us set $f_0(x) = \dfrac{f(x)}{\int_{[a,b]} f(x)\,dx}$. Of course, f_0 is a density on $[a,b]$ and $E[\phi(X)] = \int_{[a,b]} \phi(x) f_0(x)\,dx$.

But ϕ carries the convex set $\prod_{s=1}^{d}[a_s, b_s]$ onto an interval, say $[c,d]$, and $E[\phi(X)]$ must be somewhere in that interval. So there exists $\tilde{x} \in [a,b]$ such that (7) is verified. $\square$

Using Lemma 1, the third term in the right member of Equation (3) is

$$\frac{1}{6}\sum_{s,t,m=1}^{d}\int_{I(j,k)} f'''_{s,t,m}(\xi_x)(x_s-a_s)(x_t-a_t)(x_m-a_m)\,dx = \frac{1}{6}\sum_{s,t,m=1}^{d} f'''_{s,t,m;j}(\xi_{s,t,m;j})\int_{[(j-1)b,jb]} x_s x_t x_m\,dx \tag{8}$$

where $\xi_{s,t,m;j}$ is in $[a+(j-1)b, a+jb]$.

From (2), (3), (4), (5), (6), and (8), we obtain

$$\begin{aligned} E\hat{f}_{\mathbf{n}}(x) &= f(a) + \sum_{s=1}^{d} f'_s(a)(x_s-a_s) + \sum_{s=1}^{d} f''_{s,s}(a)\frac{1}{6}b_s^2 + \sum_{\substack{s,t=1\\ s\neq t}}^{d} f''_{s,t}(a)\frac{1}{2}(x_s-a_s)(x_t-a_t) \\ &\quad + \frac{1}{6}\sum_{s,t,m=1}^{d} f'''_{s,t,m}(\xi_{s,t,m;j})\int_{[(j-1)b,jb]} x_s x_t x_m\,dx. \end{aligned} \tag{9}$$

It follows that

$$bias(x) = E\hat{f}_{\mathbf{n}}(x) - f(x) = \sum_{s=1}^{d} f''_{s,s}(a)\left\{\frac{1}{6}b_s^2 - \frac{1}{2}(x_s-a_s)^2\right\} + \varepsilon(x) = b(x) + \varepsilon(x), \tag{10}$$

where

$$\begin{aligned} \varepsilon(x) &= \frac{1}{b_1\cdots b_d}\sum_{j_1,\cdots,j_d\in\{0,1\}^d} c_{j_1,\cdots,j_d}(x)\frac{1}{6}\sum_{s,t,m=1}^{d} f'''_{s,t,m}(\xi_{s,t,m;j})\int_{[(j-1)b,jb]} x_s x_t x_m\,dx \\ &\quad - \frac{1}{6}\sum_{s,t,m=1}^{d} f'''_{s,t,m}(\xi_x)(x_s-a_s)(x_t-a_t)(x_m-a_m). \end{aligned} \tag{11}$$

We want to calculate $\int_{\mathbb{R}^d} \{b(x)\}^2 \, dx$. We first evaluate

$$\int_{I(k)} \{b(x)\}^2 \, dx = \int_{I(k)} \left[\sum_{s=1}^{d} f''_{s,s}(a) \left\{ \frac{1}{6} b_s^2 - \frac{1}{2}(x_s - a_s)^2 \right\} \right]^2 dx \tag{12}$$

A few elementary calculations give us

$$\begin{aligned} \int_{I(k)} \{b(x)\}^2 \, dx &= \sum_{s=1}^{d} \left\{ f''_{s,s}(a) \right\}^2 \int_{I(k)} \left\{ \frac{1}{6} b_s^2 - \frac{1}{2}(x_s - a_s)^2 \right\}^2 dx \\ &\quad + \sum_{\substack{s,t=1 \\ s \neq t}}^{d} f''_{s,s}(a) f''_{t,t}(a) \int_{I(k)} \left\{ \frac{1}{6} b_s^2 - \frac{1}{2}(x_s - a_s)^2 \right\} \left\{ \frac{1}{6} b_t^2 - \frac{1}{2}(x_t - a_t)^2 \right\} dx \\ &= \left(\sum_{s=1}^{d} \frac{49}{2880} b_s^4 \left\{ f''_{s,s}(a) \right\}^2 + \sum_{\substack{s,t=1 \\ s \neq t}}^{d} \frac{1}{64} b_s^2 b_t^2 f''_{s,s}(a) f''_{t,t}(a) \right) b_1 \cdots b_d. \end{aligned} \tag{13}$$

We'll need the following lemma, which is a multidimensional generalization of results reached by Friedman and Diaconis (1981).

Lemma 2 *Suppose that $\phi : \mathbb{R}^d \to \mathbb{R}$ is continuous and its first order partial derivatives $\phi'_1, \cdots, \phi'_d$ are continuous and integrable. Then*

$$\sum_k \phi(\xi_k) b_1 \cdots b_d = \int_{\mathbb{R}^d} \phi(x) \, dx + O\left(\sum_{s=1}^{d} b_s \int_{\mathbb{R}^d} |\phi'_s(x)| \, dx \right), \tag{14}$$

where the sum is of over all cells, the union of which is $\mathbb{R}^d$, each cell has volume $b_1 \cdots b_d$ and ξ_k is an arbitrary point in cell number k.

An explicit bound is available if the mixed higher order derivatives

$$\phi^{(d)}_{1,\cdots,d}(x) = \frac{\partial^d}{\partial x_1 \cdots \partial x_d} \phi(x)$$

exist, and all the functions $\phi''_{i,j}(x) = \dfrac{\partial^2}{\partial x_i \partial x_j} \phi(x)$ for $i < j$,

$$\phi'''_{i,j}(x) = \frac{\partial^3}{\partial x_i \partial x_j \partial x_k} \phi(x), \quad \text{for } i < j < k \text{ and } \phi^{(d)}_{1,\cdots,d}(x) = \frac{\partial^d}{\partial x_1 \cdots \partial x_d} \phi(x)$$

exist and are square integrable. We have

$$\left|\sum_k \phi(\xi_k) b_1 \cdots b_d - \int_{\mathbb{R}^d} \phi(x)\,dx\right| \le \sum_{i=1}^d b_i \int_{\mathbb{R}^d} \left|\phi_1'(x)\right|\,dx + \sum_{1\le i<j\le d} b_i b_j \int_{\mathbb{R}^d} \left|\phi_{i,j}''(x)\right|\,dx$$

$$+ \sum_{1\le i<j<k\le d} b_i b_j b_k \int_{\mathbb{R}^d} \left|\phi_{i,j,k}'''(x)\right|\,dx + \cdots + b_1\cdots b_d \int_{\mathbb{R}^d} \left|\phi_{1,\cdots,d}^{(d)}(x)\right|\,dx \quad (15)$$

Proof. We first consider a cell $I = \prod_{s=1}^d [a_s, c_s]$ with $c_s = a_s + b_s$ and an arbitrary point $\xi = (\xi_1, \cdots, \xi_d) \in I$. We have

$$\begin{aligned} \phi(x) - \phi(\xi) &= \sum_{s=1}^d \{\phi(\xi_1,\cdots,\xi_{s-1},x_s,x_{s+1},\cdots,x_d) - \phi(\xi_1,\cdots,\xi_{s-1},\xi_s,x_{s+1},\cdots,x_d)\} \\ &= \sum_{s=1}^d \int_{\xi_s}^{x_s} \frac{\partial}{\partial x_s}\phi(\xi_1,\cdots,\xi_{s-1},y_s,x_{s+1},\cdots,x_d)\,dy_s. \end{aligned} \quad (16)$$

We then deduce immediately that

$$|\phi(x) - \phi(\xi)| \le \sum_{s=1}^d \int_{\xi_s}^{x_s} \left|\frac{\partial}{\partial x_s}\phi(\xi_1,\cdots,\xi_{s-1},y_s,x_{s+1},\cdots,x_d)\right|\,dy_s. \quad (17)$$

By Fubini's theorem, we obtain

$$\begin{aligned} \left|\int_I \phi(x)\,dx - \phi(\xi) b_1\cdots b_d\right| &= \left|\int_I [\phi(x) - \phi(\xi)]\,dx\right| \le \int_I |\phi(x) - g(\xi)|\,dx \\ &\le \sum_{s=1}^d \int_I \int_{a_s}^{c_s} \left|\frac{\partial}{\partial x_s}\phi(\xi_1,\cdots,\xi_{s-1},y_s,x_{s+1},\cdots,x_d)\right|\,dy_s\,dx \\ &= \sum_{s=1}^d (c_s - a_s) \int_I \left|\frac{\partial}{\partial x_s}\phi(\xi_1,\cdots,\xi_{s-1},x_s,x_{s+1},\cdots,x_d)\right|\,dx. \end{aligned} \quad (18)$$

We can use this bound for each cell :

$$\begin{aligned}\left|\int_I \phi(x)\,dx - \sum_k \phi(\xi_k) b_1 \cdots b_d\right| &\leq \sum_k \left|\int_{I_k} [\phi(x) - \phi(\xi_k)]\ dx\right| \\ &\leq b_1 \int_{I_k} \left|\frac{\partial}{\partial x_1}\phi(x)\,dx\right| + b_2 \sum_k \int_{I_k} \left|\frac{\partial}{\partial x_2}\phi(\xi_{k,1}, x_2, \cdots, x_d)\right|\, dx \\ &\quad + \cdots + b_d \sum_k \int_{I_k} \left|\frac{\partial}{\partial x_d}\phi(\xi_{k,1}, x_2, \cdots, x_{k,d-1}, x_d)\right|\, dx.\end{aligned} \tag{19}$$

We can note that the integrability of $\left|\frac{\partial}{\partial x_i}\phi(x)\right|$ was necessary in the previous development. Thus we have demonstrated the first part, i.e. equation (14).

Remembering the first part of the proof, we can proceed to the following algebraic decomposition. There are $3^d - 1$ terms in the right side of (20) and it is needed to be shown repeatedly and thoroughly :

$$\begin{aligned}\phi(x) \;-\; \phi(\xi) &= \sum_{s=1}^{d} \int_{\xi_s}^{x_s} \frac{\partial}{\partial x_s}\phi(x_1, \cdots, x_{s-1}, y_s, x_{s+1}, \cdots, x_d)\, dy_s \\ &- \sum_{1=s<t=d} \int_{\xi_t}^{x_t} \int_{\xi_s}^{x_s} \frac{\partial^2}{\partial x_s \partial x_t}\phi(x_1, \cdots, y_s, \cdots, y_t, \cdots, x_d)\, dy_s\, dy_t \\ &+ \sum_{1=s<t<m=d} \int_{\xi_m}^{x_m} \int_{\xi_t}^{x_t} \int_{\xi_s}^{x_s} \frac{\partial^3}{\partial x_s \partial x_t \partial x_m}\phi(x_1, \cdots, y_s, \cdots, y_t, \cdots, y_m, \cdots, x_d)\, dy_s\, dy_t\, dy_k \\ &+ \cdots + (-1)^{d+1} \int_{\xi_d}^{x_d} \cdots \int_{\xi_1}^{x_1} \frac{\partial^d}{\partial x_1 \cdots \partial x_d}\phi(y_1, \cdots, y_d)\, dy_1 \cdots dy_d.\end{aligned} \tag{20}$$

Now, consider first the particular cell I. Then we have

$$\left|\int_I \{\phi(x) - \quad \phi(\xi)\}\, dx\right| \le \sum_{s=1}^{d} \int_I \int_{a_s}^{c_s} \left|\frac{\partial}{\partial x_s}\phi(x_1, \cdots, x_{s-1}, y_s, x_{s+1}, \cdots, x_d)\right| dy_s\, dx$$

$$+ \sum_{1=s<t=d} \int_I \int_{a_t}^{c_t} \int_{a_s}^{c_s} \left|\frac{\partial^2}{\partial x_s \partial x_t}\phi(x_1, \cdots, y_s, \cdots, y_t, \cdots, x_d)\right| dy_s\, dy_t\, dx$$

$$+ \cdots + \int_I \int_{a_d}^{c_d} \cdots \int_{a_1}^{c_1} \left|\frac{\partial^d}{\partial x_1 \cdots \partial x_d}\phi(y_1, \cdots, y_d)\right| dy_1 \cdots dy_d\, dx$$

$$= \sum_{s=1}^{d} (c_s - a_s) \int_I \left|\frac{\partial}{\partial x_s}\phi(x)\right| dx + \sum_{1=s<t=d} (c_s - a_s)(c_t - a_t) \int_I \left|\frac{\partial^2}{\partial x_s \partial x_t}\phi(x)\right| dx$$

$$+ \cdots + (c_1 - a_1) \cdots (c_d - a_d) \int_I \left|\frac{\partial^d}{\partial x_1 \cdots \partial x_d}\phi(x)\right| dx. \tag{21}$$

By summing up over all cells, we finally obtain (15). □

From (13), using Lemma 2 and summing up over all cells, we obtain

$$\begin{aligned}\int_{\mathbb{R}^d} \{b(x)\}^2\, dx \;&=\; \sum_{s=1}^{d} \frac{49}{2880} b_s^4 \left[\int_{\mathbb{R}^d} \left\{f''_{s,s}(x)\right\}^2 dx + O\left(\sum_{s=1}^{d} b_s\right)\right] \\ &\quad + \sum_{1\le s<t\le d}^{d} \frac{1}{32} b_s^2 b_t^2 \left[\int_{\mathbb{R}^d} f''_{s,s}(x) f''_{t,t}(x)\, dx + O\left(\sum_{s=1}^{d} b_s\right)\right],\end{aligned} \tag{22}$$

or, equivalently,

$$\int_{\mathbb{R}^d} \{b(x)\}^2\, dx = \sum_{s=1}^{d} \frac{49}{2880} b_s^4 \int_{\mathbb{R}^d} \left\{f''_{s,s}(x)\right\}^2 dx + \sum_{1\le s<t\le d}^{d} \frac{1}{32} b_s^2 b_t^2 \int_{\mathbb{R}^d} f''_{s,s}(x) f''_{t,t}(x)\, dx + O\left(\sum_{s=1}^{d} b_s^5\right). \tag{23}$$

Now consider $\varepsilon(x) = \sum_{s,t,m=1}^{d} \varepsilon_{s,t,m}(x)$. Fixing j, from (11), we have

$$\varepsilon_{s,s,s}(x) = \frac{1}{6}\left\{\sum_{j_1,\cdots,j_d \in \{0,1\}^d} c_{j_1,\cdots,j_d}(x) f'''_{s,s,s}(\xi_{s,s,s;j}) \frac{1}{4}(-1)^{j_s+1} b_s^3 - f'''_{s,s,s}(\xi_x)(x_s - a_s)^3\right\}. \tag{24}$$

Using the inequality $(a-b)^2 \le 2(a^2+b^2)$ and integrating $\{\varepsilon_{s,s,s}(x)\}^2$ on the interval $I(k)$, we obtain

$$\int_{I(k)} \{\varepsilon_{s,s,s}(x)\}^2 \, dx \;\le\; \frac{2}{36}\left[\frac{1}{16} b_s^6 \int_{I(k)} \left(\sum_{j_1,\cdots,j_d \in \{0,1\}^d} c_{j_1,\cdots,j_d}(x) f'''_{s,s,s}(\xi_{s,s,s;j})(-1)^{j_s+1} \right)^2 dx \right.$$
$$\left. + \int_{I(k)} \left\{ f'''_{s,s,s}(\xi_x) \right\}^2 (x_s - a_s)^6 \, dx \right]. \tag{25}$$

Using lemma (1), we have

$$\int_{I(k)} \{\varepsilon_{s,s,s}(x)\}^2 \, dx \;\le\; \frac{1}{18\cdot 16}\left(\frac{1}{2}\right)^d b_s^6 \left\{ \sum_{j_1,\cdots,j_d \in \{0,1\}^d} f'''_{s,s,s}(\xi_{s,s,s;j}) \right\}^2 b_1 \cdots b_d \tag{26}$$
$$+ \frac{1}{18\cdot 7 \cdot 64} b_s^6 \left(f'''_{s,s,s}(\tilde{\xi}_x) \right)^2 b_1 \cdots b_d.$$

Finally, summing over all cells, we obtain

$$\int_{\mathbb{R}^d} \{\varepsilon_{s,s,s}(x)\}^2 \, dx = O\left(b_s^6\right). \tag{27}$$

We can do a similar analysis for all the other terms $\varepsilon_{s,t,m}(x)$, using the inequality

$$\{\varepsilon(x)\}^2 \le 2^{d^3} \sum_{s,t,m=1}^{d} \{\varepsilon_{s,t,m}(x)\}^2, \tag{28}$$

which will finally give

$$\int_{\mathbb{R}^d} (\varepsilon(x))^2 = O\left(\sum_{s=1}^{d} b_s^6\right). \tag{29}$$

Remembering that $\int_{\mathbb{R}^d} \{b(x)\}^2 \, dx = O\left(\sum_{s=1}^{d} b_s^4\right)$, by Schwarz's inequality we also have

$$\left| \int_{\mathbb{R}^d} b(x)\varepsilon(x) \, dx \right| = O\left(\sum_{s=1}^{d} b_s^5\right). \tag{30}$$

We finally obtain the behavior of the integrated squared bias :

$$\int_{\mathbb{R}^d} \{bias(x)\}^2 \, dx = \sum_{s=1}^{d} \frac{49}{2880} b_s^4 \int_{\mathbb{R}^d} \left\{ f''_{s,s}(x) \right\}^2 \, dx + \sum_{1\le s<t\le d} \frac{1}{32} b_s^2 b_t^2 \int_{\mathbb{R}^d} f''_{s,s}(x) f''_{t,t}(x) \, dx + O\left(\sum_{s=1}^{d} b_s^5\right). \tag{31}$$

3 Variance of the LBFP $\hat{f}_{\mathbf{n}}(x)$.

Let us first recall that

$$\hat{f}_{\mathbf{n}}(x) = \frac{1}{\hat{\mathbf{n}}\, b_1 \cdots b_d} \sum_{j_1, \cdots, j_d \in \{0,1\}^d} c_{j_1, \cdots, j_d}(x)\, \nu_{k_1+j_1, \cdots, k_d+j_d} \quad , \quad x \in I(k), \tag{32}$$

where $I(k) = \prod_{s=1}^{d} [x_{k,s} \, , \, x_{k,s} + b_s) = \prod_{s=1}^{d} \left[a_s - \frac{1}{2} b_s \, , \, a_s + \frac{1}{2} b_s\right)$,

$$Y_{\mathbf{i},k,j} = \begin{cases} 1 & \text{if } X_{\mathbf{i}} \in I(k,j) \\ \\ 0 & \text{otherwise} \end{cases} \qquad \text{and} \qquad p_k(j) = \int_{I(k,j)} f(x)\, dx \tag{33}$$

where $I(k,j) = \prod_{s=1}^{d} \left[x_{k,s} + \left(j_s - \frac{1}{2}\right) b_s \, , \, x_{k,s} + \left(j_s + \frac{1}{2}\right) b_s\right) = \prod_{s=1}^{d} [a_s + (j_s - 1) b_s \, , \, a_s + j_s b_s)$.
Then $\nu_{k_1+j_1, \cdots, k_d+j_d} = \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} Y_{\mathbf{i},k,j}$. Letting $j = (j_1, \cdots, j_d)$, from (32) we easily have

$$\begin{aligned} \mathrm{var} \hat{f}_{\mathbf{n}}(x) &= \frac{1}{\{\hat{\mathbf{n}} b_1 \cdots b_d\}^2} \sum_{j \in \{0,1\}^d} (c_j(x))^2 \ \mathrm{var}\left(\sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} Y_{\mathbf{i},k,j}\right) \\ &\quad + \frac{1}{(\hat{\mathbf{n}} b_1 \cdots b_d)^2} \sum_{j \neq j'} c_j(x) c_{j'}(x) \, \mathrm{cov}\left(\sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} Y_{\mathbf{i},k,j} \, , \, \sum_{\mathbf{i}' \in \mathcal{I}_{\mathbf{n}}} Y_{\mathbf{i}',k,j'}\right). \end{aligned} \tag{34}$$

For convenience, define

$$r_{1\mathbf{n}} = \frac{1}{(\hat{\mathbf{n}} b_1 \cdots b_d)^2} \sum_{\mathbf{i} \neq \mathbf{i}'} \mathrm{cov}\left(Y_{\mathbf{i},k,j} \, , \, Y_{\mathbf{i}',k,j}\right), \tag{35}$$

and, for $j \neq j'$,

$$r_{2\mathbf{n}} = \frac{1}{(\hat{\mathbf{n}} b_1 \cdots b_d)^2} \sum_{\mathbf{i} \neq \mathbf{i}'} \mathrm{cov}\left(Y_{\mathbf{i},k,j} \, , \, Y_{\mathbf{i}',k,j'}\right). \tag{36}$$

For notational simplicity, we have suppressed the dependence of $r_{1,n}$ and $r_{2,n}$ on k, j, j'.
First of all, we have that

$$\mathrm{var}\left(\sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} Y_{\mathbf{i},k,j}\right) = \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \mathrm{var}\left(Y_{\mathbf{i},k,j}\right) + \sum_{\mathbf{i} \neq \mathbf{i}'} \mathrm{cov}\left(Y_{\mathbf{i},k,j} \, , \, Y_{\mathbf{i}',k,j}\right), \tag{37}$$

where

$$\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \operatorname{var}\left(Y_{\mathbf{i},k,j}\right) = \hat{\mathbf{n}} p_k(j)\left\{1-p_k(j)\right\}. \tag{38}$$

Therefore the first term of the right member of (34) becomes

$$\frac{1}{\hat{\mathbf{n}}\left(b_1\cdots b_d\right)^2}\sum_{j\in\{0,1\}^d}\left\{c_j(x)\right\}^2 p_k(j)\left\{1-p_k(j)\right\} + \sum_{j\in\{0,1\}^d}\left(c_j(x)\right)^2 r_{1\mathbf{n}}. \tag{39}$$

Now, for $j \neq j'$, we have

$$\operatorname{cov}\left(\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} Y_{\mathbf{i},k,j}\,,\ \sum_{\mathbf{i}'\in\mathcal{I}_{\mathbf{n}}} Y_{\mathbf{i}',k,j'}\right) = \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\operatorname{cov}\left(Y_{\mathbf{i},k,j}\,,\ Y_{\mathbf{i},k,j'}\right) + \sum_{\mathbf{i}\neq\mathbf{i}'}\operatorname{cov}\left(Y_{\mathbf{i},k,j}\,,\ Y_{\mathbf{i}',k,j'}\right). \tag{40}$$

But for $j \neq j'$ we also have

$$\operatorname{cov}\left(Y_{\mathbf{i},k,j}\,,\ Y_{\mathbf{i}',k,j'}\right) = E\left(Y_{\mathbf{i},k,j}Y_{\mathbf{i},k,j'}\right) - E\left(Y_{\mathbf{i},k,j}\right)E\left(Y_{\mathbf{i},k,j'}\right) = -p_k(j)p_k(j'). \tag{41}$$

Thus the second term of the right member of (34) is

$$-\frac{1}{\left(\hat{\mathbf{n}}b_1\cdots b_d\right)^2}\sum_{j\neq j'} c_j(x)c_{j'}(x)\,\frac{1}{\hat{\mathbf{n}}\left(b_1\cdots b_d\right)^2}\,p_k(j)p_k(j') + \sum_{j\neq j'} c_j(x)c_{j'}(x)\,r_{2\mathbf{n}}. \tag{42}$$

We can notice that

$$\begin{aligned}\frac{1}{\hat{\mathbf{n}}}\left(\sum_{j\in\{0,1\}^d} c_j(x)p_k(j)\frac{1}{b_1\cdots b_d}\right)^2 \quad &= \quad \frac{1}{\hat{\mathbf{n}}\left(b_1\cdots b_d\right)^2}\sum_{j\in\{0,1\}^d}\left(c_j(x)\right)^2 p_k^2(j) \\ &\quad + \frac{1}{\left(\hat{\mathbf{n}}b_1\cdots b_d\right)^2}\sum_{j\neq j'} c_j(x)c_{j'}(x)\,\frac{1}{\hat{\mathbf{n}}\left(b_1\cdots b_d\right)^2}\,p_k(j)p_k(j').\end{aligned} \tag{43}$$

We have therefore shown that

$$\begin{aligned}\operatorname{var}\hat{f}_{\mathbf{n}}(x) \quad &= \quad \frac{1}{\hat{\mathbf{n}}\left(b_1\cdots b_d\right)^2}\sum_{j\in\{0,1\}^d} c_j^2(x)p_k(j) - \frac{1}{\hat{\mathbf{n}}\left(b_1\cdots b_d\right)^2}\left(\sum_{j\in\{0,1\}^d} c_j(x)p_k(j)\right)^2 \\ &\quad + \sum_{j\in\{0,1\}^d} c_j^2(x)\,r_{1\mathbf{n}} + \sum_{j\neq j'} c_j(x)c_{j'}(x)\,r_{2\mathbf{n}}.\end{aligned} \tag{44}$$

Using a Taylor-Lagrange expansion around the point $a = x_k + \frac{1}{2}b$, there exist $\xi_{s,s,j}$ and $\xi_{s,t,j}$ somewhere in the interval $[a + (j-1)b \ , \ a + jb)$ such that

$$\begin{aligned} f(x) &= f(a) + \sum_{s=1}^{d} f'_s(a)(x_s - a_s) + \frac{1}{2}\sum_{s=1}^{d} f''_{s,s}(\xi_{s,s,j})(x_s - a_s)^2 \\ &\quad + \frac{1}{2}\sum_{s \neq t} f''_{s,t}(\xi_{s,t,j})(x_s - a_s)(x_t - a_t). \end{aligned} \tag{45}$$

Expression (45) is used to evaluate $p_k(j) = \int_{I(k,j)} f(x)\, dx$. So we have

$$\begin{aligned} p_k(j) &= f(a)b_1 \cdots b_d + \sum_{s=1}^{d} f'_s(a) \int_{I(k,j)} (x_s - a_s)\, dx + \frac{1}{2}\sum_{s=1}^{d} f''_{s,s}(\xi_{s,s,j}) \int_{I(k,j)} (x_s - a_s)^2\, dx \\ &\quad \frac{1}{2}\sum_{s \neq t} f''_{s,t}(\xi_{s,t,j}) \int_{I(k,j)} (x_s - a_s)(x_t - a_t)\, dx. \end{aligned} \tag{46}$$

We calculate below the different integrals involved in (46). First,

$$\int_{I(k,j)} (x_s - a_s)\, dx = \int_{[(j-1)b \ , \ jb)} y_s\, dy = \frac{1}{2}(-1)^{j_s+1} b_s (b_1 \cdots b_d). \tag{47}$$

For $s = t$, we have

$$\int_{I(k,j)} (x_s - a_s)^2\, dx = \int_{[(j-1)b \ , \ jb)} y_s^2\, dy = \frac{1}{3} b_s^2 (b_1 \cdots b_d). \tag{48}$$

For $s \neq t$, we have

$$\int_{I(k,j)} (x_s - a_s)(x_t - a_t)\, dx = \int_{[(j-1)b \ , \ jb)} y_s y_t\, dy = \frac{1}{4}(-1)^{j_s+j_t} b_s b_t (b_1 \cdots b_d) \tag{49}$$

Thus, using (47), (48), (49), we obtain

$$\begin{aligned} p_k(j) &= f(a)(b_1 \cdots b_d) + \frac{1}{2}\sum_{s=1}^{d} f'_s(a)(-1)^{j_s+1} b_s (b_1 \cdots b_d) + \frac{1}{6}\sum_{s=1}^{d} f''_{s,s}(\xi_{s,s,j}) b_s^2 (b_1 \cdots b_d) \\ &\quad + \frac{1}{8}\sum_{s \neq t} f''_{s,t}(\xi_{s,t,j})(-1)^{j_s+j_t} b_s b_t (b_1 \cdots b_d). \end{aligned} \tag{50}$$

Now, we want to integrate $\mathrm{var}\hat{f}_{\mathbf{n}}(x)$ on $I(k)$, then on $\mathbb{R}^d$. First of all, in (44), we replace $p_k(j)$ by its value given by (50), which yields

$$\begin{aligned}
\mathrm{var}\hat{f}_{\mathbf{n}}(x) \quad = \quad & \frac{1}{\hat{\mathbf{n}}\,(b_1\cdots b_d)} \sum_{j\in\{0,1\}^d} c_j^2(x)\left[f(a)+\frac{1}{2}\sum_{s=1}^{d} f'_s(a)(-1)^{j_s+1}b_s\right. \\
& \left. +\frac{1}{6}\sum_{s=1}^{d} f''_{s,s}(\xi_{s,s,j})b_s^2+\frac{1}{8}\sum_{s\neq t} f''_{s,t}(\xi_{s,t,j})(-1)^{j_s+j_t}b_sb_t\right] \\
& -\frac{1}{\hat{\mathbf{n}}}\left(f(a)+\frac{1}{2}\sum_{j\in\{0,1\}^d} c_j(x)\sum_{s=1}^{d} f'_s(a)(-1)^{j_s+1}b_s\right. \\
& \left. +\frac{1}{6}\sum_{j\in\{0,1\}^d} c_j(x)\sum_{s=1}^{d} f''_{s,s}(\xi_{s,s,j})b_s^2+\frac{1}{8}\sum_{j\in\{0,1\}^d} c_j(x)\sum_{s\neq t} f''_{s,t}(\xi_{s,t,j})(-1)^{j_s+j_t}b_sb_t\right)^2 \\
& +\sum_{j\in\{0,1\}^d} c_j^2(x)\,r_{1\mathbf{n}}+\sum_{j\neq j'} c_j(x)c_{j'}(x)\,r_{2\mathbf{n}}.
\end{aligned} \tag{51}$$

Let us calculate the next integral :

$$\int_{I(k)}\sum_{j\in\{0,1\}^d} c_j^2(x)f(a)\,dx = f(a)\prod_{s=1}^{d}\int_{x_{k,s}}^{x_{k,s}+b_s}\left[2\left(\frac{x_s-x_{k,s}}{b_s}\right)^2-2\left(\frac{x_s-x_{k,s}}{b_s}\right)+1\right]dx_s. \tag{52}$$

Using the change of variable $y=\dfrac{x_s-x_{k,s}}{b_s}$ we get

$$\int_{I(k)}\sum_{j\in\{0,1\}^d} c_j^2(x)f(a)\,dx = f(a)\left(\frac{2}{3}\right)^d(b_1\cdots b_d). \tag{53}$$

Another integral to evaluate is

$$\frac{1}{2}\int_{I(k)}\sum_{j\in\{0,1\}^d}\sum_{s=1}^{d} f'_s(a)c_j^2(x)(-1)^{j_s+1}b_s = \frac{1}{2}\sum_{s=1}^{d} f'_s(a)\int_{I(k)}\left[u_s^2-(1-u_s)^2\right]\prod_{i=1,i\neq s}^{d}\left[u_i^2+(1-u_i)^2\right]dx. \tag{54}$$

By Fubini's theorem, the integral on the right-hand side of (54) is equal to

$$\int_{x_{k,s}}^{x_{k,s}+b_s} \left[\left(\frac{x_s - x_{k,s}}{b_s} \right)^2 - \left(1 - \frac{x_s - x_{k,s}}{b_s} \right)^2 \right] dx_s \prod_{i=1, i\neq s}^{d} \int_{x_{k,i}}^{x_{k,i}+b_i} \left[\left(\frac{x_i - x_{k,i}}{b_i} \right)^2 + \left(1 - \frac{x_i - x_{k,i}}{b_i} \right)^2 \right] dx_i. \tag{55}$$

We now calculate the first integral in (55), again using a change of variable $y = \dfrac{x_s - x_{k,s}}{b_s}$:

$$\int_{x_{k,s}}^{x_{k,s}+b_s} \left[\left(\frac{x_s - x_{k,s}}{b_s} \right)^2 - \left(1 - \frac{x_s - x_{k,s}}{b_s} \right)^2 \right] dx_s = b_s \int_0^1 (2y-1)\, dy = 0. \tag{56}$$

Thus we have

$$\frac{1}{2} \int_{I(k)} \sum_{j\in\{0,1\}^d} \sum_{s=1}^{d} f'_s(a) c_j^2(x) (-1)^{j_s+1} b_s \, dx = 0. \tag{57}$$

Now, we have to evaluate the following integral :

$$\frac{1}{6} \sum_{j\in\{0,1\}^d} \sum_{s=1}^{d} f''_{s,s}(\xi_{s,s,j}) b_s^2 \int_{I(k)} c_j^2(x)\, dx = \frac{1}{6} \sum_{s=1}^{d} f''_{s,s}(\xi_{s,s,j}) b_s^2 \prod_{s=1}^{d} \int_{x_{k,s}} \left(\frac{x_s - x_{k,s}}{b_s} \right)^{2j_s} \left(\frac{x_s - x_{k,s}}{b_s} \right)^{2(1-j_s)} dx_s. \tag{58}$$

Using obvious changes of variables and taking $j_s = 1$ or $j_s = 0$, we obtain

$$\frac{1}{6} \sum_{j\in\{0,1\}^d} \sum_{s=1}^{d} f''_{s,s}(\xi_{s,s,j}) b_s^2 \int_{I(k)} c_j^2(x)\, dx = \frac{1}{6} \sum_{j\in\{0,1\}^d} \sum_{s=1}^{d} f''_{s,s}(\xi_{s,s,j}) b_s^2 \left(\frac{1}{3} \right)^d (b_1 \cdots b_d). \tag{59}$$

In the same way, the following term is easily obtained :

$$\frac{1}{8} \sum_{j\in\{0,1\}^d} \sum_{s\neq t} f''_{s,t}(\xi_{s,t,j}) \int_{I(k)} c_j^2(x) (-1)^{j_s+j_t} b_s b_t \, dx = \sum_{s\neq t} \frac{1}{8} b_s b_t \sum_{j\in\{0,1\}^d} f''_{s,t}(\xi_{s,t,j}) (-1)^{j_s+j_t} \left(\frac{1}{3} \right)^d (b_1 \cdots b_d) \tag{60}$$

Therefore the first part of the right member of (51), after integrating on $I(k)$, is

$$\begin{aligned} &\int_{I(k)} \frac{1}{\hat{\mathbf{n}}\,(b_1 \cdots b_d)} \sum_{j\in\{0,1\}^d} c_j^2(x) \left[f(a) + \frac{1}{2} \sum_{s=1}^{d} f'_s(a) (-1)^{j_s+1} b_s \right. \\ &\left. + \frac{1}{6} \sum_{s=1}^{d} f''_{s,s}(\xi_{s,s,j}) b_s^2 + \frac{1}{8} \sum_{s\neq t} f''_{s,t}(\xi_{s,t,j}) (-1)^{j_s+j_t} b_s b_t \right] dx = \\ &f(a) \left(\frac{2}{3} \right)^d (b_1 \cdots b_d) + \frac{1}{6} \sum_{j\in\{0,1\}^d} \sum_{s=1}^{d} f''_{s,s}(\xi_{s,s,j}) b_s^2 \left(\frac{1}{3} \right)^d (b_1 \cdots b_d) \\ &+ \sum_{s\neq t} \frac{1}{8} b_s b_t \sum_{j\in\{0,1\}^d} f''_{s,t}(\xi_{s,t,j}) (-1)^{j_s+j_t} \left(\frac{1}{3} \right)^d (b_1 \cdots b_d). \end{aligned} \tag{61}$$

The development of the square of the second part of the right member of (51) is

$$f^2(a) + f(a)\sum_{s=1}^{d} f'_s(a)b_s \sum_{j\in\{0,1\}^d} c_j(x)(-1)^{j_s+1} + f(a)\frac{1}{3}\sum_{s=1}^{d} f''_{s,s}(\xi_{s,s,j})b_s^2$$
$$+f(a)\frac{1}{4}\sum_{j\in\{0,1\}^d} c_j(x)\sum_{s\neq t} f''_{s,t}(\xi_{s,t,j})(-1)^{j_s+j_t}b_s b_t + \cdots\ , \tag{62}$$

where the $\cdots$ denotes the terms of order b_s^3 or higher terms.

It is easy to see that

$$\int_{I(k)} f^2(a)\,dx = f^2(a)(b_1\cdots b_d). \tag{63}$$

We also have that

$$f(a)\sum_{s=1}^{d} f'_s(a)b_s \int_{I(k)} \sum_{j\in\{0,1\}^d} c_j(x)(-1)^{j_s+1}\,dx = 0 \tag{64}$$

We then calculate the following integral :

$$\int_{I(k)} f(a)\frac{1}{3}\sum_{s=1}^{d} f''_{s,s}(\xi_{s,s,j})b_s^2\,dx = f(a)\frac{1}{3}\sum_{s=1}^{d} f''_{s,s}(\xi_{s,s,j})b_s^2(b_1\cdots b_d), \tag{65}$$

and finally,

$$\int_{I(k)} f(a)\frac{1}{4}\sum_{j\in\{0,1\}^d}\sum_{s\neq t} f''_{s,t}(\xi_{s,t,j})c_j(x)(-1)^{j_s+j_t}b_s b_t = 0. \tag{66}$$

Hence the second part of the right member of (51), after integrating on $I(k)$, is

$$-\frac{1}{\hat{\mathbf{n}}}\left(f^2(a)(b_1\cdots b_d) + f(a)\frac{1}{3}\sum_{s=1}^{d} f''_{s,s}(\xi_{s,s,j})b_s^2(b_1\cdots b_d) + \cdots\right). \tag{67}$$

We can combine equations (61) and (67) to evaluate the integral of $\mathrm{var}\hat{f}_{\mathbf{n}}(x)$ on $I(k)$, giving

$$\begin{aligned}\int_{I(k)} \mathrm{var}\hat{f}_{\mathbf{n}}(x)\,dx \quad = \quad & \frac{1}{\hat{\mathbf{n}}(b_1\cdots b_d)}\Bigg[f(a)\left(\frac{2}{3}\right)^d (b_1\cdots b_d)\\ & + \frac{1}{6}\sum_{s=1}^{d} b_s^2 \sum_{j\in\{0,1\}^d} f''_{s,s}(\xi_{s,s,j})\left(\frac{1}{3}\right)^d (b_1\cdots b_d)\\ & + \sum_{s\neq t}\frac{1}{8}b_s b_t \sum_{j\in\{0,1\}^d} f''_{s,t}(\xi_{s,t,j})(-1)^{j_s+j_t}\left(\frac{1}{3}\right)^d (b_1\cdots b_d)\Bigg]\\ & -\frac{1}{\hat{\mathbf{n}}}\left(f^2(a)(b_1\cdots b_d) + f(a)\frac{1}{3}\sum_{s=1}^{d} f''_{s,s}(\xi_{s,s,j})b_s^2(b_1\cdots b_d) + \cdots\right)\\ & + \int_{I(k)}\sum_{j\in\{0,1\}^d} c_j^2(x)\, r_{1\mathbf{n}}\,dx + \int_{I(k)}\sum_{j\neq j'} c_j(x)c_{j'}\, r_{2\mathbf{n}}\,dx.\end{aligned} \tag{68}$$

Let us now examine the following integral :

$$\int_{I(k)}\sum_{j\in\{0,1\}^d} c_j^2(x)\, r_{1\mathbf{n}}\,dx \le C\left(\frac{1}{3}\right)^d (b_1\cdots b_d)\frac{1}{\hat{\mathbf{n}}}(b_1\cdots b_d)^{-1+\varepsilon} w_{k,j}, \tag{69}$$

that is,

$$\int_{I(k)}\sum_{j\in\{0,1\}^d} c_j^2(x)\, r_{1\mathbf{n}}\,dx \le C\,\frac{1}{\hat{\mathbf{n}}}(b_1\cdots b_d)\max\left\{(f(\xi_{k,j}))^{1/2}, Mf(\zeta_{k,j})\right\}(b_1\cdots b_d)^{-1+\varepsilon}. \tag{70}$$

The upper bound for the last term is exactly the same, viz.

$$\int_{I(k)}\sum_{j\neq j'} c_j(x)c_{j'}\, r_{2\mathbf{n}}\,dx \le C\,\frac{1}{\hat{\mathbf{n}}}(b_1\cdots b_d)\max\left\{(f(\xi_{k,j}))^{1/2}, Mf(\zeta_{k,j})\right\}(b_1\cdots b_d)^{-1+\varepsilon}. \tag{71}$$

To obtain the value of the integral $\int_{\mathbb{R}^d}\mathrm{var}\hat{f}_{\mathbf{n}}(x)\,dx$, it only remains to sum on k ($k\in\mathbb{Z}$). For that, we use Lemma 2. In (68), we have successively

$$\sum_k \frac{1}{\hat{\mathbf{n}}(b_1\cdots b_d)} f(a)\left(\frac{2}{3}\right)^d (b_1\cdots b_d) = \frac{1}{\hat{\mathbf{n}}(b_1\cdots b_d)}\left(\frac{2}{3}\right)^d\left[1 + O\left(\sum_{s=1}^{d} b_s\int_{\mathbb{R}^d}|f'_s(x)|\,dx\right)\right] \tag{72}$$

and

$$\sum_k \frac{1}{\hat{\mathbf{n}}(b_1\cdots b_d)}\frac{1}{6}\sum_{s=1}^d b_s^2 \sum_{j\in\{0,1\}^d} f''_{s,s}(\xi_{s,s,j})\left(\frac{1}{3}\right)^d (b_1\cdots b_d) =$$

$$\frac{1}{\hat{\mathbf{n}}(b_1\cdots b_d)}\frac{1}{6}\sum_{s=1}^d b_s^2\left(\frac{2}{3}\right)^d\left[\int_{\mathbb{R}^d} f''_{s,s}(x)\,dx + O\left(\sum_{s=1}^d b_s\int_{\mathbb{R}^d}\left|f'''_{s,s}(x)\,dx\right|\,dx\right)\right], \tag{73}$$

then

$$\sum_k \frac{1}{\hat{\mathbf{n}}(b_1\cdots b_d)}\sum_{s\neq t}\frac{1}{8}b_s b_t \sum_{j\in\{0,1\}^d} f''_{s,t}(\xi_{s,t,j})(-1)^{j_s+j_t}\left(\frac{1}{3}\right)^d (b_1\cdots b_d)$$

$$\leq \sum_{s\neq t}\frac{1}{8}b_s b_t\left(\frac{2}{3}\right)^d\left[\int_{\mathbb{R}^d} f''_{s,t}(x)\,dx + O\left(\sum_{s=1}^d b_s\int_{\mathbb{R}^d}\left|f'''_{s,t}(x)\right|\,dx\right)\right]. \tag{74}$$

For the second term of the right member of (68), we have

$$\frac{1}{\hat{\mathbf{n}}}\sum_k\left(f^2(a)(b_1\cdots b_d) + f(a)\frac{1}{3}\sum_{s=1}^d f''_{s,s}(\xi_{s,s,j})b_s^2(b_1\cdots b_d)+\cdots\right)$$

$$\leq \frac{1}{\hat{\mathbf{n}}}\left\{\int_{\mathbb{R}^d} f^2(x)\,dx + O\left(\sum_{s=1}^d b_s\int (f^2)'(x)\,dx\right)\right\} +$$

$$\frac{1}{\hat{\mathbf{n}}}\frac{1}{3}\sum_{s=1}^d b_s^2\left[\int_{\mathbb{R}^d} f(x)f''_{s,s}(x)\,dx + O\left(\sum_{s=1}^d b_s\int_{\mathbb{R}^d}\left|f'(x)f''_{s,s}(x)+f(x)f'''_{s,s}(x)\right|\,dx\right)\right] + O\left(\sum_{s=1}^d b_s^2\right). \tag{75}$$

Let us give an upper bound for the last two terms on the right-hand side of (68) :

$$\sum_k\int_{I(k)}\sum_{j\in\{0,1\}^d} c_j^2(x)\,r_{1\mathbf{n}}\,dx + \sum_k\int_{I(k)}\sum_{j\neq j'} c_j(x)c_{j'}\,r_{2\mathbf{n}}\,dx$$

$$\leq C\,\frac{1}{\hat{\mathbf{n}}}\sum_k (b_1\cdots b_d)\max\left\{(f(\xi_{k,j}))^{1/2}, Mf(\zeta_{k,j})\right\}(b_1\cdots b_d)^{-1+\varepsilon}. \tag{76}$$

Using one more time Lemma 2, we obtain the bound

$$O\left\{\frac{1}{\hat{\mathbf{n}}}(b_1\cdots b_d)^{-1+\varepsilon}\left[\left\|f^{1/2}\right\|_1 + O\left(\sum_{s=1}^d b_s\left\|(f^{1/2})'\right\|_1\right) + \|f\|_1 + O\left(\sum_{s=1}^d b_s\|f'\|_1\right)\right]\right\}. \tag{77}$$

In summary, we have

$$
\begin{aligned}
\int_{\mathbb{R}^d} \mathrm{var}\hat{f}_{\mathbf{n}}(x)\,dx \;\; \le \;\; & \frac{1}{\hat{\mathbf{n}}(b_1\cdots b_d)}\left(\frac{2}{3}\right)^d \left[1 + O\left(\sum_{s=1}^d b_s \int_{\mathbb{R}^d} \left|f_s'(x)\right|\,dx\right)\right] \\
& + \frac{1}{\hat{\mathbf{n}}(b_1\cdots b_d)}\frac{1}{6}\sum_{s=1}^d b_s^2 \left(\frac{2}{3}\right)^d \left[\int_{\mathbb{R}^d} f_{s,s}''(x)\,dx + O\left(\sum_{s=1}^d b_s \int_{\mathbb{R}^d} \left|f_{s,s}'''(x)\,dx\right|\,dx\right)\right] \\
& + \sum_{s\neq t}\frac{1}{8} b_s b_t \left(\frac{2}{3}\right)^d \left[\int_{\mathbb{R}^d} f_{s,t}''(x)\,dx + O\left(\sum_{s=1}^d b_s \int_{\mathbb{R}^d} \left|f_{s,t}'''(x)\right|\,dx\right)\right] \\
& + \frac{1}{\hat{\mathbf{n}}}\left\{\int_{\mathbb{R}^d} f^2(x)\,dx + O\left(\sum_{s=1}^d b_s \int (f^2)'(x)\,dx\right)\right\} \\
& + \frac{1}{\hat{\mathbf{n}}}\frac{1}{3}\sum_{s=1}^d b_s^2 \left[\int_{\mathbb{R}^d} f(x) f_{s,s}''(x)\,dx + O\left(\sum_{s=1}^d b_s \int_{\mathbb{R}^d} \left|f'(x) f_{s,s}''(x) + f(x) f_{s,s}'''(x)\right|\,dx\right)\right] \\
& + O\left(\sum_{s=1}^d b_s^2\right) + O\left\{\frac{1}{\hat{\mathbf{n}}}(b_1\cdots b_d)^{-1+\varepsilon}\left[\left\|f^{1/2}\right\|_1 + O\left(\sum_{s=1}^d b_s \left\|(f^{1/2})'\right\|_1\right)\right.\right. \\
& \left.\left. + \left\|f\right\|_1 + O\left(\sum_{s=1}^d b_s \left\|f'\right\|_1\right)\right]\right\}.
\end{aligned} \tag{78}
$$

We can then finally deduce

$$\left| \int_{\mathbb{R}^d} \mathrm{var} \hat{f}_{\mathbf{n}}(x)\, dx - \frac{1}{\hat{\mathbf{n}}(b_1 \cdots b_d)} \left(\frac{2}{3}\right)^d \right| \leq \frac{1}{\hat{\mathbf{n}}} \int_{\mathbb{R}^d} f^2(x)\, dx$$

$$+ O\left(\hat{\mathbf{n}}^{-1}(b_1 \cdots b_d)^{-1} \sum_{s=1}^{d} b_s \int_{\mathbb{R}^d} |f_s'(x)|\, dx \right) + O\left(\hat{\mathbf{n}}^{-1}(b_1 \cdots b_d)^{-1} \sum_{s=1}^{d} b_s^2 \right)$$

$$+ O\left(\hat{\mathbf{n}}^{-1}(b_1 \cdots b_d)^{-1} \sum_{s=1}^{d} b_s \left\| (f^2)' \right\|_1 \right) + O\left(\hat{\mathbf{n}}^{-1}(b_1 \cdots b_d)^{-1+\varepsilon} \left[\left\| f^{1/2} \right\|_1 \right] \right)$$

$$+ O\left\{ \hat{\mathbf{n}}^{-1}(b_1 \cdots b_d)^{-1+\varepsilon} \left[\left\| f^{1/2} \right\|_1 + O\left(\sum_{s=1}^{d} b_s \left\| (f^{1/2})' \right\|_1 \right) + \|f\|_1 + O\left(\sum_{s=1}^{d} b_s \|f'\|_1 \right) \right] \right\}. \tag{79}$$

4 On the inequalities (28) and (29).

Let us first consider the case $d = 1$. Then, with obvious notations, the standard frequency polygon is defined by

$$\hat{f}_{\mathbf{n}}(x) = \frac{1}{\hat{\mathbf{n}} b} \left\{ \left(\frac{1}{2} + \frac{x-a}{b} \right) \nu_1 + \left(\frac{1}{2} - \frac{x-a}{b} \right) \nu_0 \right\} \tag{80}$$

and we thus have

$$\hat{f}_{\mathbf{n}}(x) - \hat{f}_{\mathbf{n}}(a_k) = \left(\frac{\nu_1 - \nu_0}{\hat{\mathbf{n}} b} \right) \left(\frac{x - a_k}{b} \right). \tag{81}$$

Immediately,

$$\left| \hat{f}_{\mathbf{n}}(x) - \hat{f}_{\mathbf{n}}(a_k) \right| \leq 2b^{-2} |x - a_k| \qquad \text{a.s.} \tag{82}$$

From (81) we have

$$E\hat{f}_{\mathbf{n}}(x) - E\hat{f}_{\mathbf{n}}(a_k) = \frac{1}{\hat{\mathbf{n}} b} \left(EY_{\mathbf{i},1} - EY_{\mathbf{i},0} \right) \left(\frac{x - a_k}{b} \right) = \frac{1}{b}(p_1 - p_0) \left(\frac{x - a_k}{b} \right). \tag{83}$$

The mean-value theorem implies that there exists $\xi_0 \in I(k,0)$ and $\xi_1 \in I(k,1)$ such that $p_i/b = f(\xi_i)$ for $i = 1, 2$. From Assumption 2^* it follows that

$$\left| E\hat{f}_{\mathbf{n}}(x) - E\hat{f}_{\mathbf{n}}(a_k) \right| \leq |f(\xi_1) - f(\xi_0)| \left| \frac{x - a_k}{b} \right| \leq \frac{C}{b} |x - a_k|. \tag{84}$$

The same bound is obtained for all x and x_k such that $|x - x_k| < b$.

Now, we suppose that $d = 2$. Then

$$\hat{f}_{\mathbf{n}}(x) - \hat{f}_{\mathbf{n}}(a_k) = \frac{1}{\hat{\mathbf{n}} b_1 b_2} \sum_{j_1, j_2} \{c_j(x) - c_j(a_k)\} \nu_{k_1+j_1, k_2+j_2}. \tag{85}$$

We obtain

$$\begin{aligned}
c_{0,0}(x) - c_{0,0}(a_k) &= -\frac{x_1 - a_1}{b_1} - \frac{x_2 - a_2}{b_2} + \left(\frac{x_1 - a_1}{b_1}\right)\left(\frac{x_2 - a_2}{b_2}\right) \\
c_{1,0}(x) - c_{1,0}(a_k) &= \frac{x_1 - a_1}{b_1} - \frac{x_2 - a_2}{b_2} - \left(\frac{x_1 - a_1}{b_1}\right)\left(\frac{x_2 - a_2}{b_2}\right) \\
c_{0,1}(x) - c_{0,1}(a_k) &= -\frac{x_1 - a_1}{b_1} + \frac{x_2 - a_2}{b_2} - \left(\frac{x_1 - a_1}{b_1}\right)\left(\frac{x_2 - a_2}{b_2}\right) \\
c_{1,1}(x) - c_{1,1}(a_k) &= \frac{x_1 - a_1}{b_1} + \frac{x_2 - a_2}{b_2} + \left(\frac{x_1 - a_1}{b_1}\right)\left(\frac{x_2 - a_2}{b_2}\right).
\end{aligned} \tag{86}$$

Thus

$$\begin{aligned}
\hat{f}_{\mathbf{n}}(x) - \hat{f}_{\mathbf{n}}(a_k) &= \frac{1}{\hat{\mathbf{n}} b_1 b_2} \left[\left(\frac{x_1 - a_1}{b_1} \cdot \frac{x_2 - a_2}{b_2}\right)(\nu_{0,0} - \nu_{1,0} + \nu_{1,1} - \nu_{0,1}) \right. \\
&\left. + \frac{x_1 - a_1}{2b_1}(\nu_{1,1} - \nu_{0,0} + \nu_{1,0} - \nu_{0,1}) + \frac{x_2 - a_2}{b_2}(\nu_{1,1} - \nu_{0,0} - \nu_{1,0} + \nu_{0,1})\right]
\end{aligned} \tag{87}$$

and

$$\begin{aligned}
\left|\hat{f}_{\mathbf{n}}(x) - \hat{f}_{\mathbf{n}}(a_k)\right| &\leq \frac{1}{\hat{\mathbf{n}} b_1 b_2} \left[\left|\frac{x_1 - a_1}{b_1} \cdot \frac{x_2 - a_2}{b_2}\right| (|\nu_{0,0} - \nu_{1,0}| + |\nu_{1,1} - \nu_{0,1}|) \right. \\
&\left. + \left|\frac{x_1 - a_1}{2b_1}\right| (|\nu_{1,1} - \nu_{0,0}| + |\nu_{1,0} - \nu_{0,1}|) + \left|\frac{x_2 - a_2}{b_2}\right| (|\nu_{1,1} - \nu_{0,0}| + |\nu_{1,0} - \nu_{0,1}|)\right].
\end{aligned} \tag{88}$$

Assume that $\mathbf{n}$ is large enough so that b_1 and b_2 are less than 1. Note then that

$$|(x_1 - a_1)(x_2 - a_2)| \leq 2\,|(x_1 - a_1)(x_2 - a_2)| \leq (x_1 - a_1)^2 + (x_2 - a_2)^2 = \|x - a\|_2^2 \leq \|x - a_k\|_2. \tag{89}$$

The last two terms can be grouped together in (88), yielding

$$\frac{b_2\left|x_1-a_1\right|+b_1\left|x_2-a_2\right|}{2b_1b_2}\leq\frac{1}{2b_1b_2}\left(\left|x_1-a_1\right|+\left|x_2-a_2\right|\right)\leq\frac{C}{b_1b_2}\left\|x-a_k\right\|_2, \qquad (90)$$

because all norms are equivalent in a finite dimensional vector space. Thus we have

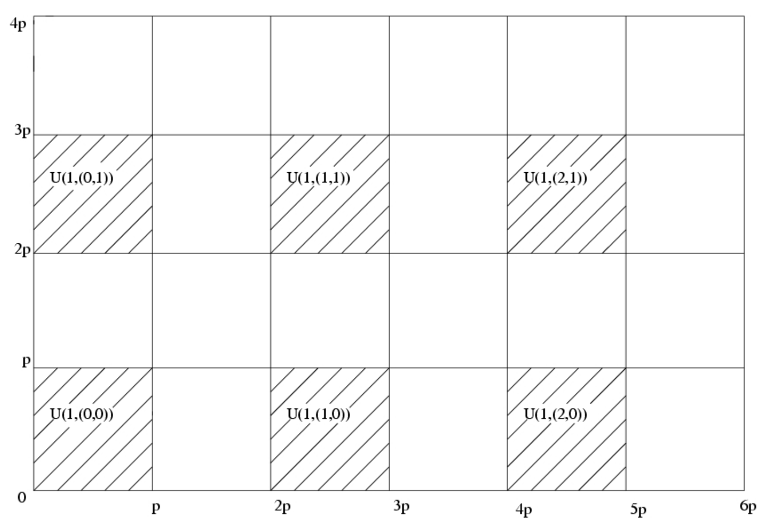
$$\left|\hat{f}_{\mathbf{n}}(x)-\hat{f}_{\mathbf{n}}(a_k)\right|\leq\frac{C}{\hat{\mathbf{n}}(b_1b_2)^2}\left\|x-a_k\right\|\left(\sum_{j_1,j_2,j_1',j_2'}\left|\nu_{j_1,j_2}-\nu_{j_1',j_2'}\right|\right). \qquad (91)$$

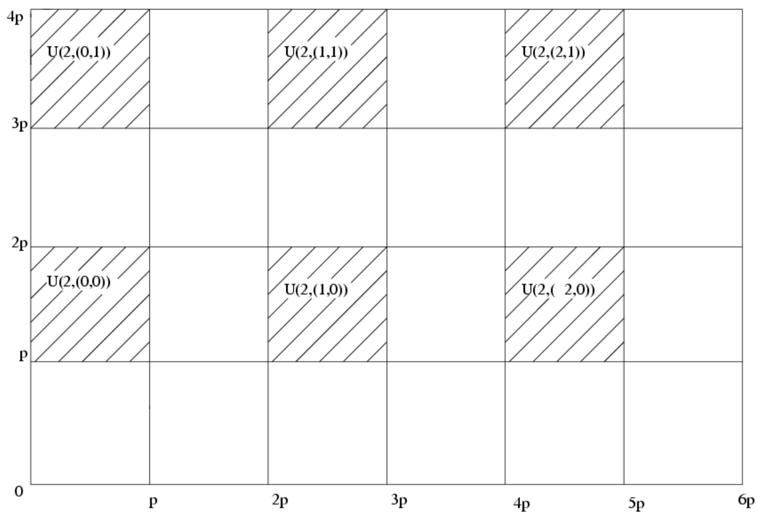
Taking the expectation in (87), using the mean-value as was done for $d=1$, we obtain

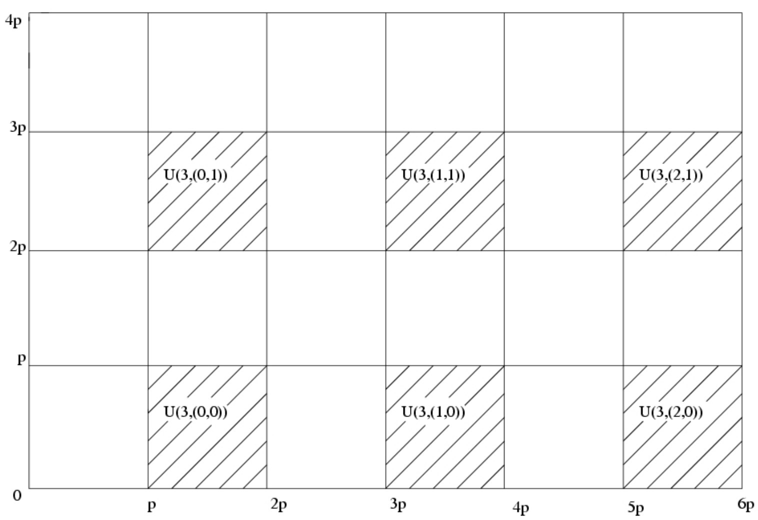
$$\left|E\hat{f}_{\mathbf{n}}(x)-E\hat{f}_{\mathbf{n}}(a_k)\right|\leq\frac{C}{b_1b_2}\sum_{j_1,j_2,j_1',j_2'}\left|f(\xi_{j_1,j_2})-f(\xi_{j_1',j_2'})\right|\left\|x-a_k\right\|\leq\frac{C}{b_1b_2}\left\|x-a_k\right\|. \qquad (92)$$

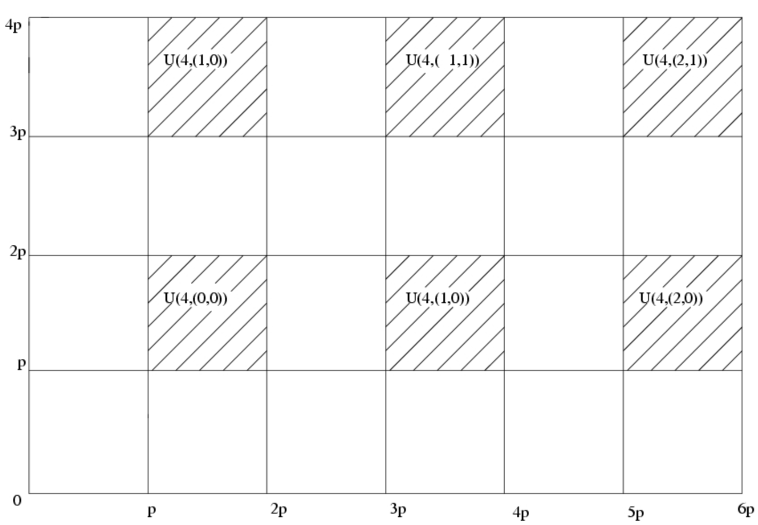
For $d>2$, we reason in the same manner. Products of the form $(x_i-a_i)(x_j-a_j)$ then appear, and we can use the same arguments as in the case with $d=2$.

The blocking technique used in the paper is shown below. In the following figures, we have omitted the $\mathbf{n}$ and the x. The first number in U is chosen between 1 and 2^N, the couple in brackets is the $\mathbf{j}$ of dimension two to be able to make a drawing.




4p
3p
2p
p
0
U(2,(0,1))
U(2,(1,1))
U(2,(2,1))
U(2,(0,0))
U(2,(1,0))
U(2,(2,0))
p
2p
3p
4p
5p
6p


4p
3p
2p
p
0
U(3,(0,1))
U(3,(1,1))
U(3,(2,1))
U(3,(0,0))
U(3,(1,0))
U(3,(2,0))
p
2p
3p
4p
5p
6p


4p
3p
2p
p
0
U(4,(1,0))
U(4,(1,1))
U(4,(2,1))
U(4,(0,0))
U(4,(1,0))
U(4,(2,0))
p
2p
3p
4p
5p
6p