Metrika manuscript No.
(will be inserted by the editor)

On the Hausdorff dimension of exceptional random sets generated by multivariate spacings

Claire Coiffard

Received: date / Accepted: date

Abstract We consider multivariate spacings blocks in the usual sense (refer to Deheuvels (Z. Wahrsch. Verw. Gebiete 64:411-424, 1983)). We consider the sets of exceptional points in the neighborhood of which such spacings are, infinitely often, unusually large. Our main result, in the spirit of Hawkes (Math. Proc. Camb. Phil. Soc.:293-303, 1981), shows that these sets constitute random fractals, whose Hausdorff dimensions are explicitly evaluated.

Keywords Spacings · Fractal · Hausdorff dimension · multidimensional variables

Mathematics Subject Classification (2000) 28A80

1 Introduction

Consider a sequence $U_1, U_2, \ldots$ of independent identically distributed random variables with a uniform distribution on $[0,1]$. If $0 = U_0^{(n)} < U_1^{(n)} < \ldots < U_n^{(n)} < U_{n+1}^{(n)} = 1$ denote the order statistics corresponding to $0, 1, U_1, \ldots, U_n$, then, for $i = 1, \ldots, n+1$, the *uniform spacing intervals* $[U_{i-1}^{(n)}, U_i^{(n)})$ have been extensively studied in the literature (Deheuvels, 1985; Greenwood, 1946; Pyke, 1965). By *uniform spacing (length)* is meant the Lebesgue measure (or length) of either of these intervals. For convenience, set $|A|$ for the the Lebesgue measure of a Borel subset A of $\mathbb{R}$. The spacings are denoted by

$$D_i^{(n)} = \left|[U_{i-1}^{(n)}, U_i^{(n)})\right| = U_i^{(n)} - U_{i-1}^{(n)}, \text{ for } i = 1, \ldots, n+1. \qquad (1)$$

The study of spacing-based statistics such as the maximal uniform spacing of order n, denoted by $M_n = \max_{1 \le i \le n+1} D_i^{(n)}$, has been initiated by Lévy (1939)

C. Coiffard (✉)
LSTA Université Pierre et Marie Curie, 175 rue du Chevaleret, 75013 Paris, FRANCE
Tel.: +(33)1 44 27 85 66
E-mail: claire.coiffard@upmc.fr

and Darling (1953). In 1978, Slud (1978) showed that $nM_n - \log n = O(\log_2 n)$, where $\log_j$ is the j-th iterated logarithm. Further refinements of this result are due to Deheuvels (1982) and Devroye (1981) (see also (Deheuvels, 1984; Deheuvels and Devroye, 1984)) who showed that

$$\limsup_{n\to\infty} \frac{nM_n - \log n}{2\log_2 n} = 1 \qquad \text{a.s.}$$

and

$$\liminf_{n\to\infty} \frac{nM_n - \log n}{\log_3 n} = -1 \qquad \text{a.s.}$$

Hawkes (1981) showed that uniform spacings generate *fractal random objects.* For each $x \in [0,1)$, let $u_n(x)$ be the uniquely defined spacing interval containing x, and denote by $Z_n(x) = |u_n(x)|$, the corresponding spacing length. The exceptional points of the process $\{Z_n(x) : x \in [0,1)\}$ fulfill the following property. (Versions of this result are also found in earlier papers such as that of Lévy (1937)).

Theorem A *With probability* 1 *we have*

$$\limsup_{n\to\infty} \frac{nZ_n(x)}{\log\log n} = 1, \tag{2}$$

for almost all $x \in [0,1)$.

To understand the meaning of Theorem A, we need to recall the following Theorem B, due to Lévy (1939).

Theorem B *With probability* 1 *we have*

$$\lim_{n\to\infty} \sup_{x\in[0,1)} \frac{nZ_n(x)}{\log n} = 1. \tag{3}$$

In addition to Theorem A, Hawkes (1981) gave a description, stated below in Theorem C, of the sets of points where the limit in (3) does not hold. Introduce the random sets

$$D(c) = \Big\{x \in [0,1) : \limsup_{n\to\infty} \frac{nZ_n(x)}{\log n} = c\Big\}, \tag{4}$$

and

$$U(c) = \Big\{x \in [0,1) : \limsup_{n\to\infty} \frac{nZ_n(x)}{\log n} > c\Big\}. \tag{5}$$

Hawkes (1981), characterized the *Hausdorff dimension* of these sets. Recall (Falconer, 1990) that the *Hausdorff dimension* of $E \subset [0,1]$ is defined by

$$\dim E = \inf\{c > 0 : s^c - \mathrm{mes}(E) = 0\} = \sup\{c > 0 : s^c - \mathrm{mes}(E) = \infty\},$$

where $s^c - mes(E)$ denotes the *Hausdorff measure* of E, defined by

$$s^c - \mathrm{mes}(E) = \lim_{\delta\to 0} \inf\Big\{\sum_{i\geq 1} d(U_i)^c : E \subseteq \bigcup_{i\geq 1} U_i, d(U_i) \leq \delta\Big\}. \tag{6}$$

We set for $d(U_i)$ the diameter of U_i, defined as the supremum of the Euclidean distance between two elements of U_i. The infimum in (6) is taken over all collections $\{U_i : i \geq 1\}$ of subsets with diameter $d(U_i) < \delta$ for all $i \geq 1$, and such that $E \subseteq \bigcup_{i \geq 1} U_i$.

Theorem C *For each $0 \leq c \leq 1$ we have, almost surely,*

$$\dim D(c) = \dim U(c) = 1 - c. \tag{7}$$

The purpose of this paper is to extend the results of Theorem C to a multivariate framework. We mention here that our results ,in the spirit of Deheuvels and Mason (1995), rely on strong limit theorems. Some weak versions of these results should rely on the corresponding description of exceptional sets. This will not be considered in the present paper.

2 Multivariate spacings

The systematic investigation of multivariate spacings started with Deheuvels (1983). We will see that his definition stated below is not the only possible one. In particular, one should cite the work of Janson (1987), and Deheuvels et al (1988). Here is the definition given by Deheuvels (1983).

Definition 1 *Let $\mathbf{U}_1, \ldots, \mathbf{U}_n$ be a random sample from the uniform distribution over $(0,1)^d$. A spacing block of order n is defined as any subset of $(0,1)^d$ of the form $\Pi_{r=1}^d]a_r, a_r + c[$ referred to as a square block) which does not contain any point among $\mathbf{U}_1, \ldots, \mathbf{U}_n$ and cannot be enlarged by a strict inclusion in a square block with the same properties (included in $(0,1)^d$ and having a void intersection with $\{\mathbf{U}_1, \ldots, \mathbf{U}_n\}$). The corresponding spacing length is defined as the length c of the side of the spacing block.*

Note that there exists with probability 1, an infinity of distinct spacing blocks associated with a given spacing length. Deheuvels (1983) studied the asymptotic behavior of these multivariate spacings. Let $M_{k,n}$ be the k-th maximal spacing associated to $\mathbf{U}_1, \ldots, \mathbf{U}_n$. He obtained the following theorem.

Theorem D *For a fixed $k \geq 1$,*

$$n\{M_{k,n}\}^d - \log n = O(\log \log n) \qquad a.s.\ as\ n \to \infty. \tag{8}$$

Now, let $\mathbf{z} \in (0,1)^d$. We denote by $\chi_n(\mathbf{z})$ the spacing length associated to $\mathbf{z}$. Consider the following sets.

$$V(c) = \{\mathbf{z} \in (0,1)^d : \limsup_{n \to \infty} n\chi_n^d(\mathbf{z}) / \log n > c\}, \tag{9}$$

$$L(c) = \{\mathbf{z} \in (0,1)^d : \limsup_{n \to \infty} n\chi_n^d(\mathbf{z}) / \log n = c\}. \tag{10}$$

In this paper, we establish the following theorem.

Theorem 1 *If* $0 \le c \le 1$, *then*

$$\dim V(c) = \dim L(c) = d(1-c). \tag{11}$$

This theorem provides a natural extension of Theorem C in a multivariate framework. Possible extensions of this result to more general spacings are mentioned in §5.

3 Preliminaries

In this section, we gather some facts and principles which will be used in the forthcoming section 4 to establish (11). First, we provide some methods to calculate the Hausdorff dimension of a set A. To evaluate the dimension $\dim A$ of a set, we split the proof into two parts. To bound $\dim A$ from above is, generally, simple. Indeed, if $s^t - mes(A) = 0$, then $\dim A \le t$. It is enough for any $\delta > 0$ to construct a δ-cover $\{U_i : i \ge 0\}$ of A such that

$$\lim_{\delta \to 0} \sum_{i \ge 0} d(U_i)^t = 0.$$

Lower bounds of the form $\dim A \ge t$ are typically more difficult to obtain. The solution is likely to follow from the "*mass distribution principle* ". (Remember that $d(A)$ denotes the diameter of A.)

Fact 1 (Mass distribution principle) *Let* μ *be a finite positive measure on* A. *For a fixed* t, *suppose that there exists* $\gamma > 0$ *and* $\delta > 0$ *such that*

$$\mu(U) \le \gamma d(U)^t,$$

for all sets U *with* $d(U) < \delta$. *Then,* $s^t - mes(A) \ge \mu(A)/\gamma > 0$ *and*

$$\dim A \ge t. \tag{12}$$

Proof See, e.g., Falconer (1990).

The following two facts are based on the "*mass distribution principle*", in Fact 1. We nned first to introduce a distance between hypercubes. Let A and B be two hypercubes of $[0,1]^d$. The distance between A and B is

$$d(A, B) = \inf_{\mathbf{x} \in A, \mathbf{y} \in B} |\mathbf{x} - \mathbf{y}|. \tag{13}$$

Fact 2 *Let* $[0,1]^d = E_0$ *and for all* $j \ge 1$, E_j *is a finite union of disjoint closed hypercubes* Q_i, $i = 1, \ldots, \ell(j)$. *Suppose that* E_{j-1} *contains at least* m_j *hypercubes of* E_j, *and for all* $i \ne i' \in \{1, \ldots, \ell(j)\}$, $d(Q_i, Q_{i'}) \ge \varepsilon_j$, *where* $0 < \varepsilon_{j+1} < \varepsilon_j$ *for each* $j \ge 1$. *Then,*

$$\dim \bigcap_{j=1}^{\infty} E_j \ge \liminf_{j \to \infty} d \frac{\log(m_1 \ldots m_{j-1})}{-\log(m_j \varepsilon_j^d)}. \tag{14}$$

Proof This result is similar to Example 4.6 p.64 (Falconer, 1990) in a multivariate framework. Suppose that E_{j-1} contains exactly m_j hypercubes of E_j. If not, we may throw out excess hypercubes which will not change the lower bound. We may define a positive measure μ on $\bigcap_{j\geq 1} E_j$, by assigning a mass $(m_1 \dots m_j)^{-1}$ to each of the $m_1 \dots m_j$ hypercubes of E_j.

Let U be a hypercube such that $0 < \text{vol}(U) < \varepsilon_1^d$, vol$(U)$ standing for the Lebesgue measure of U. We want to estimate $\mu(U)$. Let j be the integer such that

$$\varepsilon_j^d \leq \text{vol}(U) \leq \varepsilon_{j-1}^d.$$

The number of hypercubes of E_j that intersect U is

(i) at most m_j since U intersects at most one hypercube of E_{j-1},
(ii) at most $(d(U)/(\sqrt{d}\varepsilon_j) + 1)^d \leq 2^d \text{vol}(U)/\varepsilon_j^d$.

Therefore,

$$\begin{aligned}\mu(U) &\leq (m_1 \dots m_j)^{-1} \min\left\{\frac{2^d \text{vol}(U)}{\varepsilon_j^d}, m_j\right\} \\ &\leq (m_1 \dots m_j)^{-1} \Big(\frac{2^d \text{vol}(U)}{\varepsilon_j^d}\Big)^s m_j^{1-s},\end{aligned}$$

for every $0 \leq s \leq 1$. Hence,

$$\frac{\mu(U)}{d(U)^{ds}} \leq \frac{2^{ds} d^{-ds/2}}{(m_1 \dots m_{j-1}) m_j^s \varepsilon_j^{ds}}. \tag{15}$$

But (15) is bounded by a constant if

$$s < \liminf_{j\to\infty} \log(m_1 \dots m_{j-1}) / -\log(m_j \varepsilon_j^d).$$

The result is deduced directly from the application of the "mass distribution principle" (**Fact 1**).

Fact 3 *Fix $0 < s < 1$ and let $n_1, n_2, \dots$ be a rapidly increasing sequence of integers, say such that $n_{j+1} \geq \max\{n_j^j, 4n_j^{1/s}\}$ for each j. For each j, set $H_j \subset \mathbb{R}^d$ the hypercubes with side length $n_j^{-1/s}$, and the length between the midpoints of consecutive hypercubes being n_j^{-1}. Then,*

$$\dim \bigcap_{j\geq 1} H_j \geq ds. \tag{16}$$

Proof This is similar to Example 4.7 p. 65 (Falconer, 1990). Let $E_0 = [0,1]^d$, and for each $j \geq 1$, E_j consists of the hypercubes of H_j included in E_{j-1}. Then, each hypercube I of E_{j-1} contains at least $(n_{j-1}^{-1/s} n_j - 2)^d \geq 2^d$ hypercubes of E_j. Furthermore, the distance between two consecutive hypercubes of E_j (see (13)) is at least $n_j^{-1} - n_j^{-1/s} \geq 1/2n_j^{-1}$ for j large enough. Then we may apply Fact 2, noting that setting $n_{j-1}^{-d/s} n_j^{-d}$ rather than $(n_{j-1}^{-1/s} n_j - 2)^d$ does not affect the limit.

$$\dim \bigcap_{j=1}^{\infty} H_j \geq \dim \bigcap_{j=1}^{\infty} E_j = \liminf_{j\to\infty} d \frac{\log((n_1 n_2 \ldots n_{j-2})^{d-d/s} n_{j-1}^d)}{-\log(n_{j-1}^{-d/s} n_j^d n_j^{-d}) - \log 2^d}. \qquad (17)$$

Provided that n_j is sufficiently rapidly increasing, the term in $\log n_{j-1}$ is dominant in the numerator and denominator, we have (16).

4 Proof of Theorem 1

4.1 Upper bound

In this part, we show the inequality

$$\dim V(c) \leq d(1-c). \qquad (18)$$

If (18) holds, the fact that $L(c) \subseteq \bigcap_{n\geq 1} V(c - 1/n)$ shows that for all $n \geq 1$, $\dim L(c) \leq d(1 - c + 1/n)$ and hence

$$\dim L(c) \leq d(1-c). \qquad (19)$$

We now show that (18) holds. Let $p \geq 1$ be a positive integer. Set $v_j = j^{2p}$ for each $j \geq 1$. We introduce the sets

$$V_{p,j}(c) = \Big\{ \mathbf{z} \in (0,1)^d : v_j \chi_{v_j}^d(\mathbf{z}) > (c + \frac{1}{p}) \log v_j \Big\}.$$

and

$$V_p(c) = \Big\{ \mathbf{z} \in (0,1)^d : v_j \chi_{v_j}^d(\mathbf{z}) > (c + \frac{1}{p}) \log v_j, \text{ a.s. in } j \Big\}.$$

We first establish the following equality.

$$V(c) = \bigcup_{p\geq 1} V_p(c). \qquad (20)$$

The inclusion $\bigcup_{p\geq 1} V_p(c) \subseteq V(c)$ is obvious. To show the opposite inclusion, let $\mathbf{z} \in V(c)$. For some integer $p \geq 1$,

$$\limsup_{n\to\infty} n\chi_n^d(\mathbf{z})/\log n > c + 1/p.$$

But for all j, there exists an integer n such that $n - 1 \leq v_j = j^{2p} \leq n$. Then,

$$\frac{v_j \chi_{v_j}^d(\mathbf{z})}{\log v_j} \geq \frac{(n-1)\chi_n^d(\mathbf{z})}{\log n} = \frac{n\chi_n^d(\mathbf{z})}{\log n} \frac{n-1}{n}.$$

Then, we have

$$\limsup_{j\to\infty} v_j \chi_{v_j}^d(\mathbf{z})/\log v_j > c + 1/p,$$

and $V(c) \subseteq \bigcup_{p\geq 1} V_p(c)$. To obtain (18), all we have to do is to show that the inequality $s^{d(1-c)} - mes(V_p(c)) = 0$ holds for every $p \geq 1$. We first introduce some additional notation. Remember that v_j stands for the sample size and we set $\tau_j^{-1} = (c \log v_j / v_j)^{1/d}$. Denote by $\lfloor u \rfloor \leq u \leq \lfloor u \rfloor + 1$ the integer part of u. For $0 \leq i_1, \ldots, i_d \leq \lfloor v_j \rfloor$ and $j \geq 1$, we set

$$J(i_1, \ldots, i_d; j) = \prod_{r=1}^{d} \Big[\frac{i_r}{\tau_j}, \frac{i_r + 1}{\tau_j}\Big[. \tag{21}$$

Finally, we set for $0 \leq i_1, \ldots, i_d \leq \lfloor v_j \rfloor$ and $j \geq 1$,

$$\mathbb{1}_{i_1,\ldots,i_d;j} = \begin{cases} 1 & \text{when } J(i_1, \ldots, i_d; j) \text{ contains a spacing block with} \\ & \quad \text{side length greater than } \Big((c + \frac{1}{p})\frac{\log v_j}{v_j}\Big)^{1/d}, \\ 0 & \text{otherwise,} \end{cases} \tag{22}$$

and

$$I(i_1, \ldots, i_d; j) = \begin{cases} J(i_1, \ldots, i_d; j) & \text{if } \mathbb{1}_{i_1,\ldots,i_d;j} = 1, \\ \varnothing & \text{if } \mathbb{1}_{i_1,\ldots,i_d;j} = 0. \end{cases}$$

Note that $\bigcup_{i_1=0}^{\lfloor \tau_j \rfloor} \ldots \bigcup_{i_d=0}^{\lfloor \tau_j \rfloor} I(i_1, \ldots, i_d; j)$ is a $\sqrt{d}/\tau_j$-cover of $V_{p,j}(c)$. We want to show the inequality

$$s^{d(1-c)} - mes\Big(\bigcup_{j\geq 1} \bigcup_{i_1=0}^{\lfloor \tau_j \rfloor} \cdots \bigcup_{i_d=0}^{\lfloor \tau_j \rfloor} I(i_1, \ldots, i_d; j)\Big) < \infty. \tag{23}$$

As follows from the properties of Hausdorff measure, recalling (22), all we need is to show that

$$\sum_{j\geq 1} \sum_{i_1=0}^{\lfloor \tau_j \rfloor} \cdots \sum_{i_d=0}^{\lfloor \tau_j \rfloor} \Big(\sqrt{d}\frac{1}{\tau_j}\Big)^{d(1-c)} \mathbb{1}_{i_1,\ldots,i_d;j} < \infty.$$

Set $S_j = \sum_{i_1=0}^{\lfloor \tau_j \rfloor} \ldots \sum_{i_d=0}^{\lfloor \tau_j \rfloor} \mathbb{1}_{i_1,\ldots,i_d;j}$. The proof of (23) reduces to

$$\sum_{j\geq 1} \Big(\sqrt{d}\frac{1}{\tau_j}\Big)^{d(1-c)} ES_j < \infty. \tag{24}$$

In view of showing (24), we shall prove the next lemma.

Lemma 1 *Let* $\mathbf{U}_1, \ldots, \mathbf{U}_{v_j}$ *be a sample of length* v_j *and* J *a hypercube of* $[0, 1]^d$ *with side length* $\tau_j^{-1} = (c \log v_j / v_j)^{1/d}$. *Let* B_j *be the number of spacings blocks of* J *which length is greater than* $((c + 1/p) \log v_j / v_j)^{1/d}$. *We then have*

$$P(B_j \geq 1) \leq v_j^{-(c+1/p)}, \tag{25}$$

for all j *large enough.*

Proof Suppose that the event $\{B_j \geq 1\}$ holds. Then, there exists a spacing block $\Delta_{v_j,p}$ with side length greater than

$$\Big((c+1/p)\frac{\log v_j}{v_j}\Big)^{1/d},$$

and such that $\Delta_{v_j,p} \cap J \neq \varnothing$. So let $R(\Delta_{v_j,p})$ be the hypercube with side length $\big((c+1/p)\log v_j/v_j\big)^{1/d}$, and with one of his angles matching with one of the angles of $\Delta_{v_j,p}$. It is then obvious that

$$v_j\mu_{v_j}(R(\Delta_{v_j,p})) = 0.$$

Recall that the empirical measure of the sample $\mathbf{U}_1, \ldots, \mathbf{U}_n$, on a set A, is defined by

$$\mu_n(A) = \frac{1}{n}\#\{\mathbf{U}_i \in A : 1 \leq i \leq n\}.$$

So, we obtain the following inclusion.

$$\{B_j \geq 1\} \subseteq \{v_j\mu_{v_j}(R(\Delta_{v_j,p})) = 0\}.$$

But we know that

$$\begin{aligned} P(v_j\mu_{v_j}(R(\Delta_{v_j,p})) = 0) &= \Big(1 - \frac{(c+1/p)\log v_j}{v_j}\Big)^{v_j} \\ &= \big(1+o(1)\big)\frac{1}{v_j^{c+1/p}}. \end{aligned}$$

for j large enough. Then

$$P(B_j \geq 1) \leq P(v_j\mu_{v_j}(R(\Delta_{v_j,p})) = 0) \leq \frac{1}{v_j^{c+1/p}},$$

for j large enough.

Noting that $ES_j = \lfloor \tau_j \rfloor^d P(B_j \geq 1)$, (25) implies

$$\begin{aligned} \sum_{j\geq 1}\Big(\frac{\sqrt{d}}{\tau_j}\Big)^{d(1-c)} ES_j &\leq \sum_{j\geq 1} d^{d(1-c)/2}\tau_j^c P(B_j \geq 1) \\ &\leq \sum_{j\geq 1} d^{d(1-c)/2}(c\log v_j)\frac{1}{v_j^{1/p}}. \end{aligned}$$

Recalling that $v_j = j^{2p}$, we conclude that (24) holds. To conclude, we have $\dim V(c) \leq d(1-c)$ and $\dim L(c) \leq d(1-c)$.

4.2 Lower bound

In this part, we will establish the inequality

$$\dim L(c) \geq d(1-c). \tag{26}$$

which is enough to complete the proof of Theorem 1. Indeed, we may suppose that (26) holds. Noting that $\bigcup_{n\geq 1} L(c+1/n) \subseteq V(c)$, we deduce that $\dim V(c) \geq d(1-c)$. By combining this last result with (18) and (19) we readily obtain (11).

The idea underlying our proof is the following. We construct a mesh on $[0,1]^d$. For a large integer K_j, the hypercubes $\Pi_{r=1}^{d}[i_r/K_j;(i_r+1)/K_j)$ are the components of the mesh. We show that "most" of the hypercubes of the mesh contain a subspacing $T(i_1,\ldots,i_d,j)$, from a sample of size N_j of side length about $(c\log N_j)/N_j$. Thus, roughly speaking, the set

$$A_j = \bigcup_{i_1,\ldots,i_d \text{ even}} T(i_1,\ldots,i_d,j),$$

consists of $(1/2K_j)^d$ hypercubes of side length $(c\log N_j)/N_j$, separated by a distance of at least $1/K_j$. We will have to verify that we can apply Fact 3 for A_j, and that the number of hypercubes of the mesh which may not contain a spacing of the stated length is "small".

First we need some results on the distribution of spacings. Let $\{\delta_j, j\geq 1\}$ be a sequence decreasing to 0. The sequence $\{N_j, j\geq 1\}$ denotes the sample size. It fulfills the following conditions.

N1) the sequence $\{N_j, j\geq 1\}$ increases fast enough so the sequence $n_j^{-1} = \nu_j^{-1} = (c\log N_j/N_j)^{1-c}$ fulfills the conditions of Fact 3, (which means that $n_{j+1} \geq \max\{n_j^j, 4n_j^s\}$ for each j),

N2) the series $\sum_{j\geq 1} N_j^{-(c+\delta_j)}$ is convergent.

The following lemma will be an argument to say that the number of hypercubes of the mesh which may not contain a subspacing of the stated length is "small".

Lemma 2 *Suppose that $0\leq c\leq 1$ and $\{\delta_j, j\geq 1\}$ is a sequence decreasing to 0. Let J be a hypercube of $[0,1]^d$ of side length $\nu_j^{-1} = (c\log N_j/N_j)^{(1-c)/d}$, the sample size being N_j. Let M_j be the number of subspacings of J, with side length in*

$$\Big[\Big(c\frac{\log N_j}{N_j}\Big)^{\frac{1}{d}}, \Big((c+\delta_j)\frac{\log N_j}{N_j}\Big)^{\frac{1}{d}}\Big].$$

Then, $\sum_{j\geq 1} P(M_j=0)$ is a convergent series.

Proof For $j\geq 1$, let C_j be the number of spacings of J, with side length strictly lower to $(c\log N_j/N_j)^{1/d}$. We can see that

$$P(M_j\geq 1) \geq P(B_j=0)\times P(C_j=0).$$

Suppose that $\{C_j \geq 1\}$ holds. Therefore, there exists a spacing Υ with side length strictly lower to $(c \log N_j/N_j)^{1/d}$. Let Q be a hypercube of $[0,1]^d$ with side length $(c \log N_j/N_j)^{1/d}$, and such that $\Upsilon \subset Q$. Then,

$$\begin{aligned} P(C_j \geq 1) &\leq P(N_j \mu_{N_j}(Q) \geq 2) \\ &\leq \Big(\frac{c \log N_j}{N_j}\Big)^2. \end{aligned}$$

So,

$$P(C_j = 0) \geq 1 - \Big(\frac{c \log N_j}{N_j}\Big)^2. \tag{27}$$

Use Lemma 1 with the formal replacement of $1/p$ by δ_j, τ_j by ν_j and v_j by N_j. Then (25) and (27) jointly imply that

$$\begin{aligned} P(M_j \geq 1) &\geq \Big(1 - \frac{1}{N_j^{c+\delta_j}}\Big)\Big(1 - \Big(\frac{c \log N_j}{N_j}\Big)^2\Big) \\ &\geq 1 - \Big(\frac{c \log N_j}{N_j}\Big)^2 - \frac{1}{N_j^{c+\delta_j}} + \frac{(c \log N_j)^2}{N_j^{2+c+\delta_j}}. \end{aligned}$$

So, we have the inequality,

$$\begin{aligned} P(M_j = 0) &\leq \Big(\frac{c \log N_j}{N_j}\Big)^2 + \frac{1}{N_j^{c+\delta_j}} - \frac{(c \log N_j)^2}{N_j^{2+c+\delta_j}} \\ &\leq \frac{1}{N_j^{c+\delta_j}}, \end{aligned}$$

for j large enough. Finally, we use the condition N2) to show that this lemma holds.

We now give the details of the proof of Theorem 1. Remember that for $j \geq 1$, $\nu_j^{-1} = (c \log N_j/N_j)^{(1-c)/d}$. For all $0 \leq i_1, \dots, i_d \leq \lfloor \nu_j \rfloor$, we set

$$J(i_1, \dots, i_d, j) = \prod_{r=1}^{d} [i_r \nu_j^{-1}, (i_r + 1)\nu_j^{-1}).$$

If $J(i_1, \dots, i_d, j)$ contains a spacing (from the sample of size N_j) with side length in

$$\Big[\Big(c\frac{\log N_j}{N_j}\Big)^{1/d}, \Big((c+\delta_j)\frac{\log N_j}{N_j}\Big)^{1/d}\Big],$$

(it is the case if $M_j \geq 1$), we choose this spacing and call it $T(i_1, \dots, i_d, j)$. Otherwise, we choose any hypercube with side length $(c \log N_j/N_j)^{1/d}$, and call it $S(i_1, \dots, i_d, j)$. Next, we define

$$S_j = \bigcup_{i_1, \dots, i_d \text{ even}} S(i_1, \dots, i_d, j); \quad T_j = \bigcup_{i_1, \dots, i_d \text{ even}} T(i_1, \dots, i_d, j);$$

and finally

$$R_j = S_j \cup T_j; \quad R = \bigcap_{j\geq 1} R_j; \quad S = \bigcap_{j\geq 1} S_j; \quad T = \bigcap_{j\geq 1} T_j. \tag{28}$$

By construction, R_j consists of hypercubes of side length in

$$\Big[(c\log N_j/N_j)^{1/d}, ((c+\delta_j)\log N_j/N_j)^{1/d}\Big],$$

separated by a distance greater than $(c\log N_j/N_j)^{(1-c)/d}$. Note that Fact 3 is still true when the hypercubes H_j have side length equivalent to $n_j^{-1/s}$. Using the condition N1), we can apply Fact 3 with $n_j^{-1} = \nu_j^{-1} = (c\log N_j/N_j)^{(1-c)/d}$ and $s = 1-c$. We get

$$\dim R \geq d(1-c). \tag{29}$$

Now, by (28) $R = S\cup T$ and $T\cap V(c) \subset L(c)$. Moreover, by (29), $s^{d(1-c)} - mes(R) > 0$. Thus, if we show $s^{d(1-c)} - mes(S) = 0$, we will be able to conclude that $s^{d(1-c)} - mes(L(c)) > 0$ and so that (26) holds.

The hypercubes $S(i_1,\ldots,i_d,j)$, $0\leq i_1,\ldots,i_d\leq \nu_j$, are together a cover of S. We want

$$\sum_{j\geq 1}\sum_{i_1=0}^{\nu_j}\cdots\sum_{i_d=0}^{\nu_j} d(S(i_1,\ldots,i_d,j))^{d(1-c)} < \infty, \tag{30}$$

where $d(A)$ denotes the diameter of A. But (30) holds if

$$\sum_{j\geq 1}\mathbb{E}\sum_{i_1=0}^{\nu_j}\cdots\sum_{i_d=0}^{\nu_j} d(S(i_1,\ldots,i_d,j))^{d(1-c)} = \sum_{j\geq 1}\Big(\frac{1}{2}\nu_j\Big)^d P(M_j=0)\Big(\frac{c\log N_j}{N_j}\Big)^{1-c} < \infty. \tag{31}$$

Recalling $\nu_j^{-1} = (c\log N_j/N_j)^{(1-c)/d}$, we replace ν_j in (31) and using Lemma 2, we obtain

$$\sum_{j\geq 1}\Big(\frac{1}{2}\nu_j\Big)^d P(M_j=0)\Big(\frac{c\log N_j}{N_j}\Big)^{1-c} = \sum_{j\geq 1}\Big(\frac{1}{2}\Big)^d P(M_j=0) < \infty.$$

We have established that

$$s^{d(1-c)} - mes(L(c)) > 0 \text{ so } \dim L(c) \geq d(1-c).$$

This completes the proof of Theorem 1.

5 Extensions

As mentioned earlier, there are several different definitions for multivariate spacings. We only considered here squared shapes spacings. In their article, Deheuvels et al (1988) allow the shapes to be more general, like balls, rectangles, polyhedra. It would be interesting to obtain similar results for these types of spacings. We conjecture that our results should remain valid in this extended framework. However, the technicality of proofs becomes much more involved.

Acknowledgment

We thank the referee for his insightful comments.

References

Darling D (1953) On a class of problems related to the random division of an interval. Ann Math Statist 24:239–253
Deheuvels P (1982) Strong limiting bounds for maximal uniform spacings. Ann Probab 10(4):1058–1065
Deheuvels P (1983) Strong bounds for multidimensional spacings. Z Wahrsch Verw Gebiete 64:411–424
Deheuvels P (1984) Strong limit theorems for maximal spacings from a general univariate distribution. Ann Probab 12(4):1181–1193
Deheuvels P (1985) Spacings and applications. Probability and Statistical Decision Theory A, (Bad Tatzmannsdorf, 1983):1–30
Deheuvels P, Devroye L (1984) Strong law for the maximal k-spacing, when $k \leq c \log n$. Z Wahrsch Verw Gebiete 66:315–334
Deheuvels P, Mason D (1995) On the fractal nature of empirical increments. Ann Probab 23(1):
Deheuvels P, Einmahl J, Mason D, Ruymgaart F (1988) The almost sure behavior of maximal and minimal multivariate k_n-spacings. J of Multivariate anal 24:155–176
Devroye L (1981) Laws of the iterated logarithm for order statistics of uniform spacings. Ann Probab 9(5):860–867
Falconer K (1990) Fractal Geometry-Mathematical Foundations and Applications. Wiley, New-York
Greenwood M (1946) The statistical study of infectious diseases. J Roy Statist Soc, Ser A 109:85–110
Hawkes J (1981) On the asymptotic behavior of sample spacings. Math Proc Camb Phil Soc 90:293–303
Janson S (1987) Maximal spacing in several dimensions. Ann Probab 15(1):274–280
Lévy P (1937) Théorie de l'addition des variables aléatoires. Gauthier-Villars, Paris

Lévy P (1939) Sur la division d'un segment par des points choisis au hasard. CR Acad Sci Paris 208:147–149

Pyke R (1965) Spacings. J Roy Statist Soc, Ser B 27:395–449

Slud E (1978) Entropy and maximal spacings for random partitions. Z Wahrsch Verw Gebiete 41:341–352