Discrepancy of Point Sequences on Fractal Sets

Hansjörg Albrecher[1] Jiří Matoušek[2] Robert F. Tichy[1]

Dedicated to Prof. Kálmán Győry on the occasion of his 60th birthday

Abstract

We consider asymptotic bounds for the discrepancy of point sets on a class of fractal sets. By a method of R. Alexander, we prove that for a wide class of fractals, the L_2-discrepancy (and consequently also the worst-case discrepancy) of an N-point set with respect to halfspaces is at least of the order $N^{-1/2-1/2s}$, where s is the Hausdorff dimension of the fractal. We also show that for many fractals, this bound is tight for the L_2-discrepancy. Determining the correct order of magnitude of the worst-case discrepancy remains a challenging open problem.

Keywords:
discrepancy, fractals, halfspaces

1 Introduction

In the last few years the mathematical notion of fractal sets has turned out to be a powerful tool to effectively describe a variety of physical phenomena [8, 11]. As many state equations for physical structures lead to high-dimensional integrals, the practical calculation of their numerical solutions are of great importance. For a large class of functions, Quasi-Monte-Carlo-methods converge much faster than classical Monte-Carlo by using suitably chosen, low-discrepancy (in a sense evenly distributed) point sets instead of random points (for a survey on the subject see e.g. [14]). It is thus essential to study discrepancy bounds for point sequences on fractal sets.
In the sequel we will recall some of the important definitions and notions of fractals and discrepancy. Section 2 investigates bounds for the discrepancy with respect to halfspaces and Section 3 discusses the discrepancy for some other set systems.

Let $|\cdot|$ denote the usual norm on the d-dimensional Euclidean space $\mathbf{R}^d$. By the diameter of $U \subset \mathbf{R}^d$ we mean $\|U\| = \sup_{x,y\in U} |x-y|$, and a δ-covering of a given Borel set F is a countable family of sets $\{U_i\}$, each of diameter at most δ, whose union covers the set F. For $F \subset \mathbf{R}^d$ and $s \geq 0$ we now define for every $\delta > 0$

$$\mathcal{H}^s_\delta(F) = \inf\left\{ \sum_{i=1}^{\infty} \|U_i\|^s : U_i \text{ is a } \delta\text{-covering of } F \right\},$$

[1]Supported by the Austrian Science Foundation Project 12005-MAT
[2]Research supported by Charles University grants No. 158/99 and 159/99.

and subsequently the *s-dimensional Hausdorff measure* of F by

$$\mathcal{H}^s(F) = \lim_{\delta \to 0} \mathcal{H}^s_\delta(F). \tag{1}$$

This limit exists for every Borel set F and possesses the scaling property $\mathcal{H}^s(\lambda F) = \lambda^s \mathcal{H}^s(F)$ for $\lambda > 0$. $\mathcal{H}^s$ is a monotonically decreasing function in s and there exists a unique value of s where $\mathcal{H}^s$ jumps from ∞ to 0. This value is called the *Hausdorff dimension* of F

$$\dim_H F = \inf\{s : \mathcal{H}^s(F) = 0\} = \sup\{s : \mathcal{H}^s(F) = \infty\}.$$

The Hausdorff measure generalizes the concept of the Lebesgue measure in $\mathbf{R}^d$ ($\mathcal{H}^d$ is, up to a constant, equal to the d-dimensional Lebesgue measure in $\mathbf{R}^d$). A set $F \subset \mathbf{R}^d$ is called a *fractal set* if $\dim_H F$ is bigger than its topological dimension, which is always an integer. For a description of properties and techniques for practical calculations of $\dim_H F$, as well as for some different concepts of dimension, see [7].

In the following we outline some basic facts and concepts of fractal geometry: A *β-similitude* in $\mathbf{R}^d$ is a mapping $\psi : \mathbf{R}^d \to \mathbf{R}^d$ of the form

$$\psi(x) = \beta^{-1} U(x) + \alpha, \tag{2}$$

where $U : \mathbf{R}^d \to \mathbf{R}^d$ is a unitary, linear map and $\beta \in \mathbf{R}$, $\alpha \in \mathbf{R}^d$ are constants. Given a finite family $\psi = \{\psi_1, \dots, \psi_a\}$ of β-similitudes and a set $A \subset \mathbf{R}^d$, define

$$\psi(A) = \bigcup_{i=1}^{a} \psi_i(A).$$

Hutchinson [10] proved that for a given family ψ with $\beta > 1$ there exists a unique compact set $F \subset \mathbf{R}^d$ such that $F = \psi(F)$ and, moreover, if $A \subset \mathbf{R}^d$ is any compact set, the iterates $\psi^n(A)$ converge to F in the Hausdorff metric. Based on this result we can define a *self-similar* fractal with volume scaling factor a and linear scaling factor β to be a pair (ψ, F) consisting of a system $\psi = \{\psi_1, \dots, \psi_a\}$ of β-similitudes and its unique fixed point F, for which $\mathcal{H}^s(\psi_i(F) \cap \psi_j(F)) = 0 \; \forall i \neq j$ holds.
The *similarity dimension* of a self-similar fractal F is defined by

$$\dim_S F = \frac{\log a}{\log \beta},$$

and it is always greater than or equal to the Hausdorff dimension $\dim_H F$. Equality holds, if the system $\psi = \{\psi_1, \dots, \psi_a\}$ satisfies the so-called *open-set condition*, namely if there is a non-empty, bounded open set V such that

$$\psi(V) = \bigcup_{i=1}^{a} \psi_i(V) \subset V, \tag{3}$$

with the union disjoint (see e.g. [15]).
Typical examples of self-similar fractals that satisfy the open-set condition (3) are the Sierpiński gasket $G \subset \mathbf{R}^2$ and the Vićsek set $V_5 \subset \mathbf{R}^2$. The Sierpiński gasket is the fixed point of the system $\psi = \{\frac{1}{2}x + (0,0),\, \frac{1}{2}x + (\frac{1}{2},0),\, \frac{1}{2}x + (\frac{1}{4}, \frac{\sqrt{3}}{4})\}$; cf.[7]. The Vićsek set is obtained by partitioning a square into a 3×3-chessboard, deleting all

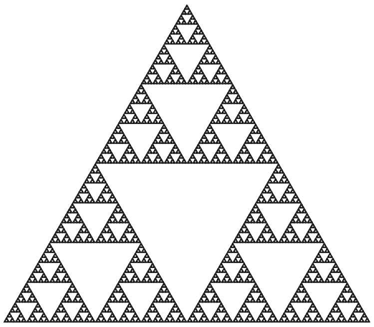

Figure 1: The Sierpiński gasket G

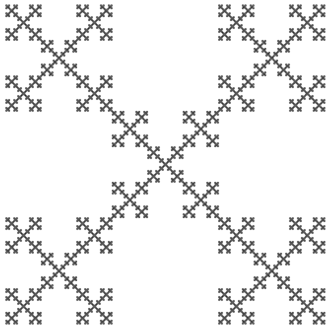

Figure 2: The Vićsek set V_5

four black squares, and repeating this procedure for every remaining square. Here we have $U : \mathbf{R}^2 \to \mathbf{R}^2$, $U(x) = x$, $\beta = 3$, $a = 5$, $\alpha_1 = (0,0)$, $\alpha_2 = (\frac{2}{3}, 0)$, $\alpha_3 = (\frac{1}{3}, \frac{1}{3})$, $\alpha_4 = (0, \frac{2}{3})$, $\alpha_5 = (\frac{2}{3}, \frac{2}{3})$, and $\psi(V_5) = V_5$.

The notion of discrepancy on fractals has first been introduced in [9] (where the special case of the Sierpiński gasket G has been considered): Let $\mathcal{D}$ be some system of Borel sets A, such that the boundary of each A is a μ-null set. Then the (volume) *discrepancy* of a point sequence $X = (x_1, \ldots, x_N)$ on a bounded fractal set F (embedded in the Euclidean space $\mathbf{R}^d$) with respect to $\mathcal{D}$ is defined by

$$D_N^{\mathcal{D}}(X) = \sup_{A \in \mathcal{D}} \left| \mu(A) - \frac{1}{N} \sum_{n=1}^{N} \chi_A(x_n) \right|,$$

where χ_A is the characteristic function of the set A and μ is the Hausdorff measure $\mathcal{H}^s$ normalized so that $\mu(F) = 1$. Important examples of systems $\mathcal{D}$ are the set $\mathcal{B}$ of all balls $B(x, \varepsilon) = \{y \in F \mid |x - y| < \varepsilon\}$ or the set $\mathcal{U}$ of all upper halfspaces $H \subset \mathbf{R}^d$ whose bounding hyperplanes intersect the convex hull of F (note that all other upper halfspaces always have zero discrepancy). For a detailed survey on discrepancy systems $\mathcal{D}$ we refer to [6]. We will also write $D_N^{\mathcal{D}}[F]$ for $\inf_X D_N^{\mathcal{F}}(X)$, where the infimum is over all N-term sequences $X = (x_1, \ldots, x_N)$, $x_1, x_2, \ldots, x_N \in F$.
In the definition above, D_N is a "worst-case" discrepancy. However, in many cases it is more convenient to work with an "average" discrepancy. An average discrepancy can be introduced by fixing a probability measure ω_0 on our set system $\mathcal{D}$ and then defining

$$D_{p,N}^{\mathcal{D}}(X) = \left(\int_{\mathcal{D}} \left| \mu(A) - \frac{1}{N} \sum_{n=1}^{N} \chi_A(x_n) \right|^p \mathrm{d}\omega_0(A) \right)^{\frac{1}{p}} \tag{4}$$

to be the L_p-discrepancy of X with respect to $\mathcal{D}$ ($p \geq 1$). Obviously $D_{p,N}^{\mathcal{D}}(X) \leq D_N^{\mathcal{D}}(X)$.
For the case $\mathcal{D} = \mathcal{U}$, we will use the particular measure ω_0 introduced below. As is well-known, there exists a motion-invariant measure ω on the $(d-1)$-dimensional hyperplanes in $\mathbf{R}^d$, which is unique up to scaling (see e.g. [13]). The ω-measure of the set of hyperplanes intersecting a given line segment in $\mathbf{R}^d$ is proportional to the length of the segment. In the plane, more generally, the ω-measure of the set of lines

intersecting a given convex set K is proportional to the perimeter of K (so-called *Crofton's formula*). The measure ω on hyperplanes induces a measure, denoted by ω_0, on the set $\mathcal{U}$ of the upper halfspaces whose boundaries intersect the convex hull of F. We assume that the scaling is chosen so that $\omega_0(\mathcal{U}) = 1$, i.e. ω_0 is a probability measure on $\mathcal{U}$.

2 Discrepancy with respect to halfspaces

2.1 Lower bounds

The following theorem is a straightforward generalization of the results of Alexander [2], [3].

Theorem 1 *Let $F \subset \mathbf{R}^d$ be a compact set of Hausdorff dimension $s > 1$, and let μ be the s-dimensional Hausdorff measure $\mathcal{H}^s$ on F, normalized so that $\mu(F) = 1$. Suppose that*

$$\mu(B(x,r)) \leq Cr^s \ \forall x \in F \ \forall r > 0 \tag{5}$$

(with $B(x,r)$ denoting the Euclidean ball of radius r centered at x and C being a constant). Then, for any finite point sequence $X = (x_1, \dots, x_N)$, $x_1, x_2, \dots, x_N \in F$, we have

$$D_N^{\mathcal{U}}(X) \geq D_{2,N}^{\mathcal{U}}(X) \gg N^{-\frac{1}{2}-\frac{1}{2s}}, \tag{6}$$

where $\mathcal{U}$ is the system of all upper halfspaces $H \subseteq \mathbf{R}^d$ whose boundaries intersect the convex hull of F, and the constant implied by $\gg$ depends on F, d, and s.

Proof: Let ν be the signed (discrepancy) measure

$$\nu = \nu^+ - \nu^- = \left(\frac{1}{N} \sum_{n=1}^{N} \delta_{\mathbf{x}_n} \right) - \mu, \tag{7}$$

on $\mathbf{R}^d$ (concentrated on F), where δ_x denotes the Dirac measure at a point x. We observe that $\nu(F) = 0$. We introduce the functional I by

$$I(\nu) = \int\limits_{\mathbf{R}^d} \int\limits_{\mathbf{R}^d} |p - q| \,\mathrm{d}\nu(p) \,\mathrm{d}\nu(q).$$

As shown by Alexander [1], the L_2-discrepancy can be expressed using $I(\nu)$; in our case

$$\left(D_{2,N}^{\mathcal{U}}(X)\right)^2 = \int_{H \in \mathcal{U}} \nu(H)^2 \,\mathrm{d}\omega_0(H) = -\gamma_F I(\nu), \tag{8}$$

where $\gamma_F > 0$ is a constant depending on the scaling factor of the measure ω_0 (for example, if $F \subset \mathbf{R}^2$ the convex hull of F has perimeter 1 then $\gamma_F = 1$). It remains to prove that

$$-I(\nu) \gg N^{-1-\frac{1}{s}}. \tag{9}$$

First we recall several lemmata of Alexander.

Lemma 1 (Alexander [3]) *Let ϕ be a signed discrete measure concentrated on the points $r_1, r_2, \ldots, r_k \in \mathbf{R}^d$ and let ψ be a signed bounded measure with compact support on $\mathbf{R}^d$ and with $\psi(\mathbf{R}^d) = 0$. For $x \in \mathbf{R}^d$, let ψ_x denote the translated measure given by $\psi_x(M) = \psi(M - x)$, and define the convolution $\psi * \phi$ by setting $\psi * \phi = \sum_{i=1}^k \phi(r_i)\psi_{r_i}$. Then we have*

$$-I(\psi * \phi) \leq -|\phi|^2 I(\psi),$$

where $|\phi| = \sum_{i=1}^k |\phi(r_i)|$ denotes the total variation of ϕ.

In the sequel, ϕ is a measure on $\mathbf{R}$ with finite support. We will work in the $(d+1)$-dimensional space $\mathbf{R}^{d+1}$. We consider $\mathbf{R}^d$ embedded in $\mathbf{R}^{d+1}$ as the coordinate hyperplane $x_{d+1} = 0$, while the real line $\mathbf{R}$ supporting ϕ is identified with the x_{d+1}-axis in $\mathbf{R}^{d+1}$. In this way, ϕ can also be considered as a measure in $\mathbf{R}^{d+1}$. If ψ is a measure in $\mathbf{R}^d$ then $\psi \times \phi$ and $\psi * \phi$ can be regarded as the same signed measure on $\mathbf{R}^{d+1}$.
Analogous to the definition of I, we let J be the "cross-term" functional

$$J(\nu_1, \nu_2) = \int_{\mathbf{R}^d} |p - q| \, \mathrm{d}\nu_1(p) \, \mathrm{d}\nu_2(q).$$

Lemma 2 (Alexander [3]) *Let ψ_1, ψ_2 be bounded signed measures on $\mathbf{R}^d$ and let ϕ be a signed measure on $\mathbf{R}$ with finite support. Then*

$$J(\psi_1 \times \phi, \psi_2 \times \phi) = \int_{\mathbf{R}^d} \int_{\mathbf{R}^d} J(\phi_p, \phi_q) \, \mathrm{d}\psi_1(p) \, \mathrm{d}\psi_2(q).$$

The next lemma provides some key properties of $J(\phi_p, \phi_q)$. We recall that the tth moment of a measure ϕ as above is $\int_{-\infty}^{\infty} x^t \, \mathrm{d}\phi(x)$.

Lemma 3 (Alexander[3]) *Let ϕ be a signed measure with finite support contained in the interval $[-\frac{1}{4}, \frac{1}{4}]$, with $\phi(\mathbf{R}) = 0$, with variation $|\phi| = 1$, and such that the first t moments of ϕ are 0. Let $a \in \mathbf{R}^d$ be a vector; we consider it as a vector in $\mathbf{R}^{d+1}$ orthogonal to the x_{d+1}-axis. Then $J(\phi, \phi_a)$ only depends on $|a|$, $-J(\phi, \phi_a)$ is a strictly decreasing function in $|a|$, we have $-J(\phi, \phi_a) \geq 0$ for all $a \in \mathbf{R}^d$, and*

$$-J(\phi, \phi_a) < |a|^{-2t-1} \quad \textit{for } |a| \geq 2.$$

Finally, we need to remark that measures ϕ as in the lemma (with the first t moments vanishing) exist such that $-I(\phi)$ is a positive constant (depending on t). An explicit construction of such a ϕ, using a finite-differencing formula, is provided in [5].
Following the proof of Theorem A in Alexander [3], we now look, instead of F and the sequence $X = (x_1, \ldots, x_N)$, at the similar set $F^* = K^{\frac{1}{s}} F$ and the sequence $Y = (y_1, \ldots, y_N)$, where $y_n = K^{\frac{1}{s}} x_n$, and $K > 0$ is a parameter to be specified later. Notice that, by the scaling property of fractal sets, $\mu(F^*) = K$. Furthermore, if we define

$$\bar{\nu} = \bar{\nu}^+ - \bar{\nu}^- = \left(\frac{K}{N} \sum_{n=1}^{N} \delta_{y_n} \right) - \mu,$$

then we get

$$I(\bar{\nu}) = K^{2+\frac{1}{s}} I(\nu).$$

In the following we will estimate $-I(\bar\nu)$. We let ϕ be a finitely supported measure on $\mathbf{R}$ as in Lemma 3 (for $t = d+1$; i.e. the support is contained in $[-\frac{1}{4}, \frac{1}{4}]$, $\phi(\mathbf{R}) = 0$, $|\phi| = 1$, the first $d+1$ moments vanish, and $-I(\phi) = c_1 > 0$). By Lemma 1 we get the inequality

$$\begin{aligned} -I(\bar\nu) &\geq -I(\bar\nu * \phi) \\ &= -I(\bar\nu \times \phi) \\ &= -I(\bar\nu^+ \times \phi) + 2J(\bar\nu^+ \times \phi, \bar\nu^- \times \phi) - I(\bar\nu^- \times \phi). \end{aligned}$$

We will show that the term $-I(\bar\nu^+ \times \phi)$ is large and that the other terms are either positive or sufficiently small in absolute value. We have, using Lemma 2,

$$-I(\bar\nu^+ \times \phi) = \int\limits_{\mathbf{R}^d}\int\limits_{\mathbf{R}^d} -J(\phi_p, \phi_q)\, \mathrm{d}\bar\nu^+(p)\, \mathrm{d}\bar\nu^+(q).$$

Since $\bar\nu^+$ is concentrated on the points $y_1, y_2, \ldots, y_N$, the last integral is actually a finite sum. We distinguish two types of terms in the sum: those with $p = q$ and those with $p \neq q$. The latter ones are all nonnegative by Lemma 3, and so their contribution (whose magnitude we cannot control in general) cannot decrease the estimate for $-I(\bar\nu)$. For the terms with $p = q$, we have $-J(\phi_p, \phi_p) = -I(\phi_p) = -I(\phi) = c_1 > 0$. Therefore,

$$-I(\bar\nu^+ \times \phi) \geq -I(\phi) \sum_{n=1}^{N} \bar\nu(y_n)^2 = c_1 \frac{K^2}{N}. \tag{10}$$

The term $-I(\bar\nu^- \times \phi)$ is easily seen to be nonnegative (using an analogue of (8) or the nonnegativity of $-J(\phi_p, \phi_q)$ from Lemma 3). For the remaining term $J(\bar\nu^+ \times \phi, \bar\nu^- \times \phi)$ we get, by Lemma 2,

$$J(\bar\nu^+ \times \phi, \bar\nu^- \times \phi) = \sum_{n=1}^{N} \bar\nu^+(\mathbf{y}_n) \int_{\mathbf{R}^d} J(\phi, \phi_{y_n - q})\, \mathrm{d}\mu(q).$$

To estimate the integral behind the summation sign, we divide the integration domain to the two regions $\{q : |q - y_n| \leq 2\}$ and $\{q : |q - y_n| > 2\}$. For the first region, we have

$$-\int\limits_{|y_n - q| \leq 2} J(\phi, \phi_{y_n - q})\, \mathrm{d}\mu(q) \leq |\phi|^2 \int\limits_{|y_n - q| \leq 2} \mathrm{d}\mu(q) \leq C2^s,$$

where the last inequality uses condition (5) and the scaling property of the Hausdorff measure:

$$\mu(F^* \cap B(x^*, K^{1/s} r)) = K\mu(F \cap B(x, r)) \leq KCr^s = C\left(K^{1/s} r\right)^s \quad \forall r > 0$$

and therefore

$$\mu(F^* \cap B(x^*, r^*)) < C(r^*)^s \quad \forall r^* > 0\, \forall x^* \in F^*.$$

The integral over the second region is handled by

$$
\begin{aligned}
-\int\limits_{|y_n-q|>2} J(\phi,\phi_{y_n-q})\,\mathrm{d}\mu(q) &= -\sum_{l=1}^{\infty}\int\limits_{2^l<|y_n-q|\le 2^{l+1}} J(\phi,\phi_{y_n-q})\,\mathrm{d}\mu(q)\\
&\le -\sum_{l=1}^{\infty} J_{2^l}\int\limits_{2^l<|y_n-q|\le 2^{l+1}} \mathrm{d}\mu(q)\\
&\le -C_1\sum_{l=1}^{\infty} J_{2^l}2^{s(l+1)},
\end{aligned}
$$

where J_{2^l} denotes the value of $J(\phi,\phi_{y_n-q})$ for $|y_n-q|=2^l$ and C_1 is a constant depending on d. Lemma 3 then yields ($t=d+1$)

$$
-J_{2^l}\le (2^l)^{-2(d+1)-1}<2^{-(d+1)(l+1)}.
$$

Since $s<d$, we have

$$
-J_{2^l}<2^{-(s+1)(l+1)}.
$$

Summing up the geometric series yields

$$
-\int\limits_{|y_n-q|\ge 2} J(\phi,\phi_{y_n-q})\,\mathrm{d}\mu(q)<\frac{1}{2}C_1
$$

and for the entire integration domain

$$
-\int\limits_{\mathbf{R}^d} J(\phi,\phi_{y_n-q})\,\mathrm{d}\mu(q)\le C2^s+\frac{1}{2}C_1=C_2.
$$

Then

$$
-2J(\bar{\nu}^+\times\phi,\bar{\nu}^-\times\phi)\le 2C_2K. \tag{11}
$$

For the choice $K=\frac{3C_2}{c_1}N$, (10) and (11) together imply

$$
-I(\bar{\nu})\ge c_1\frac{K^2}{N}-2C_2K>\frac{C_2^2}{c_1}N.
$$

Since $I(\bar{\nu})=K^{2+\frac{1}{s}}I(\nu)$, we have established (9), and this completes the proof. □

Remark to the proof. An alternative version of the proof of Theorem 1 is possible using the ideas of Chazelle, Matoušek, and Sharir [5]. The main difference is in that the measure ϕ is written down explicitly, using a finite differencing formula. Formally, one doesn't need to speak about the convolution (or product) of measures as in the proof above; instead, one may think of placing shifted signed copies of F in a suitable higher-dimensional space. Finally, while the nonnegativity of $-J(\phi_p,\phi_q)$ in Lemma 3 is not entirely easy to prove, the nonnegativity of the corresponding functional $-G(p,q)$ in the proof of Chazelle et al. is rather easy. For reader's convenience, we sketch this alternative presentation of the proof of Theorem 1 here.

Let t be a sufficiently large integer, and we consider the set $\tilde{F} = F \times [-1,1]^t \subset \mathbf{R}^{d+t}$. Let $\mathcal{U}_t$ denote the set of all upper halfspaces in $\mathbf{R}^{d+t}$ with boundary intersecting the convex hull of $\tilde{F}$. Regard ν (as defined above) as a signed measure in $\mathbf{R}^{d+t}$ (concentrated in the d-dimensional coordinate hyperplane $x_{d+1} = x_{d+2} = \cdots = 0$) and let ω_0 be a probability measure on $\mathcal{U}_t$ (defined in the same way as ω_0 was defined for $\mathcal{U}$). So the L_2-discrepancy of X can be written as

$$D_{2,N}^{\mathcal{U}}(X) = \gamma_F \int_{H \in \mathcal{U}_t} \nu(H)^2 \, \mathrm{d}\omega_0(H)$$

Let now $w = cN^{-1/s}$ be a parameter, where the constant $c > 0$ is chosen sufficiently small. For $j = 1, 2, \ldots, t$, let $\mathbf{w}_j \in \mathbf{R}^{d+t}$ be the vector with w in the $(d+j)$-th position and with 0's elsewhere. We define a new "replicated" measure $\tilde{\nu}$ by setting

$$\tilde{\nu}(A) = \sum_{b \in \{0,1\}^t} (-1)^{\sum_{j=1}^t b_j} \nu\Big(A - \sum_{j=1}^t b_j \mathbf{w}_j\Big).$$

As in [5], it can be shown that (analogue of Lemma 1)

$$\int_{\mathcal{U}_t} \nu(H)^2 \, \mathrm{d}\omega_0(H) \gg \int_{\mathcal{U}_t} \tilde{\nu}(H)^2 \, \mathrm{d}\omega_0(H),$$

and hence it is enough to deal with $\tilde{\nu}$. Following (8), we have

$$\int_{\mathcal{U}_t} \tilde{\nu}(H)^2 \, \mathrm{d}\omega_0(H) = -\int_{\mathbf{R}^{d+t}} \int_{\mathbf{R}^{d+t}} |p-q| \, \mathrm{d}\tilde{\nu}(p) \, \mathrm{d}\tilde{\nu}(q). \tag{12}$$

By the definition of $\tilde{\nu}$, (12) can be further expressed as

$$-\int_{\mathbf{R}^d} \int_{\mathbf{R}^d} G(p,q) \, \mathrm{d}\nu(p) \, \mathrm{d}\nu(q), \tag{13}$$

where

$$G(p,q) = 2^t \sum_{j=0}^{t} (-1)^j \binom{t}{j} \sqrt{|p-q|^2 + jw^2}.$$

The following properties of $G(p,q)$ are proved in [5], using basic properties of finite differencing:

(i) $-G(p,p) = c_3 w$ for a positive constant c_3 (dependent on d and t).

(ii) $|G(p,q)| = \mathcal{O}(|p-q|)$, with the constant of proportionality depending on d and t.

(iii) For $|p-q| \geq t$, we have $|G(p,q)| = \mathcal{O}(w^{2t}/|p-q|^{2t-1})$, with the constant of proportionality depending on d and t.

(iv) $G(p,q) \leq 0$ for all p, q.

Substituting the definition of ν into (13), we get

$$\int_{\mathcal{U}_t} \tilde{\nu}(H)^2 \, \mathrm{d}\omega_0(H) = -\int_{\mathbf{R}^d} \int_{\mathbf{R}^d} G(p,q) \, \mathrm{d}\nu(p) \, \mathrm{d}\nu(q) = E_{\mu\mu} + E_{\mu X} + E_{X2} + E_{XX}, \tag{14}$$

where

$$\begin{aligned}
E_{\mu\mu} &= -\int_{\mathbf{R}^d}\int_{\mathbf{R}^d} G(p,q)\,\mathrm{d}\mu(p)\,\mathrm{d}\mu(q), \\
E_{\mu X} &= \frac{1}{N}\cdot\sum_{p\in X}\int_{q\in F} G(p,q)\,\mathrm{d}\mu(q), \\
E_{X2} &= -\frac{1}{N^2}\sum_{p\in X} G(p,p), \text{ and} \\
E_{XX} &= -\frac{1}{N^2}\sum_{p,q\in X:\, p\neq q} G(p,q).
\end{aligned}$$

By (i), we have $E_{X2} = \frac{c_2 w}{N}$. By (iv), we get $E_{XX} \geq 0$ and $E_{\mu\mu} \geq 0$. Finally, for $E_{\mu X}$, we have

$$|E_{\mu X}| \leq \frac{1}{N}\sum_{p\in X}\int_{q\in\mathbf{R}^d} |G(p,q)|\,\mathrm{d}\mu(q)$$

and, using (ii) and (iii),

$$\begin{aligned}
\int_{q\in\mathbf{R}^d} |G(p,q)|\,\mathrm{d}\mu(q) &\leq \int_{|p-q|\leq tw} \mathcal{O}(w)\,\mathrm{d}\mu(q) + \int_{|p-q|>tw} \mathcal{O}\left(\frac{w^{2t}}{|p-q|^{2t-1}}\right)\mathrm{d}\mu(q) \\
&= \mathcal{O}\left(w\mu(B(p,tw)) + \sum_{k=0}^{\infty}\frac{w^{2t}}{(2^k tw)^{2t-1}}\,\mu(B(p,2^k tw))\right) \\
&= \mathcal{O}\left(w^{1+s} + w\sum_{k=1}^{\infty} 2^{-k(2t-1)}(2^k tw)^s\right) \\
&= \mathcal{O}(w^{1+s})
\end{aligned}$$

provided that $2t-1 > s$. Thus, we get that $|E_{\mu X}| \leq c_4 w^{1+s}$, with a constant c_4 depending on d, t, and s but not on the constant c in the definition of w. By choosing c small enough, we get that the right-hand side of (14) is $\Omega(N^{-1-1/s})$, and Theorem 1 follows. □

Remark: Theorem 1 gives a lower bound for the L_2-discrepancy w.r.t. halfspaces for all fractal sets satisfying (5) with finite $\mathcal{H}^s$. Hutchinson [10] has shown that all self-similar fractals F (as defined in Section 1) for which the open-set condition (3) is fulfilled, satisfy (5) and $0 < \mathcal{H}^s(F) < \infty$. Thus Theorem 1 is applicable for this large class of fractal sets.
It can be shown (see [7]) that every fractal F contains a compact subset E with $\mathcal{H}^s(E) > 0$, for which condition (5) is satisfied. However, the following construction shows that not necessarily $E = F$: Consider the Sierpiński gasket $G \subset \mathbf{R}^2 \subset \mathbf{R}^3$ (with Hausdorff dimension $s = \log 3/\log 2$) and generate infinitely, but countably many copies of G by rotating G in $\mathbf{R}^3$ around one of its base lines. Since a countable union of sets with Hausdorff dimension s still has Hausdorff dimension s, the union $G_1 \subset \mathbf{R}^3$ of these copies has Hausdorff dimension $s = \log 3/\log 2$, and, as can easily be verified, condition (5) does not hold for the points $x \in G_1$ where all the copies of G meet. Thus (5) is not a property of all fractal sets.

2.2 Optimality for the L_2-discrepancy

In this section, we show that the the lower bound for the L_2-discrepancy in Theorem 1 is asymptotically optimal for many fractals. To get a low-discrepancy N-point sequence, we use a random sampling similar to the one employed by Beck [4] in a slightly different context (let us remark that a similar sampling is used, too, in computer graphics, where it is called a *jittered sampling*).

Let us consider self-similar sets F with parameters a and $\beta > 1$ as defined in Section 1. Using the mappings $\psi_1, \ldots, \psi_a$, the set F is naturally subdivided into a^k subsets of level k, $k = 1, 2, \ldots$. The subsets of level 1 are $\psi_1(F)$, $\psi_2(F)$, $\ldots$, $\psi_a(F)$, the sets of level 2 are $\psi_i(\psi_j(F))$, $i, j = 1, 2, \ldots, a$, etc. For example, in the Sierpiński gasket, the sets of level k are the (nonempty) intersections of G with the equilateral triangles arising by k-times iterated subdivision of the top-level triangle.

We will only consider the values $N = a^k$, $k = 1, 2, \ldots$. If $F_1, F_2, \ldots, F_{a^k}$ are the sets of level k, we sample the point x_n from F_n at random, according to the probability distribution given by the restriction of the fractal measure μ to F_n, and the choices being mutually independent for all $n = 1, 2, \ldots, N = a^k$.

Let $H \in \mathcal{U}$ be a fixed halfspace, and let us consider the expectation of the squared discrepancy $\left(\mu(H) - \frac{1}{N}\sum_{n=1}^{N} \chi_H(x_n)\right)^2$. If F_n is completely contained in H or disjoint from H, then the contribution of x_n to the quantity $\mu(H) - \frac{1}{N}\sum_{n=1}^{N} \chi_H(x_n)$ is 0. If the boundary of H intersects the convex hull of F_n, then the contribution of x_n to this quantity is a random variable with zero expectation whose values lie in $[-\frac{1}{N}, \frac{1}{N}]$. Thus, if the boundary of H intersects the convex hull of $\kappa(H)$ of the F_n, then $\mu(H) - \frac{1}{N}\sum_{n=1}^{N} \chi_H(x_n)$ is the sum of $\kappa(H)$ independent random variables with zero expectation. The expected squared discrepancy of H is thus the variance of this sum, and it is bounded by $\kappa(H)/N^2$.

The expectation of $(D_N^{\mathcal{U}}(X))^2$, over a random choice of the N-point sequence X as above, is at most $\frac{1}{N^2}\int_{\mathcal{U}} \kappa(H)\,\mathrm{d}\omega_0(H)$. If $\mathcal{U}_n \subseteq \mathcal{U}$ is the set of the upper halfspaces whose boundary intersects the convex hull of F_n, we have $\omega_0(\mathcal{U}_n) = \beta^{-k}$ (recall that $\omega_0(\mathcal{U}) = 1$). Therefore

$$\int_{\mathcal{U}} \kappa(H)\,\mathrm{d}\omega_0(H) = \sum_{n=1}^{a^k} \omega_0(\mathcal{U}_n) = a^k\beta^{-k} = N^{1-1/s},$$

where $s = \log a/\log\beta$ is the similarity dimension of F. It follows that there exist an N-point X with $D_{2,N}^{\mathcal{U}}(X) \le N^{-1/2-1/2s}$. If F is a self-similar fractal sets that fulfills the open-set condition (3), the similarity dimension of F equals its Hausdorff dimension and thus we have shown that for the wide class of self-similar fractals satisfying the open-set condition (3), the lower L_2-bound (6) is also optimal.

2.3 Worst-case upper bounds

We shortly recall some definitions (see [12] for more details). Let $\mathcal{S}$ be a set system on a point (multi)set $X = \{x_1, \ldots, x_N\}$. A coloring χ is any mapping $X \to \{-1, 1\}$. The *combinatorial discrepancy* of $\mathcal{S}$ is defined by

$$\operatorname{disc}(\mathcal{S}) = \min_{\chi} \max_{S \in \mathcal{S}} |\chi(S)|,$$

where the minimum is taken over all colorings χ of X and $\chi(S) = \sum_{x \in S} \chi(x)$. The *primal shatter function* $\pi_{\mathcal{S}}(m)$ is the maximum possible number of distinct intersections of the sets of $\mathcal{S}$ with an m-point subset of X and the *dual shatter function* is the maximum number of equivalence classes on X defined by an m-element subfamily $\mathcal{Y} \subset \mathcal{S}$, where two points $x, y \in X$ are equivalent w.r.t. $\mathcal{Y}$ if x belongs to the same sets of $\mathcal{Y}$ as y does. Shatter functions can be used to derive upper bounds for the combinatorial discrepancy (cf. [12]).

In our case, we consider the set system $\mathcal{S}$ induced by halfspaces on an N-point (multi)set $\{x_1, x_2, \ldots, x_N\}$, $x_1, x_2, \ldots, x_N \in F$. If we can prove a bound of $f(N)$ for the combinatorial discrepancy of any N-point set in this case, for some function $f(N)$, it follows that the geometric discrepancy with respect to the fractal measure μ can also be bounded by $\mathcal{O}(f(N))$; that is, for all N, there are sequences $X = (x_1, x_2, \ldots, x_N)$ with $D_N^{\mathcal{U}}(X) = \mathcal{O}(f(N))$ (see [12] for a formulation with the Lebesgue measure instead of the measure μ; the generalization to the fractal measure μ is straightforward).

The primal shatter function for a set system induced by halfspaces in $\mathbf{R}^d$ is always bounded by $\mathcal{O}(m^d)$. Moreover, set systems with primal shatter function bounded by $\mathcal{O}(m^d)$ have combinatorial discrepancy $\mathcal{O}(N^{-1/2-1/2d})$ (for any $d > 1$; see [12]). Consequently, we have the following bound for the discrepancy with respect to halfspaces, for point sets on fractals in $\mathbf{R}^d$:

$$D_N^{\mathcal{U}}[F] = \mathcal{O}\left(N^{-1/2-1/2d}\right).$$

One might wonder whether this bound could be improved by showing a better estimate for the primal or dual shatter functions for point sets selected from a fractal. Unfortunately, this is not the case in general.

For the primal shatter function, let us consider a fractal embedded in $\mathbf{R}^2$ and take two parallel lines, each cutting the fractal at infinitely many points (say), and choose $m/2$ points on each of the lines. Then there are $\Omega(m^2)$ subsets cut off by a halfplane. A similar construction can be done in $\mathbf{R}^d$, showing a lower bound of $\Omega(m^d)$.

For the dual shatter function, let us consider the Sierpiński-gasket and choose $m/2$ parallel lines very close to the base of the top-level triangle, and $m/2$ other suitable lines perpendicular to the base. Then, among the regions defined by these lines in the plane, there are $\Omega(m^2)$ intersected by the considered fractal, and so the dual shatter function is $\Omega(m^2)$, too. This indicates that the shatter functions are both too large and thus are not the appropriate tool to obtain better upper bounds for the worst-case discrepancy in the fractal case. Deciding whether the lower bound in Theorem 1 is tight (possibly up to a logarithmic factor) at least for some nontrivial fractals, such as the Sierpiński gasket, is a challenging open problem.

3 Discrepancy with respect to other set systems

Axis-parallel boxes. For the discrepancy w.r.t. axis-parallel boxes there exists an upper bound for the combinatorial discrepancy in the Euclidean space $\mathbf{R}^d$ [12]. Again these results immediately lead to upper bounds for the discrepancy on fractals:

$$D_N^{\mathcal{K}_d}[F] = \mathcal{O}\left(\log^{d+1/2} N \frac{\sqrt{\log\log N}}{N}\right),$$

where $\mathcal{K}_d$ denotes the set of axis-parallel boxes.

For lower bounds a derivation of a general result (possibly by adapting Roths or Becks lower bound methods; see e.g. [12]) seems intriguing, and it is not even clear what bounds should be expected.

For the case of the Lebesgue measure and axis-parallel boxes (the classical discrepancy problem), the L_2-discrepancy is known to be of the order $(\log N)^{(d-1)/2}/N$ (by the celebrated results of Roth). The worst-case discrepancy in this situation, for $d=2$, is known to be of the order $(\log N)/N$ (results of Schmidt and Van der Corput); for larger d, the order of magnitude is not known but it is generally believed to be $(\log N)^{d-1}/N$. One might thus be tempted to conjecture that for an s-dimensional fractal measure, the discrepancy for axis-parallel boxes might be of the order $(\log N)^{s-1}/N$ (worst case) and $(\log N)^{(s-1)/2}/N$ (L_2-average case). The following simple example shows that this is false. Namely, let $C \subset [0,1]$ be the classical Cantor "middle third" set (cf. [7]), and put $F = C \times C$. There is a measure-preserving and monotone bijection g between $[0,1]$ with the Lebesgue measure and $C \setminus E$ with the Hausdorff measure, where E is a countable set (the right endpoint of the intervals forming $[0,1] \setminus C$). Then $g \times g$ is a measure-preserving map of $[0,1]$ to F minus a negligibly small set and, moreover, axis-parallel rectangles are mapped to axis-parallel rectangles. Consequently, the discrepancy with respect to the fractal measure on F behaves in the same way as the "usual" discrepancy w.r.t. the Lebesgue measure. For example, we have

$$D_N^{\mathcal{K}_2}[F] = \Theta\left(\frac{\log N}{N}\right).$$

It would be very interesting to find good lower and upper bounds for some fractal sets that do not have such a simple product structure, such as the Vićsek set or the Sierpiński gasket. In particular, is there a set of Hausdorff dimension $s \in (1,2)$ in the plane, such that the worst-case discrepancy w.r.t. its fractal measure (for axis-parallel rectangles) is $o((\log N)/N)$?

Elementary Sets. Self-similar fractals can be constructed by recursion using elementary sets and mapping them in each level of construction according to (2). It is thus natural to consider these elementary sets as the set system $\mathcal{D}$. As is shown in [9], the elementary discrepancy of a set similar to the classical Van der Corput set is $\mathcal{O}(\frac{1}{N})$, showing that there is no nontrivial lower bound for the worst-case discrepancy.

References

[1] Alexander, R. Generalized sums of distances. *Pacific J. Math.* 56, 297-304 (1975)

[2] Alexander, R. Geometric methods in the theory of uniform distribution. *Combinatorica*, 10(2), 115-136 (1990).

[3] Alexander, R. Principles of a new method in the study of irregularities of distribution. *Invent. Math.* 103, 279-296 (1991)

[4] Beck, J. Irregularities of distribution I. *Acta Math.*, 159, 1–49 (1987).

[5] Chazelle, B., Matoušek, J., and Sharir, M. An elementary approach to lower bounds in geometric discrepancy. *Discr. Comput. Geom.*, 13:363-381 (1995).

[6] Drmota, M., Tichy, R. *Sequences, Discrepancies and Applications.* Lecture Notes in Mathematics 1651, Springer (1996)

[7] Falconer, K. *Fractal Geometry.* John Wiley, Chichester (1990).

[8] Feder, J. *Fractals.* Plenum Press XV, New York (1988)

[9] Grabner, P., Tichy, R. Equidistribution and Brownian Motion on the Sierpinski Gasket. *Mh. Math.* 125, 147-164 (1998)

[10] Hutchinson, J. E. Fractals and self-similarity. *Indiana Univ. Math. J.* 30,No.5, 713-747 (1981)

[11] Mandelbrot, B. B. *The fractal geometry of nature.* San Francisco (W. H. Freeman&Co.) (1982)

[12] Matoušek, J. *Geometric Discrepancy. An Illustrated Guide.* Springer-Verlag, Berlin etc. (1999)

[13] Santaló, L. A. *Integral Geometry and Geometric Probability.* Addison-Wesley, Reading, MA, 1976.

[14] Spanier, J., Maize, E. Quasi-random methods for estimating integrals using relatively small samples. *SIAM Rev.*, 36:18-44 (1994)

[15] Yamaguti, M., Hata, M., Kigami, J. L. A. *Mathematics of Fractals.* Transl.Math.Monogr. Vol.167, AMS (1997).

H. Albrecher and R.F. Tichy
Department of Mathematics
Technical University Graz
Steyrergasse 30
8010 Graz
Austria

J. Matoušek
Department of Applied Mathematics
Charles University
Malostranské nám. 25
118 00 Praha 1
Czech Republic