

SETS OF NEAR-OPTIMAL POINTS FOR INTERPOLATION ON THE SQUARE*

S. DE MARCHI

Department of Computer Science

University of Verona

S.da Le Grazie, 15 - 37134 Verona (Italy)

E-mail: stefano.demarchi@univr.it

Firstly, we present new sets of nodes for *polynomial interpolation on the square* that are asymptotically distributed w.r.t. the *Dubiner metrics*⁸. Then, we shall deal with two particular families which show Lebesgue constants that numerically grow like $\log^2(n)$, with n the degree of the interpolating polynomial. In the non-polynomial case with *radial basis functions* we also present two families of nearly-optimal interpolation points which can be determined *independently* of the radial function. One of these families can be conceptually described as a *Leja sequence* in the bivariate case.

1. Optimal points for polynomial interpolation

The problem of finding “optimal” points for polynomial interpolation in the one-dimensional case and in higher dimensions, has attracted the attention of researchers for many years.

In the one-dimensional case and on the canonical interval $[-1, 1]$, those optimal points are characterized by the *Bernstein-Erdős conjecture*, proved by de Boor and Pinkus in the seventies⁴, saying that *degree n optimal canonical points are a unique and symmetric set*. Their positions were found only in 1989 by J.R. Angelos, E.H. Kaufman et al.¹.

If in the one-dimensional case the problem can be considered essentially solved and every good interpolation point set has asymptotically the arccosine distribution, in the multivariate setting the problem is still open. Many attempts were done recently for solving the problem on different bidimensional domains. As some examples, we remind the following recent works^{4,10,3,15} where different types of near-optimal interpolation sets were

*This work was supported by the University of Verona funds for the year 2004.

presented on different bidimensional domains.

Let $\Omega \subset \mathbb{R}^d$ be compact. We call *optimal* polynomial interpolation points a set $X_N^* \subset \Omega$ (N , cardinality), so that the *Lebesgue constant*

$$\Lambda_n(X_N) = \max_{\mathbf{x} \in \Omega} \lambda_n(\mathbf{x}; X_N), \quad \lambda_n(\mathbf{x}; X_N) = \sum_{i=1}^N |\ell_i(\mathbf{x}; X_N)|, \quad (1)$$

defined for all sets $X_N = \{\mathbf{x}_1, \dots, \mathbf{x}_N\} \subset \Omega$ (unisolvant for polynomial interpolation of degree n), attains its minimum at $X_N = X_N^*$. $\lambda_n(\mathbf{x}; X_N)$ is the *Lebesgue function* of X_N and the ℓ_i are the fundamental Lagrange polynomials of degree n , and $N = \dim(\mathbb{P}_n(\mathbb{R}^d))$, $N = \binom{n+d}{d}$.

In the present paper, we will *only* consider “good” polynomial interpolation points on the square, that is sets of points whose *Lebesgue constant* does not grow geometrically with the degree n . That is why we call them *near-optimal*.

As it is well-known, in the one-dimensional case with $\Omega = [-1, 1]$, *Chebyshev*, *Fekete*, *Leja* as well as the *zeros of Jacobi orthogonal polynomials* are near-optimal points for polynomial interpolation. The Lebesgue constants increase *logarithmically* in the dimension N of the corresponding polynomial space². All these points have asymptotically the *arccosine distribution*, that is they are asymptotically equidistributed w.r.t. the arccosine metrics^{3,11}.

*Dubiner*⁸, proposed a metrics which coincides in the one-dimensional case with the *arccosine metrics*:

$$\mu_{[-1,1]}(x, y) = |\cos^{-1}(x) - \cos^{-1}(y)|, \quad \forall x, y \in [-1, 1].$$

By using the *van der Corput-Schaake inequality*¹³, we can re-write it as

$$\mu_{[-1,1]}(x, y) = \sup_{\|P\|_{\infty, [-1,1]} \leq 1} \frac{1}{\deg P} |\cos^{-1}(P(x)) - \cos^{-1}(P(y))|. \quad (2)$$

Thus, the *Dubiner metrics* on the compact $\Omega \subset \mathbb{R}^d$ can be defined as follows. For every $\mathbf{x}, \mathbf{y} \in \Omega$

$$\mu_{\Omega}(\mathbf{x}, \mathbf{y}) = \sup_{\|P\|_{\infty, \Omega} \leq 1} \frac{1}{\deg P} |\cos^{-1}(P(\mathbf{x})) - \cos^{-1}(P(\mathbf{y}))|. \quad (3)$$

In view of the properties of such a metrics⁸, L. Bos stated the following **conjecture**⁹: *nearly-optimal interpolation points on a compact Ω are asymptotically equidistributed w.r.t. the Dubiner metrics on Ω .*

Unfortunately the Dubiner metrics is explicitly known only in very few cases. For $d = 2$, the case we are mostly concerned in the paper, with $\Omega_s = [-1, 1]^2$ the square and $\Omega_c = \{\mathbf{x} : |\mathbf{x}| \leq 1\}$ the circle, we have

- $\mu_{\Omega_s}(\mathbf{x}, \mathbf{y}) = \max\{|\cos^{-1}(x_1) - \cos^{-1}(y_1)|, |\cos^{-1}(x_2) - \cos^{-1}(y_2)|\}$
- $\mu_{\Omega_c}(\mathbf{x}, \mathbf{y}) = \left| \cos^{-1} \left(x_1 x_2 + y_1 y_2 + \sqrt{1 - x_1^2 - y_1^2} \sqrt{1 - x_2^2 - y_2^2} \right) \right|$.

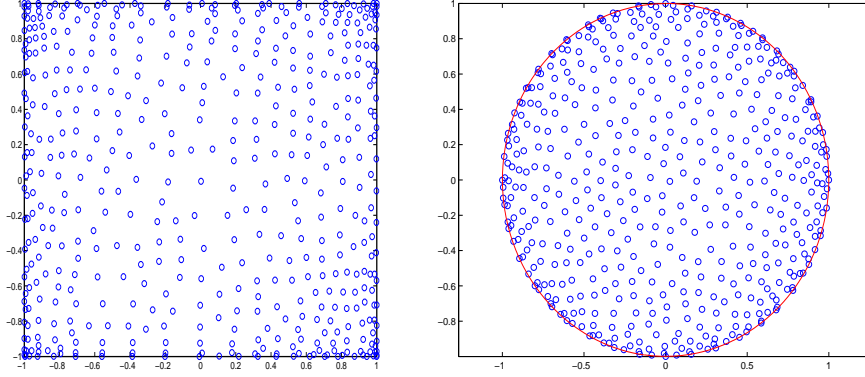


Figure 1. Left: 496 (i.e. degree 30) quasi-uniform Dubiner points for the square; Right: 496 quasi-uniform Dubiner points for the circle.

Looking at Figure 3 below, the growth n^3 for Dubiner points (i.e. $N^{3/2}$) is not satisfactory. Indeed, for Fekete points on the square it is well-known that $\Lambda_n(F_N) \leq N^{-10}$. This suggests that quasi-uniformity in the Dubiner metrics *is not sufficient* for near-optimality of the interpolation points.

1.1. Morrow-Patterson and Padua points

Morrow and Patterson¹² proposed for cubature purposes the following set of points. For n , a positive *even* integer, consider the points $X_N^{MP} = \{(x_m^{MP}, y_k^{MP})\} \subset [-1, 1]^2$ given by

$$x_m^{MP} = \cos\left(\frac{m\pi}{n+2}\right), \quad y_k^{MP} = \begin{cases} \cos\left(\frac{2k\pi}{n+3}\right) & m \text{ odd} \\ \cos\left(\frac{(2k-1)\pi}{n+3}\right) & m \text{ even} \end{cases} \quad (4)$$

$1 \leq m \leq n+1$, $1 \leq k \leq \frac{n}{2} + 1$. These are the Morrow-Patterson points and are exactly equally spaced w.r.t. the Dubiner metrics.

Bos^{9,3} proved that $\Lambda_n^{MP} = O(n^6)$. From our experiments we showed that $\Lambda_n^{MP} = O(n^2)$: moreover Λ_n^{MP} can be least-square fitted with the quadratic polynomial $(0.7n+1)^2$, which is smaller than N (i.e. the theoretical bound for Fekete points).

We constructed also a new set of points that we called *Extended Morrow-Patterson points* (EMP), which correspond to using extended Chebyshev nodes in (4), i.e. $X_N^{EMP} = \{(x_m^{EMP}, y_k^{EMP})\} \subset [-1, 1]^2$, so that

$$x_m^{EMP} = \frac{1}{\alpha_n} x_m^{MP}, \quad y_k^{EMP} = \frac{1}{\beta_n} y_k^{MP}, \quad 1 \leq m \leq n+1, \quad 1 \leq k \leq \frac{n}{2} + 1 \quad (5)$$

where the *dilation coefficients*, α_n and β_n , have expressions $\alpha_n = \cos(\pi/(n+2))$, $\beta_n = \cos(\pi/(n+3))$. The growth of the Lebesgue constant is again quadratic in the degree and numerically we showed $\lambda_n((x_1, x_2); X_N^{EMP}) \not\leq \lambda_n((x_1, x_2); X_N^{MP})$, $\forall (x_1, x_2) \in [-1, 1]^2$, while $\Lambda_n(X_N^{EMP}) < \Lambda_n(X_N^{MP})$.

For n a positive *even* integer consider the points $(x_m, y_k) \in [-1, 1]^2$ given by

$$x_m^{PD} = \cos\left(\frac{(m-1)\pi}{n}\right), \quad y_k^{PD} = \begin{cases} \cos\left(\frac{(2k-2)\pi}{n+1}\right) & m \text{ odd} \\ \cos\left(\frac{(2k-1)\pi}{n+1}\right) & m \text{ even} \end{cases} \quad (6)$$

$1 \leq m \leq n+1$, $1 \leq k \leq \frac{n}{2} + 1$. These are *modified* Morrow-Patterson points that were firstly discussed in Padua and so we have decided to call them *Padua points* (shortly, PD points). They are exactly equispaced w.r.t. the Dubiner metrics on the square.

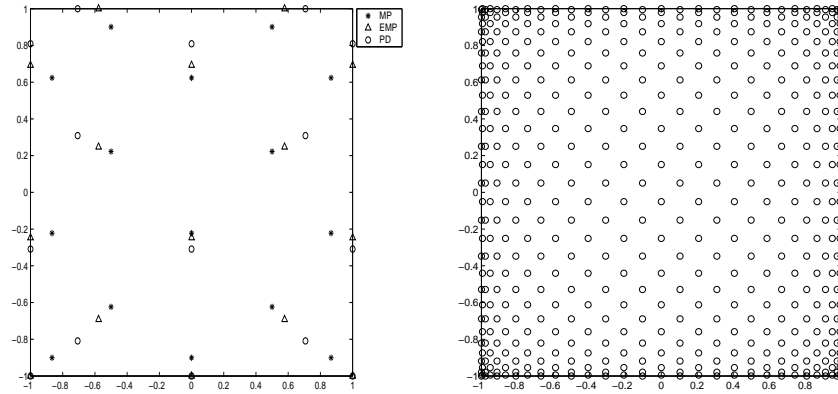


Figure 2. Left: Morrow-Patterson (MP), Extended Morrow-Patterson (EMP) and Padua (PD) points, for degree $n = 4$. Right: Padua points for degree $n = 30$.

In Tables 1-2 we show the interpolation errors in the sup-norm computed approximating two classical functions: the well-known *Franke function* and the function $f(\mathbf{x}) = |\mathbf{x}|$.

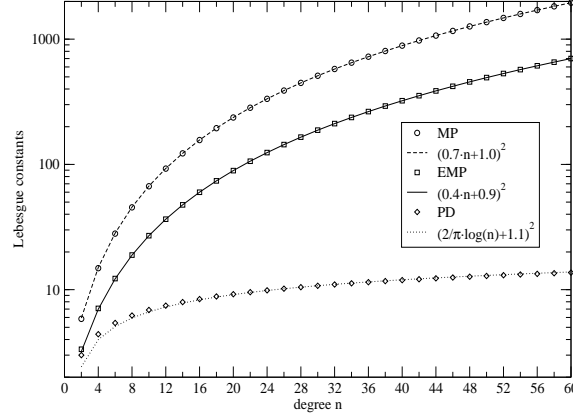


Figure 3. The behavior of the Lebesgue constants for Morrow-Patterson (MP), Extended Morrow-Patterson (EMP), Padua (PD) points up to degree 60, and their least-squares fitting curves.

Table 1. Interpolation errors for the *Franke function* and the corresponding Lebesgue constants for MP, EMP and PD points

	$n = 34$	Λ_{34}	$n = 48$	Λ_{48}	$n = 62$	Λ_{62}	$n = 76$	Λ_{76}
MP	$1.3 \cdot 10^{-3}$	649	$2.6 \cdot 10^{-6}$	1264	$1.1 \cdot 10^{-9}$	2082	$2.0 \cdot 10^{-13}$	3102
EMP	$6.3 \cdot 10^{-4}$	237	$1.3 \cdot 10^{-6}$	456	$5.0 \cdot 10^{-10}$	746	$5.4 \cdot 10^{-14}$	1106
PD	$4.3 \cdot 10^{-5}$	11	$3.3 \cdot 10^{-8}$	13	$5.4 \cdot 10^{-12}$	14	$1.9 \cdot 10^{-14}$	15

Table 2. Interpolation errors for the function $f(\mathbf{x}) = |\mathbf{x}|$.

n	34	48	62	76
MP on $[-1, 1]^2$	$4.4 \cdot 10^{-1}$	$4.4 \cdot 10^{-1}$	$4.4 \cdot 10^{-1}$	$4.4 \cdot 10^{-1}$
MP on $[0, 2]^2$	$8.8 \cdot 10^{-4}$	$2.8 \cdot 10^{-4}$	$2.6 \cdot 10^{-4}$	$1.7 \cdot 10^{-5}$
EMP on $[-1, 1]^2$	$1.4 \cdot 10^{-1}$	$1.4 \cdot 10^{-1}$	$1.4 \cdot 10^{-1}$	$1.4 \cdot 10^{-1}$
EMP on $[0, 2]^2$	$8.3 \cdot 10^{-4}$	$2.6 \cdot 10^{-4}$	$2.1 \cdot 10^{-4}$	$2.1 \cdot 10^{-5}$
PD on $[-1, 1]^2$	$3.7 \cdot 10^{-2}$	$2.7 \cdot 10^{-2}$	$2.1 \cdot 10^{-2}$	$1.7 \cdot 10^{-2}$
PD on $[0, 2]^2$	$7.3 \cdot 10^{-4}$	$3.7 \cdot 10^{-4}$	$7.0 \cdot 10^{-6}$	$4.6 \cdot 10^{-6}$

Some comments are now necessary. First of all, in both examples Padua points reveal to be the most reliable both from the interpolation error point of view that the growth of the Lebesgue constants. Concerning Table 2, it is worth noting that when the singularity of the function $f(\mathbf{x}) = |\mathbf{x}|$ is at the corner of the domain, which is the case when the domain considered is $[0, 2]^2$ instead of $[-1, 1]^2$, then the approximation error is smaller, due to

the fact that all points sets accumulate at the corners of the square.

2. Optimal points for radial basis function interpolation

In this section we assume the reader quite confident with the theory and applications of radial basis functions. For a comprehensive reference we suggest to see the recent book by Buhmann ⁵.

Let $\Omega \subset \mathbb{R}^d$ be a non-empty, bounded and open domain. Let $X = \{x_1, \dots, x_N\} \subseteq \Omega$, N distinct *data sites*, the *point set*. Consider the values $\{f_1, \dots, f_N\}$ *data values*, the interpolation problem with radial basis functions can be formulate as follows. Fix a symmetric function $\Phi : \Omega \times \Omega \rightarrow \mathbb{R}$ and form the interpolant

$$s_{f,X} = \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j) \quad (7)$$

The α_j are determined by the interpolation conditions

$$s_{f,X}(x_j) = f_j, \quad 1 \leq j \leq N.$$

They are unique if the interpolation matrix $A_{\Phi,X} := (\Phi(x_i, x_j))_{1 \leq i,j \leq N}$ is invertible.

If $A_{\Phi,X}$ is *positive definite* $\forall X \subseteq \Omega$, then Φ is called a *positive definite* kernel. It is often *radial* in the sense $\Phi(x, y) = \phi(\|x - y\|_2)$. In what follows we confine to the positive definite case since *every conditionally positive definite kernel has an associated normalized positive definite kernel*¹⁴.

Some useful notations. Take $V_X = \text{span}\{\Phi(\cdot, x) : x \in X\}$. The interpolant $s_{f,X}$ can be written in terms of *cardinal functions* $u_j \in V_X$, $u_j(x_i) = \delta_{ji}$, i.e. $s_{f,X} = \sum_{j=1}^N f(x_j)u_j$. For the purpose of *stability* and *error analysis* the following quantities are important:

- *separation distance*: $q_X = \min_{x_j, x_k \in X, j \neq k} \|x_j - x_k\|_2$;
- *fill-distance*: $h_{X,\Omega} = \sup_{x \in \Omega} \min_{x_j \in X} \|x - x_j\|_2$;
- *uniformity*: $\rho_{X,\Omega} = \frac{q_X}{h_{X,\Omega}}$.

Remark. q_X and $h_{X,\Omega}$ roughly coincide for evenly distributed data sets ⁷.

Problem: are there any “good” or optimal point sets, X^* , for the interpolation problem? Our approach is essentially based on two lines of investigations: *power function estimates* and *geometric arguments*.

2.1. Power function estimates

The kernel Φ defines on the space $V_\Omega = \text{span}\{\Phi(\cdot, x) : x \in \Omega\}$ an inner product so that Φ is a reproducing kernel of V_Ω . Let $\text{clos}(V_\Omega) = N_\Phi(\Omega)$ be the *native Hilbert space*. If $f \in N_\Phi(\Omega)$, then

$$f(x) - s_{f,X}(x) = \left(f, \Phi(\cdot, x) - \sum_{j=1}^N u_j(x) \Phi(\cdot, x_j) \right)_{\Phi}, \quad (8)$$

and by the Cauchy-Schwarz inequality

$$|f(x) - s_{f,X}(x)| \leq P_{\Phi,X}(x) \|f\|_{\Phi}, \quad (9)$$

$P_{\Phi,X}(x)$ is the **power function** which represents the norm of the *point-wise error functional*. There exist error estimates that bound $P_{\Phi,X}(x)$ in terms of the *fill distance* $h_{X,\Omega}$ ¹⁶. Moreover, if $X \subseteq Y$ then $P_{\Phi,X}(x) \geq P_{\Phi,Y}(x)$, $\forall x \in \Omega$. If Φ is translation invariant, integrable and has Fourier transform such that

$$c_\phi(1 + \|\omega\|_2^2)^{-\beta} \leq \hat{\phi}(\omega) \leq C_\phi(1 + \|\omega\|_2^2)^{-\beta}$$

with $\beta > d/2$, $C_\phi \geq c_\phi > 0$, then $n_\Phi(\mathbb{R}^d)$ is norm-equivalent to the space $W_2^\beta(\mathbb{R}^d)$. Therefore,

$$\|f - s_{f,X}\|_{L^\infty(\Omega)} \leq C h_{X,\Omega}^{\beta-d/2} \|f\|_{W_2^\beta(\mathbb{R}^d)},$$

that is we can bound the error by means of the fill-distance.

Theorem 2.1.⁷ *Let Ω closed and bounded in \mathbb{R}^d , satisfying an interior cone condition and Φ has Fourier transform as before. Then, for every $\alpha > \beta$ there exists a constant $M_\alpha > 0$ with the following property: if $\epsilon > 0$ and $X = \{x_1, \dots, x_N\} \subseteq \Omega$ are given such that*

$$\|f - s_{f,X}\|_{L^\infty(\Omega)} \leq \epsilon \|f\|_{\Phi}, \quad \text{for all } f \in W_2^\beta(\mathbb{R}^d), \quad (10)$$

then the fill distance of X satisfies

$$h_{X,\Omega} \leq M_\alpha \epsilon^{\frac{1}{\alpha - \frac{d}{2}}}. \quad (11)$$

Remark. The previous theorem says that optimally distributed data sites are sets that cannot have a large region in Ω without centers, i.e. $h_{X,\Omega}$ is sufficiently small. In the next subsections we outline two methods for computing numerically near-optimal distributed centers for radial basis function interpolation.

2.2. The greedy and the geometric greedy algorithms

The *greedy method* produces well-distributed point sets by generating larger and larger data sets by adding the points where the power function has its maxima w.r.t. preceding set.

- **initial step:** $X_1 = \{x_1\}$, $x_i \in \Omega$, *arbitrary*.
- **iteration step:** $X_j = X_{j-1} \cup \{x_j\}$ with $P_{\Phi, X_{j-1}}(x_j) = \|P_{\Phi, X_{j-1}}\|_{L_\infty(\Omega)}$, $j \geq 2$.

The convergence of the greedy method is guaranteed by the fact that $\|P_j\|_{L_\infty(\Omega)} \leq C j^{-\frac{1}{d}}$ for some $C > 0$. Thus $\lim_{j \rightarrow \infty} \|P_{\Phi, X_j}\|_{L_\infty(\Omega)} = 0$ when Ω convex, $\Phi \in C^2(\Omega \times \Omega)$ or $\Phi \in C^2(\Omega_1 \times \Omega_1)$, $\Omega \subseteq \Omega_1$ convex.

Practical experiments show that the greedy algorithm fills the currently largest hole in the point set, close to the center of the hole. Thus, we can try to see if points suitably extracted from a discretization of Ω can be useful for our purposes. This is the idea behind the *geometric greedy algorithm*.

- **initial step:** $X_0 = \emptyset$ and define $\text{dist}(x, \emptyset) := A$, $A > \text{diam}(\Omega)$.
- **iteration step:** given $X_n \in \Omega$, $|X_n| = n$ pick $x_{n+1} \in \Omega \setminus X_n$ s.t.
 $x_{n+1} = \max_{x \in \Omega \setminus X_n} \text{dist}(x, X_n)$. Then, form $X_{n+1} := X_n \cup \{x_{n+1}\}$.

Remark: this geometric algorithm works very well for subsets X_n of Ω , with *small* fill-distance $h_{X, \Omega}$ and *large* separation distance q_X .

Letting q_n and h_n the separation and fill distances for the sets X_n , the convergence of the geometric greedy algorithms reduces to prove that

$$h_n \geq q_n \geq \frac{1}{2} h_{n-1} \geq \frac{1}{2} h_n, \quad \forall n \geq 2. \quad (12)$$

The proof of (12) can be found in the paper ⁷. Furthermore, one can prove that the geometric greedy algorithm produces sequences of points uniformly distributed in a general metrics ν^3 .

2.3. Examples

Let $\Omega = [-1, 1] \times [-1, 1]$ be discretized by a regular grid of $71 \times 71 = 5041$ points. The radial functions we consider are the gaussian (with scale 1) and the compactly supported Wendland function (with scale 15). The greedy algorithm (shortly, g.a.) has been executed until $\|P_N\|_{L_\infty(\Omega)}^2 \leq \eta$, where η a prescribed threshold. The geometric greedy algorithm (shortly g.g.a.) builds sets X_n and the corresponding errors in the sup-norm are evaluated at these points set.

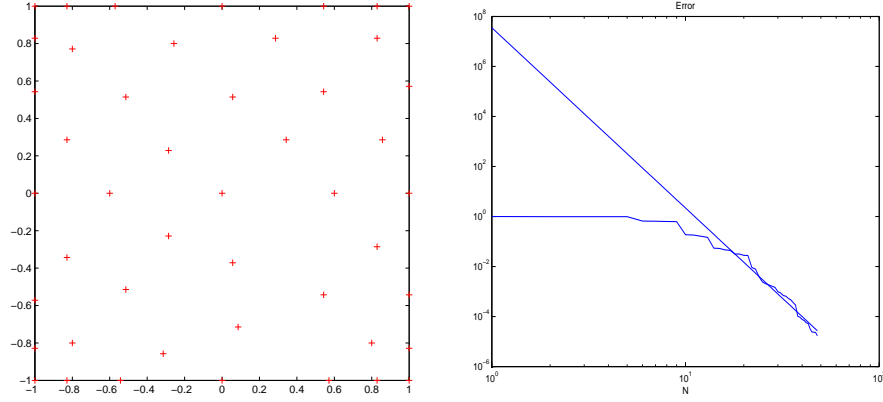


Figure 4. (Left) $N=48$ optimal points when $\eta = 2 \cdot 10^{-5}$; (Right) the error as function of N , decays as $N^{-7.2}$.

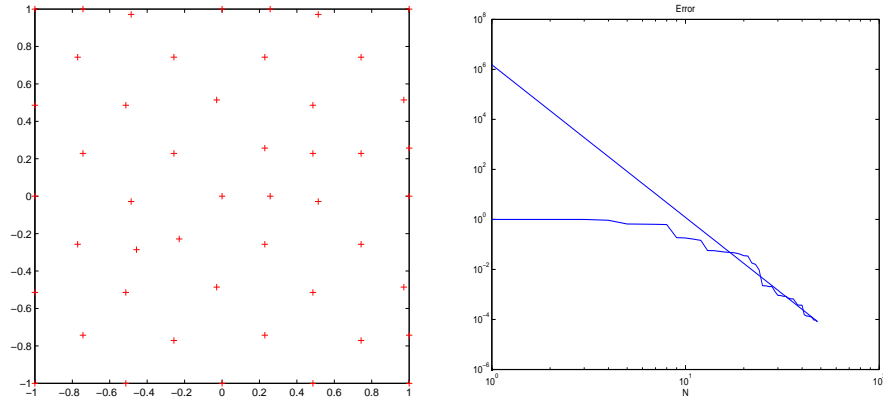


Figure 5. (Left) geometric greedy data X_{48} , (Right) the error is larger by a factor 4 and decays as $n^{-6.1}$.

2.4. Final remarks

- The g.g.a. is *independent* on the kernel and generates asymptotically equidistributed optimal sequences. It still inferior to the g.a. that considers the power function.
- The $(n + 1)$ th point computed by the g.g.a. is such that $h_{X_n, \Omega} = \max_{x \in \Omega} \min_{y \in X_n} \|x - y\|_2$. In the paper³, we proved that they are *quasi-uniform* in the Dubiner metrics and connected to **Leja** sequences⁶.

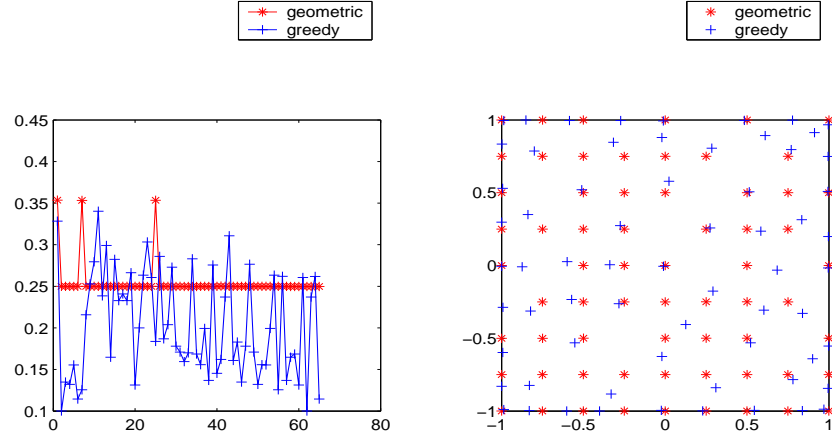


Figure 6. (Right) 65 points as computed by the g.g.a. with $\eta = 2 \cdot 10^{-7}$ (*) and the G.A.(+), (Left) their separation distances.

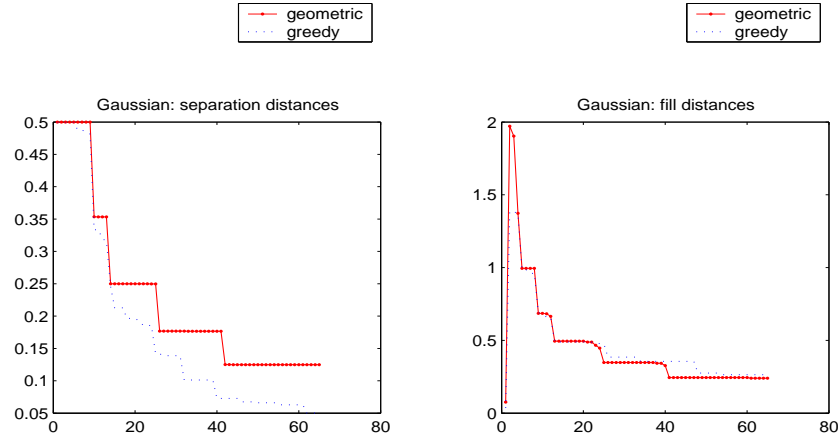


Figure 7. The separation distances q_n and the fill distances h_n .

- So far, we have no proof of the fact the g.g.a. generates a sequence with $h_n \leq Cn^{-1/d}$, as required by asymptotic optimality.

References

1. J.R. Angelos, E.H. Kaufman et. al. in *Approximation Theory VI*, Chui, Schumaker and Ward eds., (1989).
2. L. Brutman, *Ann. Numer. Math.* **4**, (1997).

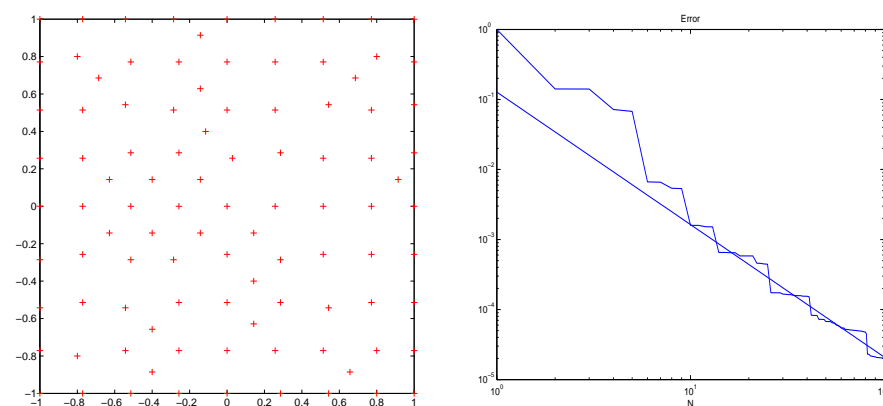


Figure 8. g.a. for the Wendland function. (Left) the $N=100$ optimal points when $\eta = 2 \cdot 10^{-5}$; (Right) the error as function of N that decays as $N^{-1.9}$.

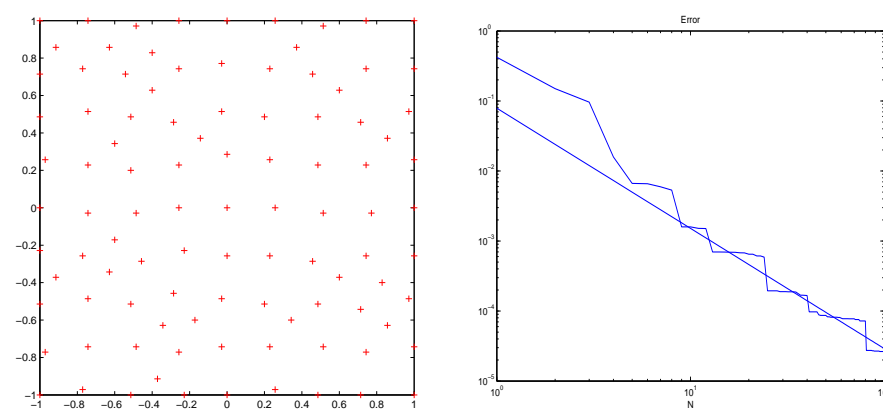


Figure 9. g.g.a. for the Wendland function. (Left) geometric greedy data X_{100} ; (Right) the error is larger by a factor 1.4 and decays as $n^{-1.72}$.

3. M. Caliarì, S. De Marchi and M. Vianello, in press on *Appl. Math. Comput.*, (2004).
4. Q. Chen and I. Babuška, *Comput. Methods Appl. Mech. Engrg.*, **128**, (1995).
5. M. D. Buhmann, *Radial Basis Functions: theory and implementations*, (2003).
6. S. De Marchi, *Appl. Math. Comput.*, **152**, (2004).
7. S. De Marchi, R. Schaback and H. Wendland, in press on *Adv. Comput. Math.*, (2004).
8. M. Dubiner, *J. Anal. Math.*, (1995).
9. L. Bos, *Private communication*, (2002).
10. Bos L, Taylor M. A, Wingate B.A. *Math. Comp.*, **70** (2001).

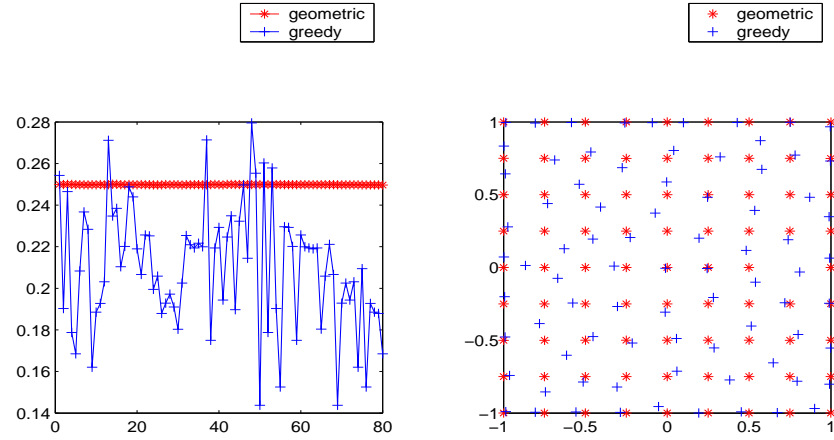


Figure 10. (right) 80 points as computed by the g.g.a. with $\eta = 0.1$ (*) and the g.a. (+), (left) their separation distances.

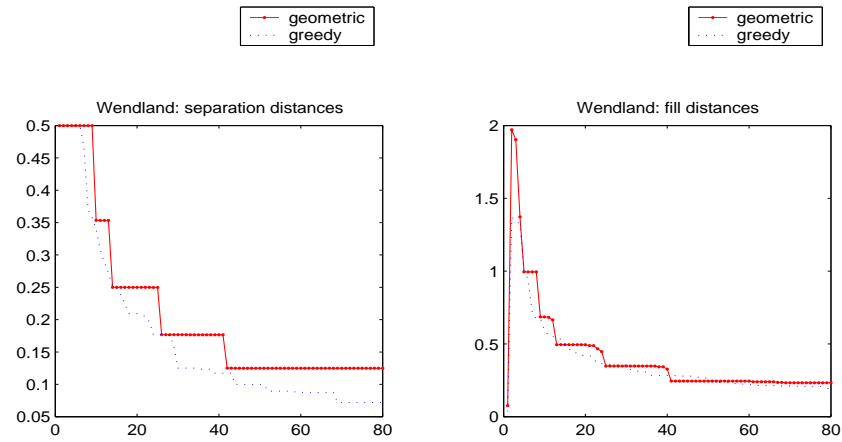


Figure 11. The separation distances q_n and the fill distances h_n .

11. G. Mastroianni and D. Occorsio, *J. Comput. Appl. Math.* **134**(1-2), (2001).
12. C.R. Morrow and T.N.L. Patterson, *SIAM J. Numer. Anal.* **15**(5), (1978).
13. J. G. van der Corput and G. Schaake, *Compositio Math.* **2**, (1935).
14. R. Schaback, *International Series of Numerical Mathematics* **132**, (1999).
15. M. A. Taylor, B. A. Wingate and R. E. Vincent, *SIAM J. Numer. Anal.* **38**(5), (2000).
16. Z. M. Wu and R. Schaback, *IMA J. Numer. Anal* **13**(1), (1993).