Hierarchical Scattered Data Filtering for Multilevel Interpolation Schemes

Armin Iske

Abstract. Multilevel scattered data interpolation requires decomposing the given data into a hierarchy of nested subsets. This paper concerns the efficient construction of such hierarchies. To this end, a recursive filter scheme for scattered data is proposed which generates hierarchies of locally optimal nested subsets. The scheme is a composition of greedy thinning, a recursive point removal strategy, and exchange, a local optimization procedure. The utility of the filter scheme for multilevel interpolation using radial basis functions is shown by numerical examples.

§1. Introduction

Scattered data approximation requires recovering a function $f : \mathbb{R}^d \to \mathbb{R}$, $d \geq 1$, from a given data vector $D_Z(f) = (f(z_1), \ldots, f(z_N))^T \in \mathbb{R}^N$ of function values sampled from f at a finite set $Z = \{z_1, \ldots, z_N\} \subset \mathbb{R}^d$ of locations. Especially when N is extremely large, and the points in Z are unevenly distributed, multilevel interpolation schemes are appropriate techniques. One such scheme was introduced in [4], where compactly supported radial basis functions were used. For a recent survey on radial basis functions, we recommend [1]. The starting point in [4] is a decomposition of the given data into a hierarchy

$$X_L \subset X_{L-1} \subset \cdots \subset X_1 \subset X_0 = Z \tag{1}$$

of $L+1$ nested subsets. As confirmed in Section 4, the performance of the multilevel interpolation scheme [4] heavily depends on the choice of the hierarchy (1). Due to available error estimates for radial basis function interpolation, we wish to keep for each level index $1 \leq j \leq L$ the covering radius

$$r(X_j, X_{j-1}) = \max_{y \in X_{j-1}} d_{X_j}(y)$$

of X_j on X_{j-1} small. Throughout this paper, we use the notation

$$d_X(y) = \min_{x \in X} \|y - x\|$$

Mathematical Methods for Curves and Surfaces: Oslo 2000
Tom Lyche and Larry L. Schumaker (eds.), pp. 211–220.

ISBN 0-8265-1378-6

for the Euclidean distance between a point $y \in \mathbb{R}^d$ and a non-empty set $X \subset \mathbb{R}^d$. This paper concerns the *efficient* construction of hierarchies (1) whose covering radii $r(X_j, X_{j-1})$, $1 \le j \le L$, are small. This shall be accomplished by recursively applying TTX-filters, a family of operators on scattered data sets which output *locally optimal* subsets. The resulting TTX-filtering is a composition of greedy thinning, a recursive point removal scheme [4,5], and exchange, a local optimization procedure. Details are provided in Section 3. The multilevel scattered data interpolation scheme [4] is briefly discussed in Section 2. Finally, the performance of TTX-filtering and its utility for multilevel interpolation is illustrated by the numerical results in Section 4.

§2. Multilevel Interpolation

Let us briefly discuss how to use the hierarchy (1) for multilevel interpolation. According to the scheme proposed in [4], also subject of the papers [3,9], a sequence $s_L, \dots, s_0$ of approximations to f is recursively computed as follows. Let $s_{L+1} \equiv 0$. For $j = L, \dots, 0$, compute an interpolant $\Delta s_j : \mathbb{R}^d \to \mathbb{R}$ to the residual $D_{X_j}(f - s_{j+1})$ on X_j, then let $s_j = s_{j+1} + \Delta s_j$. Altogether, the following $L+1$ interpolation problems are to be solved one after the other:

$$\begin{aligned} D_{X_L}(f) &= D_{X_L}(\Delta s_L) \\ D_{X_{L-1}}(f - s_L) &= D_{X_{L-1}}(\Delta s_{L-1}) \\ &\cdots \\ D_{X_1}(f - s_2) &= D_{X_1}(\Delta s_1) \\ D_{X_0}(f - s_1) &= D_{X_0}(\Delta s_0). \end{aligned} \tag{2}$$

Note that each function s_j matches f at the subset X_j, i.e. $D_{X_j}(s_j) = D_{X_j}(f)$, $0 \le j \le L$. The basic idea behind this multilevel scheme is to capture a global trend of the function f by its coarse approximation s_L, before finer details of f are gradually added by using the sequence of functions Δs_j. Thus, the interpolants $s_L, \dots, s_0$ provide a sequence of representations for f at $L+1$ different resolutions.

In [4], it was suggested to use compactly supported radial basis functions. For a fixed positive definite radial function $\phi : [0, \infty) \to \mathbb{R}$ with support $\mathrm{supp}\,(\phi) = [0,1]$, let $\phi_\rho(\cdot) = \phi(\cdot/\rho)$ for $\rho > 0$, so that $\mathrm{supp}\,(\phi_\rho) = [0, \rho]$. Each of the interpolants Δs_j, $0 \le j < L$, in (2) is then supposed to be of the form

$$\Delta s_j = \sum_{x \in X_j} c_x \phi_{\rho_j}(\| \cdot - x \|),$$

where the support radii ρ_j are monotonically decreasing: $\rho_{L-1} > \cdots > \rho_0 > 0$. Further details can be found in [4]. For the initial step $j = L$, however, we recommend to use the globally supported thin plate splines, $\phi(r) = r^2 \log(r)$, according to whose interpolation scheme the function Δs_L in (2) is of the form

$$\Delta s_L = \sum_{x \in X_L} c_x \| \cdot - x \|^2 \log(\| \cdot - x \|) + p,$$

where $p : \mathbb{R}^d \to \mathbb{R}$ denotes a linear polynomial. In contrast to interpolation by compactly supported radial basis functions, the thin plate spline method reproduces linear polynomials. This carries over to the above multilevel interpolation scheme. Indeed, if f is linear, then $s_L \equiv f$. In this case, every subsequent residual $f - s_j$ vanishes identically, which implies $\Delta s_j \equiv 0$ and thus $s_j \equiv f$ for $0 \leq j \leq L$.

We remark that for the stability of the multilevel interpolation, we require the separation distance $q(X) = \min_{x \in X} d_{X \setminus x}(x)$ in each subset $X = X_j$, $1 \leq j \leq L$, to be not too small. This is because the condition number of a collocation matrix arising in (2) blows up, whenever the separation distance of the corresponding point set is too small [10]. We shall come back to this important point in the following section.

Finally, note that the multilevel scheme (2) would allow us to replace *interpolation* with *least squares approximation* in each of its steps. This makes especially sense in situations, where there is liable to be noise in the data. However, as confirmed in our numerical examples, this leads to significantly higher computational costs (see the details in [8] concerning the required computations for least squares approximation by using radial basis functions).

§3. Hierarchical Scattered Data Filtering

In this section, we shall introduce thinning and exchange, our two main ingredients for scattered data filtering. But let us first fix some notations.

Definition 1. *Let $\mathcal{X}_Z = \{X : X \subset Z, X \neq \emptyset\}$ denote the power set of the $2^N - 1$ non-empty subsets of Z. We say that an operator $F : \mathcal{X}_Z \to \mathcal{X}_Z$ is a* filter *on Z, iff it satisfies $F(X) \subset X$ and $|F(X)| < |X|$ for all $X \in \mathcal{X}_Z$. A filter F is said to be a* thinning operator*, iff $|F(X)| = |X| - 1$ for all $X \in \mathcal{X}_Z$.*

The idea for the construction of the hierarchy (1) is to recursively apply a sequence $F_1, \ldots, F_L$ of TTX-filters, so that $X_j = F_j(X_{j-1})$, $1 \leq j \leq L$. The form of these particular filters shall be specified at the end of this section. For this purpose, it is sufficient to restrict ourselves to the special case of one sublevel, i.e. $L = 1$ in (1). Therefore, in this section we discuss how to extract merely *one* subset $X \equiv X_1$ from Z, where $|X| \ll |Z|$.

3.1. Thinning

Recall that we wish to select one subset $X \in \mathcal{X}_Z$, $|X| \ll |Z|$, whose covering radius $r(X, Z)$ on Z is small. This shall be accomplished by recursive point removals from Z. To this end, let us fix one thinning operator T, and let T^n denote its n-fold composition. Then, by

$$T^n(Z) \subset T^{n-1}(Z) \subset \cdots \subset T(Z) \subset T^0(Z) = Z$$

we obtain a nested sequence of subsets, like in (1), but with a much finer granulation. Now, the idea is to select one particular subset $X = T^k(Z) \in \mathcal{X}_Z$ by picking one suitable breakpoint $k \gg 0$ at run time (i.e. during the removal).

Ideally, we wish to select one subset X^* of size $|Z| - k$ which is optimal by minimizing its covering radius on Z among all subsets $X \in \mathcal{X}_Z$ of equal size, i.e.

$$r_k^*(Z) = r(X^*, Z) = \min_{\substack{X \in \mathcal{X}_Z \\ |X| = |Z| - k}} r(X, Z).$$

The problem of finding for a fixed $\alpha \geq 1$ an algorithm which outputs for any given finite set Z and any k, $1 \leq k < |Z|$, a subset $X \in \mathcal{X}_Z$ of size $|X| = |Z| - k$ satisfying

$$r(X, Z) \leq \alpha \cdot r_k^*(Z) \tag{3}$$

is one variant of the k-center problem (for more details see [6], Section 9.4.1). However, for any $\alpha < \alpha^* = \sqrt{2 + \sqrt{3}}$ there is no α-approximation algorithm for the k-center problem, unless P=NP (cf. [11], Section 4). There are a few approximation algorithms available for $\alpha = 2$, the first one dating back to [7], but for $\alpha^* \leq \alpha < 2$ no such algorithm is known.

Note that the universal constant α in (3) depends neither on Z nor on X. The following proposition provides one dynamic bound, of a similar form to (3), which allows us to pick several subsets $X \in \mathcal{X}_Z$ at run time whose covering radii $r(X, Z)$ approach the optimal value up to a factor $\alpha \equiv \alpha(X, Z) \ll \alpha^*$. To this end, assume without loss of generality that the points in $Z = \{z_1, \ldots, z_N\}$ are ordered such that their *significances* $\sigma_k = d_{Z \setminus z_k}(z_k)$ are monotonically increasing, i.e. $\sigma_k \leq \sigma_{k+1}$, $1 \leq k < N$.

Proposition 2. *Let $X \in \mathcal{X}_Z$ be of size $|X| = |Z| - k$. Then,*

$$\sigma_k \leq r_k^*(Z) \leq r(X, Z) \leq \alpha(X, Z) \cdot r_k^*(Z), \tag{4}$$

where $\alpha(X, Z) = r(X, Z)/\sigma_k$.

Proof: We shall prove the inequality $r(X, Z) \geq \sigma_k$ for all $X \in \mathcal{X}_Z$ of size $|X| = |Z| - k$. This implies $r_k^*(Z) \geq \sigma_k$, and moreover

$$r(X, Z) = \alpha(X, Z) \cdot \sigma_k \leq \alpha(X, Z) \cdot r_k^*(Z).$$

To this end, let $Y = Z \setminus X$. Since $d_X(y) = d_{Z \setminus Y}(y) \geq d_{Z \setminus y}(y)$ for all $y \in Y$, using $|Y| = k$ we conclude

$$r(X, Z) = \max_{z \in Z} d_X(z) = \max_{y \in Y} d_X(y) \geq \max_{y \in Y} d_{Z \setminus y}(y) \geq \sigma_k,$$

which completes our proof. $\square$

We shall record the quality indices $\alpha(X, Z)$ during the point removal at run time. The above bound (4) allows us to control for any current subset $X \in \mathcal{X}_Z$, $|X| = |Z| - k$, the relative error by

$$\left| \frac{r(X, Z) - r_k^*(Z)}{r_k^*(Z)} \right| \leq \alpha(X, Z) - 1,$$

and this provides a useful criterion for the breakpoint selection.

NP-complete optimization problems, like the k-center problem, can effectively be tackled by using greedy approximation algorithms, algorithms of polynomial complexity which make an optimal decision at each single step. This motivates the usage of greedy thinning operators.

Definition 3. *A thinning operator $T_* : \mathcal{X}_Z \to \mathcal{X}_Z$ is said to be a* greedy thinning operator *iff T_* minimizes for any $X \in \mathcal{X}_Z$ the covering radius $r(T(X), X)$ among all thinning operators T.*

The following observation provides a useful characterization of this operator class.

Proposition 4. *A thinning operator T_* is greedy iff for any $X \in \mathcal{X}_Z$, the action of T_* on X removes one point $x^* \in X$ from X which satisfies*

$$\sigma_X(x^*) = \min_{x \in X} \sigma_X(x), \tag{5}$$

where $\sigma_X(x) = d_{X \setminus x}(x)$ denotes the significance of $x \in X$ in X.

Proof: Let T be any thinning operator whose action on $X \in \mathcal{X}_Z$ is given by $T(X) = X \setminus x_T$ for a specific $x_T \in X$. Then,

$$r(T(X), X) = \max_{x \in X} d_{T(X)}(x) = d_{X \setminus x_T}(x_T) = \sigma_X(x_T).$$

Hence, T minimizes $r(T(X), X)$ iff x_T minimizes $\sigma_X(x)$ among all $x \in X$. □

The above proposition yields an additional argument in favour of using greedy thinning operators: Recall that we wish to have the separation distances $q(X)$ not too small. Since $q(X) = \sigma_X(x^*)$, where x^* in (5) is a least significant point in X, we see that a greedy thinning operator serves to eliminate closest point pairs.

3.2. Exchange

In order to balance the short-sightedness of greedy thinning, we offer one additional useful ingredient for the construction of suitable filters. The idea is, for any given $X \in \mathcal{X}_Z$, to exchange a point pair between X and the complementary set $Y = Z \setminus X$, whenever the exchange pays off in terms of the reduction of the covering radius $r(X, Z)$. This leads us to the following

Definition 5. *We say that a point pair $(x, y) \in X \times Y$ is* exchangeable *iff it satisfies*

$$r(X, Z) > r((X \setminus x) \cup y, Z).$$

A subset $X \in \mathcal{X}_Z$ is said to be locally optimal *in Z iff there is no exchangeable point pair $(x, y) \in X \times Y$. By $\mathcal{X}_Z^*$ we shall denote the set of all subsets from $\mathcal{X}_Z$ which are locally optimal in Z.*

We shall interpret each exchange as a non-trivial operation of an exchange operator $E : \mathcal{X}_Z \to \mathcal{X}_Z$ on a subset $X \in \mathcal{X}_Z$:

$$E(X) = \begin{cases} X, & \text{iff } X \text{ is locally optimal,} \\ (X \setminus x) \cup y, & \text{for an exchangeable pair } (x, y) \in X \times Y. \end{cases}$$

Let E^n be the n-fold composition of E, and E^0 the identity on $\mathcal{X}_Z$. Note that for every $X \in \mathcal{X}_Z \setminus \mathcal{X}_Z^*$, there is one unique positive index $m = m(X) < \infty$, such that

$$E^{m+1}(X) = E^m(X) \neq E^{m-1}(X).$$

In other words, the exchange process terminates after finitely many steps, in which case the subset $E^\infty(X) = E^m(X)$ is locally optimal. Indeed, this is true since Z was assumed to be finite, and in each exchange the (non-negative) covering radius is strictly reduced:

$$0 \leq r(E^m(X), Z) < r(E^{m-1}(X), Z) < \ldots < r(E(X), Z) < r(X, Z).$$

Therefore, the operator $E^\infty : \mathcal{X}_Z \to \mathcal{X}_Z^*$ is a projector onto the locally optimal subsets of Z, where we let $m(X) = 0$ for $X \in \mathcal{X}_Z^*$.

The remainder of this subsection is devoted to the characterization of exchangeable point pairs. The following theorem provides one sufficient criterion which is useful for the efficient localization of such pairs.

Theorem 6. *Let $y^* \in Y$ satisfy $d_X(y^*) > d_X(y)$ for all $y \in Y \setminus y^*$. Then, for $x \in X$ the pair $(x, y^*) \in X \times Y$ is exchangeable if it satisfies the following two conditions.*

1) $r(X, Z) > \sigma_X(x)$;
2) $r(X, Z) > d_{X \setminus x}(y)$ *for all* $y \in (Y \setminus y^*) \cap V_X(x)$.

In the above condition 2), $V_X(x) = \{z \in \mathbb{R}^d : d_X(z) = \|z - x\|\}$ stands for the Voronoi tile of $x \in X$ w.r.t. X.

Proof: Note that $r(X, Z) = d_X(y^*)$, which together with 2) implies

$$r(X, Z) > d_{X \setminus x}(y) \qquad \text{for all } y \in Y \setminus y^*. \tag{6}$$

Indeed, condition 2) covers (6) for all $y \in (Y \setminus y^*) \cap V_X(x)$, and by using

$$r(X, Z) > d_X(y) = d_{X \setminus x}(y) \qquad \text{for all } y \in Y \setminus y^* \text{ with } y \notin V_X(x),$$

we see that (6) holds. Now, (6) in combination with condition 1) implies

$$r(X, Z) > \max\left(\max_{y \in Y \setminus y^*} d_{X \setminus x}(y), d_{X \setminus x}(x)\right),$$

and therefore

$$r(X, Z) > \max_{y \in (Y \setminus y^*) \cup x} d_{X \setminus x}(y) \geq \max_{y \in (Y \setminus y^*) \cup x} d_{(X \setminus x) \cup y^*}(y) = r((X \setminus x) \cup y^*, Z)$$

which completes our proof. $\square$

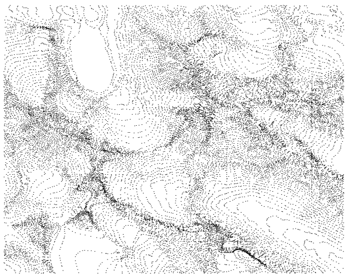
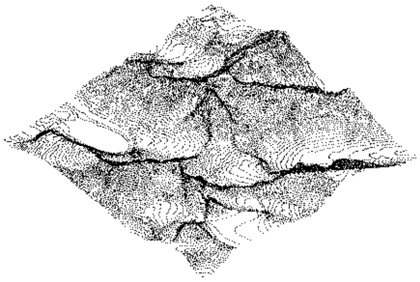

Fig. 1 and 2. The data set Hurrungane (23092 points), 2D view and 3D view.

3.3. Scattered Data Filtering

Let T_* be a fixed greedy thinning operator and E a fixed exchange operator. For the purpose of selecting one subset $X \in \mathcal{X}_Z$ with a small covering radius $r(X, Z)$, we shall use TTX-filters, filters of the form $F_n = E^\infty \circ T_*^n$. Mnemonically, the acronym TTX stands for "**T**hinning, ..., **T**hinning, E**X**change". These filters generate a (not necessarily nested) sequence $\{F_n(Z)\}_n$ of locally optimal subsets of decreasing size $|F_n(Z)| = |Z| - n$. In our numerical examples, we have used their quality indices $\alpha(F_n(Z), Z)$ for picking one subset $X = F_n(Z) \in \mathcal{X}_Z^*$ whose covering radius $r(X, Z)$ is close to the optimum $r_n^*(Z)$. We finally remark that an efficient implementation of F_n can be accomplished in two dimensions by using Delaunay triangulations and heaps, as suggested in [5] for greedy thinning.

§4. Numerical Results

We have implemented TTX-filtering for bivariate data, $d = 2$, by using Delaunay triangulations and heaps, with the exchange being based on the sufficient criterion in Theorem 6. In our numerical examples, we considered using a data set called *Hurrungane*, a sample of height values $\{f(z)\}_{z \in Z}$ of a mountain area at $|Z| = 23092$ distinct locations in the domain $\Omega = [437000, 442000] \times [6812000, 6817000]$ with minimum height $\min_{z \in Z} f(z) = 1100$ and maximum height $\max_{z \in Z} f(z) = 2400$. The data set is displayed in the Figures 1,2.

4.1. Non-hierarchical Filtering

Let us briefly illustrate how the performance of the TTX-filtering scheme compares with approximation algorithms for the k-center problem. For this purpose, we have computed the sequence $\{F_n(Z)\}_n$ of subsets. Figure 3 shows the graph of their covering radii $r(F_n(Z), Z)$, $n \in [300, 20000]$, along with that of the initial significances σ_n. Figure 4 displays the graph of the quality indices $\alpha(F_n(Z), Z)$. We find $\alpha(F_n(Z), Z) < \alpha^*$ for $n = 300, \ldots, 18599$, and

$$\min_{500 \le n \le 20000} \alpha(F_n(Z), Z) = \alpha(F_{2976}(Z), Z) = 1.1851.$$

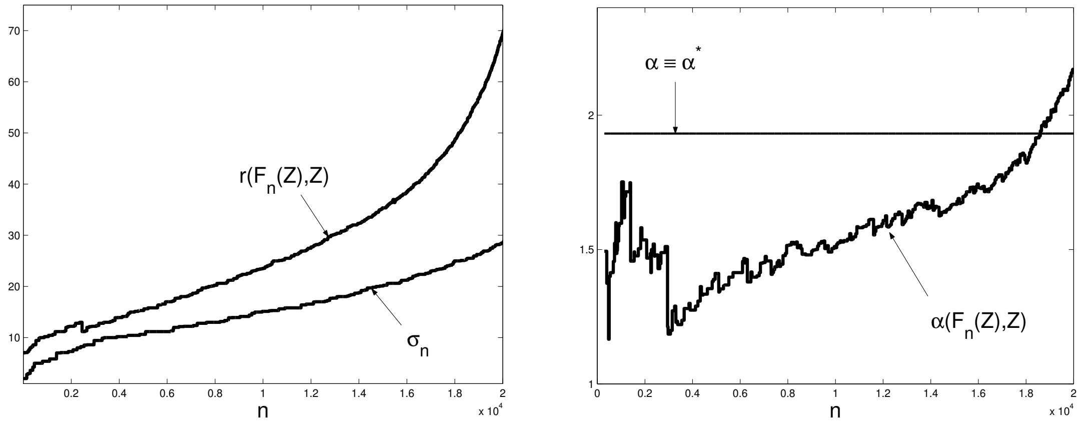

Fig. 3 and 4. The graphs σ_n, $r(F_n(Z), Z)$ and $\alpha(F_n(Z), Z)$, $n \in [300, 20000]$.

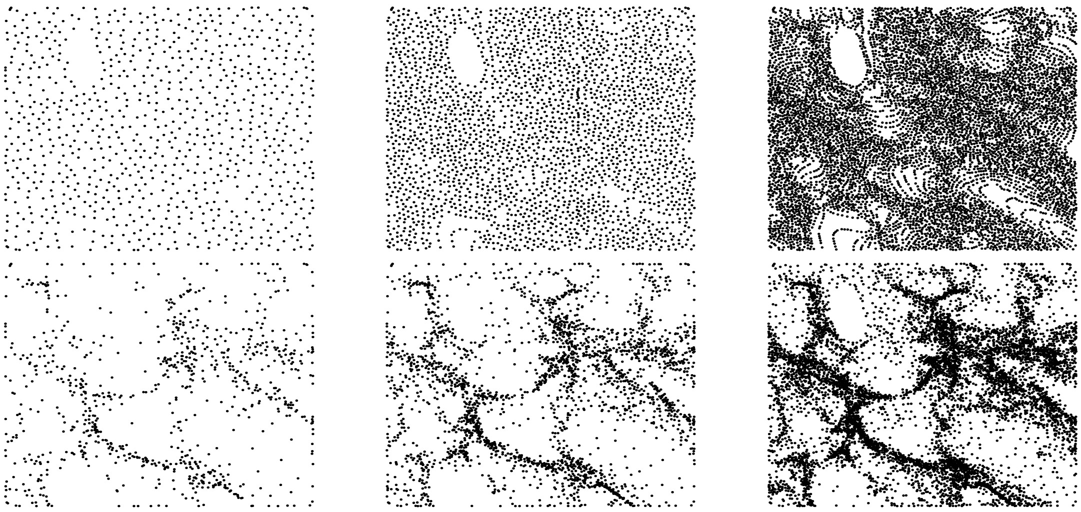

Fig. 5. The sets X_3^F, X_2^F, X_1^F (first row) and X_3^A, X_2^A, X_1^A (second row).

4.2. Hierarchical Filtering and Multilevel Interpolation

We have decomposed the given superset Z into four nested subsets, i.e. $L = 3$ in (1), by recursively using TTX-filters. The three subsets $X_3^F \subset X_2^F \subset X_1^F$ of sizes $|X_3^F| = 855$, $|X_2^F| = 2566$ and $|X_1^F| = 7697$ are displayed in the Figure 5, first row. Observe that the points of these subsets are uniformly distributed in the domain Ω.

In contrast, the second row of Figure 5 displays a nested sequence $X_3^A \subset X_2^A \subset X_1^A$ of non-uniformly distributed subsets of the same sizes, i.e. $|X_3^A| = 855$, $|X_2^A| = 2566$ and $|X_1^A| = 7697$. This sequence has been generated by using adaptive thinning [2], a recent method for selecting subsets from bivariate scattered data whose piecewise linear interpolants over their Delaunay triangulations are close to the original data. This particular thinning method prefers to remove data points in flat regions of the surface, while it keeps data points where the surface's curvature is high. Adaptive thinning is essentially

j	$r_j(F)$	$r_j(A)$	$\alpha_j(F)$	$\alpha_j(A)$	$\delta_j(F)$	$\delta_j(A)$	u(F)	u(A)
1	34.41	252.53	1.67	12.26	185.24	6.14	19.80	195.91
2	65.49	372.83	1.53	13.12	356.66	31.42	3.68	52.94
3	117.44	463.01	1.47	9.64	424.71	69.26	0.78	15.00

Tab. 1. Performance of TTX-filtering (F) and adaptive thinning (A).

j	$\eta_\infty(F)$	$\eta_\infty(A)$	$\eta_2(F)$	$\eta_2(A)$	u(F)	u(A)
3	169.63	541.51	18.44	35.87	4.81	4.84
2	126.33	342.39	10.70	18.91	1.76	30.47
1	83.81	222.05	7.97	12.60	10.83	247.78

Tab. 2. Multilevel interpolation: approximation quality and computational costs.

a data-dependent method, i.e. it depends on the function values of f at Z, whereas TTX-filtering is data-independent, because it merely depends on the data set Z.

The computational costs, in CPU units time, required for the subset selection at each level are, along with the subsets' covering radii $r_j = r(X_j, X_{j-1})$, and the quality indices $\alpha_j = \alpha(X_j, X_{j-1})$, $j = 1, 2, 3$, documented in Table 1, where 'F' stands for TTX-filtering, and 'A' for adaptive thinning. In addition, Table 1 shows the values $\delta_j = \max_{z \in Z} |L(f, X_j)(z) - f(z)|$ of the maximal deviations between f and $L(f, X_j)$ at Z, where $L(f, X_j)$ is the piecewise linear interpolant over the Delaunay triangulation of X_j satisfying $D_{X_j}(L(f, X_j)) = D_{X_j}(f)$.

Note that the covering radii $r_j(F)$ and quality indices $\alpha_j(F)$ of the subsets X_j^F are much smaller than the corresponding values $r_j(A)$ and $\alpha_j(A)$, even at significantly lower computational costs. On the other hand, however, the piecewise linear interpolants $L(f, X_j^A)$ are compared with $L(f, X_j^F)$ much closer to the given function values of f at Z, i.e. $\delta_j(A) \ll \delta_j(F)$, $j = 1, 2, 3$. The latter is the goal of adaptive thinning in [2], and therefore we refrain from expanding further details here.

Now let us turn to multilevel interpolation. Often has it been suggested to work with a data-dependent method, like adaptive thinning, instead of a data-independent one, like TTX-filtering, when it comes to the construction of the hierarchy (1). If the piecewise linear interpolant $L(f, X_j)$ was close to the function f, then the interpolant s_j in (2) would be close to f, i.e. $f \approx L(f, X_j) \approx s_j$, so the expectation.

The comparison between the two different sequences X_1^F, X_2^F, X_3^F and X_1^A, X_2^A, X_3^A shows, however, that the opposite can be true. For the purpose of illustration, we considered using the compactly supported radial basis function $\phi(r) = (1-r)_+^4(4r+1)$ [12], and we let $\rho_2 = 327.45$, $\rho_1 = 177.05$ for the support radii at the levels $j = 2, 1$ in (2). Table 2 reflects our numerical results, where

$$\eta_\infty = \max_{x \in X_{j-1}} |f(x) - s_j(x)|, \qquad \eta_2^2 = \frac{1}{|X_{j-1}|} \sum_{x \in X_{j-1}} |f(x) - s_j(x)|^2.$$

Already at the initial level $L = 3$, the thin plate spline interpolant $s_3(F)$ of f at X_3^F yields much smaller residual norms $\eta_\infty(F), \eta_2(F)$ than its counterpart $s_3(A)$. Moreover, the poor approximation quality of $s_3(A)$ cannot be recovered by the subsequent interpolants $s_2(A), s_1(A)$, and the computational costs u(A) blow up in the remaining steps of (2). Observe that the method 'A' looses dramatically against the method 'F' at both ends, approximation quality and computational costs. In fact, we haven't encountered any numerical example where this is different. In summary, the choice of the hierarchy (1) can significantly affect the performance of the multilevel interpolation scheme (2), where special attention is to be paid to the approximation quality of s_L.

Finally, coming back to one comment in Section 2, we remark that working with least squares approximation rather than interpolation does dramatically increase the computational costs. Indeed, if we e.g. replace $s_3(F)$ by the corresponding least squares approximation, this would reduce the least squares error by $\eta_2 = 15.31$, but at $u = 53.28$ CPU units time.

Acknowledgments. The author was partly supported by the European Union within the project MINGLE (Multiresolution in Geometric Modelling), contract no. HPRN-CT-1999-00117.

References

1. Buhmann, M. D., Radial basis functions, Acta Numerica, (2000), 1–38.
2. Dyn, N., M. S. Floater, and A. Iske, Adaptive thinning for bivariate scattered data, preprint, Technische Universität München, 2000.
3. Fasshauer, G. E. and J. W. Jerome, Multistep approximation algorithms: Improved convergence rates through postconditioning with smoothing kernels, Advances in Comp. Math. **10** (1999), 1–27.
4. Floater, M. S. and A. Iske, Multistep scattered data interpolation using compactly supported radial basis functions, J. Comput. Appl. Math. **73** (1996), 65–78.
5. Floater, M. S. and A. Iske, Thinning algorithms for scattered data interpolation, BIT **38** (1998), 705–720.
6. Hochbaum, D. S. (ed.), *Approximation Algorithms for NP-hard Problems*, PWS Publishing Company, Boston, 1997.
7. Hochbaum, D. S. and D. B. Shmoys, A best possible heuristic for the k-center problem, Mathematics of Operations Research **10:2** (1985), 180–184.
8. Iske, A., Reconstruction of smooth signals from irregular samples by using radial basis function approximation, in *Proceedings of the 1999 International Workshop on Sampling Theory and Applications*, Y. Lyubarskii (ed.), The Norwegian University of Science and Technology, Trondheim, 1999, 82–87.

9. Narcowich, F. J., R. Schaback, and J. D. Ward, Multilevel interpolation and approximation, Appl. Comput. Harmonic Anal. **7** (1999), 243–261.
10. Narcowich, F. J. and J. D. Ward, Norms of inverses and condition numbers for matrices associated with scattered data, J. Approx. Theory **64** (1991), 69–94.
11. Shmoys, D. B., Computing near-optimal solutions to combinatorial optimization problems, DIMACS, Ser. Discrete Math. Theor. Comput. Sci. **20** (1995), 355–397.
12. Wendland, H., Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree, Advances in Comp. Math. **4** (1995), 389–396.

Armin Iske
Zentrum Mathematik
Technische Universität München
D-80290 München, Germany
iske@ma.tum.de