

THB-spline approximations for turbine blade design with local B-spline approximations

Cesare Bracco, Carlotta Giannelli, David Großmann, Sofia Imperatore,
Dominik Mokriš, Alessandra Sestini

Abstract We consider two-stage scattered data fitting with truncated hierarchical B-splines (THB-splines) for the adaptive reconstruction of industrial models. The first stage of the scheme is devoted to the computation of local least squares variational spline approximations, exploiting a simple fairness functional to handle data distributions with a locally varying density of points. Hierarchical spline quasi-interpolation based on THB-splines is considered in the second stage of the method to construct the adaptive spline surface approximating the whole scattered data set and a suitable strategy to guide the adaptive refinement is introduced. A selection of examples on geometric models representing components of aircraft turbine blades highlights the performances of the scheme. The tests include a scattered data set with voids and the adaptive reconstruction of a cylinder-like surface.

1 Introduction

Scattered data fitting is nowadays of fundamental importance in a variety of fields, ranging from geographic applications to medical imaging and geometric modeling. The topic can be addressed in different approximation spaces, either by using spline spaces or radial basis functions. In particular, thanks to their low computational cost and also to their better control of conditioning, schemes relying on local approximations have received a lot of attention, being formulated either as partition of unity

Cesare Bracco, Carlotta Giannelli, Sofia Imperatore, Alessandra Sestini
Dipartimento di Matematica e Informatica “U. Dini”, Università degli Studi di Firenze,
Viale Morgagni 67/A, 50134 Firenze, Italy e-mail: {cesare.bracco, carlotta.giannelli, alessandra.sestini}@unifi.it, sofia.imperatore@stud.unifi.it

David Großmann, Dominik Mokriš
MTU Aero Engines AG, Munich, Germany e-mail: {David.GROSSMANN, Dominik.Mokris}@mtu.de

interpolation or as two-stage approximation methods, see for example [4, 6, 7, 8] and references therein.

In this work we are interested in two-stage surface reconstruction of industrial models starting from scattered data obtained by optical scanning acquisitions. In order to reduce the noise, the available data are always preprocessed because the interest is in highly accurate reconstructions. Anyway, since the available data can not be considered exact, interpolation is not required and a less costly approximation scheme can be used to compute any local fitting. Furthermore, considering that splines represented in B-spline form are the standard choice for industrial computer-aided design applications (see e.g. [22]), we are interested in determining the final approximation in a space spanned by suitable extensions of tensor-product B-splines.

As well known, an approximating spline can be obtained by using different approximation approaches and also operating in different kinds of spline spaces, with the related schemes usually divided into two main classes. The first is the class of *global* schemes which use simultaneously all the given information and thus require the solution of a linear system whose size is (substantially) equal to the cardinality of the considered spline space. The second class collects all kinds of *local* schemes which avoid the solution of a global linear system. Intermediate alternatives are also possible, see e.g., [17], where spline spaces on triangulations are considered together with a domain decomposition technique. Many local methods can be collocated in the field of Quasi-Interpolation (QI) which is a fundamental methodology in the context of spline approximation, see e.g. [18, 19] for an introduction. In some cases, e.g. when dealing with scattered data, QI can require the solution of small linear systems whose size does not depend on the cardinality of the entire data set or on the size of the global spline space. These local systems descend from the application of *local* approximation schemes, each one considering a small number of data whose associated parameter values belong to a modest portion of the parametric domain intersecting the support of a compactly supported basis function. The computational advantage of the QI approach is evident, since these local linear systems are of small size and independent of each other. However, the development of an effective scattered data fitting approach based on local methods is never trivial, since the quality of the final spline approximation depends in this case not only on the considered global spline space but also on the approximation power of the local scheme and on the choice of the corresponding local data set. Here we are interested in considering a method of this kind but having also the adaptivity feature, in particular relying on a QI two stage methodology and working in adaptive spline spaces, since they allow local refinement and generalize the standard tensor-product model. Since scattered data can be characterized by a highly varying distribution, by also including voids, a flexible and reliable approach which automatically (re-) constructs the geometric model may strongly improve the efficiency of the overall scheme by suitably adapting the solution to the shape and configuration of the given point clouds.

In recent years the need for local refinement both in computer aided geometric design and isogeometric analysis has motivated a lot of research on adaptive spline spaces, leading to the introduction of several adaptive spline constructions, such as spline spaces over T-meshes [9], T-splines [23], LR B-splines [10] and hierarchical

B-splines [16]. Spline spaces over T-meshes is the most straightforward approach: given a T-mesh, that is a rectangular mesh allowing T-junctions (vertices with valence 3), a space of functions which are polynomial over each cell can be easily defined. While this is very natural, it has been shown that the dimension of the corresponding space is stable (namely, it depends only on the topology of the mesh and not on the geometry) only under certain conditions, see [20] and [3]. Consequently, an efficient construction of a global basis is, in general, still an open question. T-splines [23] and LR B-splines [10] are both formally B-splines, but defined on local knot vectors depending on the topology of the mesh, and therefore allowing local refinement. They are very flexible, but their linear independence is guaranteed only under certain conditions on the mesh (different for the two types of spline), which any refinement algorithm must then preserve, see e.g. [1] and [21]. The construction of hierarchical B-splines [16] is based on a multi-level construction, where the refinement is obtained by replacing coarser elements with finer ones in the mesh hierarchy. As a consequence, the corresponding basis is composed of B-splines constructed on meshes of different levels. The multilevel nature of these splines allows the design of efficient local refinement schemes and their suitable integration into existing computer aided design software [15]. Moreover, the truncated version of hierarchical B-splines (THB-splines) [13] can be used for an easy extension to hierarchical spline spaces of any QI scheme formulated in a standard spline space [24, 25]. These features make THB-splines a natural choice to solve our scattered data fitting problem and are then the solution here considered.

The first proposal based on adaptive THB-spline fitting of scattered data for the reconstruction of industrial models was introduced in [15], where a global adaptive smoothed least-squares scheme was developed. Successively, in order to increase its locality and reduce the computational cost, the same problem has been addressed in different papers by combining a two-stage approach with hierarchical spline approximations. The first contribution where this kind of schemes was used by some of the authors was presented in [2] to deal with gridded data of Hermite type. In [4] these kinds of approximants were extended for the first time to scattered data, by using in the first stage of the scheme local polynomial least squares approximations of variable degree. A preliminary application of this scheme to industrial data reconstruction was given in [5], where its theoretical analysis was also presented.

In this paper a variant of the approach considered in [4, 5] is presented, to further decrease the number of degrees of freedom necessary to reach a certain accuracy and also to reduce the artifacts in the resulting surface. Since the local polynomials used in [4, 5] need to be converted into B-spline form for being usable by the quasi-interpolation approach considered in the second stage of scheme, we now work directly in local spline spaces. In this way, the algorithm for the first stage of the scheme is simplified. In order to avoid rank-deficiency problems by simultaneously controlling the smoothness of local approximants, a smoothing term is added to the least squares objective function. To improve the stability of the proposed method for general data configuration, we also inserted an automatic control on the choice of the local data sets.

The structure of the paper is as follows. Section 2 presents the model problem, while the first stage of the new scheme, devoted to the computation of local smoothed least squares B-spline approximations, is described in Section 3. Section 4 introduces the construction of the adaptive THB-spline surface approximating the scattered data set. A selection of examples on geometric models representing components of aircraft turbine blades is presented in Section 5, and compared with the results obtained in [5]. The numerical experiments include a new scattered data set with voids and the adaptive reconstruction of a surface closed in one parametric direction.

2 The problem

The industrial models here considered are components of aircraft turbine blades which can be suitably represented in parametric form by using just one map, with the possibility of being periodic in one parametric direction. The problem can be mathematically described as follows. Let

$$F := \{\mathbf{f}_i \in \mathbb{R}^3, i = 1, \dots, n\}$$

be a scattered set of distinct points in the 3D space which can be reasonably associated by a one-to-one map to a set $X := \{\mathbf{x}_i := (x_i, y_i) \in \Omega \subset \mathbb{R}^2, i = 1, \dots, n\}$ of distinct parameter values belonging to a closed planar parametric domain Ω . Since the choice of a suitable parametrization method for the definition of the set X , which can naturally influence the quality of the final approximation, is not our focus, in this paper we relate to classical parametrization methods based on a preliminary triangulation of the scattered data set F , see e. g., [11, 12]. Consequently, both F and X are considered input data for the approximation problem.

Focusing on two-stage spline approximation schemes, we can introduce the general idea referring for simplicity to their formulation in a standard space V of tensor product splines of bi-degree $\mathbf{d} = (d_1, d_2)$, where it is assumed $d_i \geq 1, i = 1, 2$, in order to deal with at least continuous functions. With this kind of methods, a quasi-interpolation operator Q is defined so that $Q(F, X) = \mathbf{s}$, with \mathbf{s} denoting a vector function, possibly periodic in one parametric direction, with components in the spline space V . Using a suitable spline basis $\mathcal{B} := \{B_J\}_{J \in \Gamma}$ of V , such vector spline \mathbf{s} can be expressed as follows,

$$\mathbf{s}(\mathbf{x}) := \sum_{J \in \Gamma} \lambda_J(F_J, X_J) B_J(\mathbf{x}), \quad \mathbf{x} \in \Omega, \quad (1)$$

where each coefficient vector $\lambda_J(F_J, X_J)$, $J \in \Gamma$, is computed in the first stage of the scheme by using a certain local subset $F_J \subset F$ of data and the corresponding set of parameter values $X_J \subset X$, so that $\mathbf{s}(\mathbf{x}_i) \approx \mathbf{f}_i$, $i = 1, \dots, n$.

When dealing with discrete data, measuring the accuracy of the spline approximation with the maximum of the errors $\|\mathbf{s}(\mathbf{x}_i) - \mathbf{f}_i\|_2$ at each parameter site can appear reasonable at the first sight. However, the quality of the approximant is also

strictly related to the lack of unwanted artifacts, a feature of fundamental importance for industrial applications of any approximation scheme. In this context, it is then a common practice to require the error under a prescribed tolerance only at a certain percentage of sites in X .

3 First stage: local B-spline approximations

For computing each vector coefficient λ_J , $J \in \Gamma$, necessary in (1) to define the approximation \mathbf{s} , we consider a local data subset $F_J \subset F$,

$$F_J := \{\mathbf{f}_i : i \in I_J\} \quad \text{with} \quad I_J := \{i : \mathbf{x}_i \in X \cap \Omega_J\},$$

associated to the set

$$X_J := \{\mathbf{x}_i : i \in I_J\}$$

of parameter values in a local subdomain Ω_J of Ω which has non empty intersection with the support of the basis function B_J , namely $\Omega_J \cap \text{supp}(B_J) \neq \emptyset$. By denoting with $\mathcal{B}_J := \{B_I : I \in \Lambda_J \subset \Gamma\}$ the set of B-splines in \mathcal{B} not vanishing in Ω_J (which necessarily includes B_J), the value of λ_J is defined as the vector coefficient associated with B_J in a local spline approximation $\mathbf{s}_J \in \mathcal{S}_J$, with $\mathbf{s}_J : \Omega_J \rightarrow \mathbb{R}^3$ and

$$\mathcal{S}_J := \text{span}\{B_I, I \in \Lambda_J\}.$$

Concerning this local spline space, the following proposition is proved, since it shows that \mathcal{S}_J has reasonable approximation power and in particular that it includes the restriction to Ω_J of any linear polynomial,

Proposition *The following space inclusion holds true*

$$\Pi_{\mathbf{d}}^2|_{\Omega_J} \subseteq \mathcal{S}_J := \text{span}\{B_I : I \in \Lambda_J\},$$

where $\Pi_{\mathbf{d}}^2|_{\Omega_J}$ denotes the restriction to Ω_J of the tensor product space of bivariate polynomials of bi-degree \mathbf{d} .

Proof Let c_K be a cell of \mathcal{G} such that $c_K \cap \Omega_J \neq \emptyset$. Then the definition of Λ_J implies that $c_K \subseteq \text{supp}(B_I) \Rightarrow I \in \Lambda_J$. Thus we can say that $\text{span}\{B_I : c_K \subseteq \text{supp}(B_I)\} \subseteq \mathcal{S}_J$. Since $\text{span}\{B_I : c_K \subseteq \text{supp}(B_I)\} = \Pi_{\mathbf{d}}^2|_{c_K}$, the proof is completed considering that c_K is any cell of \mathcal{G} with non vanishing intersection with Ω_J . \square

The variational fitting method adopted to determine \mathbf{s}_J in \mathcal{S}_J consists in minimizing the following objective function

$$\sum_{i \in I_J} \|\mathbf{s}_J(\mathbf{x}_i) - \mathbf{f}_i\|_2^2 + \mu E(\mathbf{s}_J), \quad (2)$$

where $\mu > 0$ is a smoothing coefficient and $E(\mathbf{s}_J)$ the thin-plate energy,

$$E(\mathbf{s}_J) := \int_{\Omega_J} \left\| \frac{\partial^2 \mathbf{s}_J}{\partial x^2} \right\|_2^2 + 2 \left\| \frac{\partial^2 \mathbf{s}_J}{\partial x \partial y} \right\|_2^2 + \left\| \frac{\partial^2 \mathbf{s}_J}{\partial y^2} \right\|_2^2 dx dy .$$

As recalled in the Appendix, the assumption of a positive μ ensures that this local approximation problem admits always a unique solution, provided that the sites belonging to X_J are not collinear.

Since the scheme is locally applied, an automatic (data-dependent) selection of the parameter μ could be considered. For example, the choice may take into account the cardinality $|X_J|$ of the local sample or the area of Ω_J , which influence the value of the first and of the second addend in (2), respectively. In view of this influence however, we may observe that a constant value of μ implies that the balancing between the fitting and the smoothing term in the objective function usually increases when $|X_J|$ or the area of Ω_J increases, being this true in the second case because second derivatives are involved in the smoothing term. Both these behaviors seem reasonable and are confirmed by the quality of the results obtained in our experiments, where a constant value for μ is suitably chosen.

Differently from [4, 5], in order to better avoid overfitting, a lower bound n_{\min} for the cardinality of X_J is now required, being $n_{\min} \geq 3$ the only additional input parameter required by the algorithm, besides μ . To fulfill this condition, Ω_J is initialized as $\text{supp}(B_J)$ and enlarged until $|X_J| \geq n_{\min}$. Note that the refinement strategy presented in Section 4 automatically guarantees that the inequality $|X_J| \geq n_{\min}$ becomes fulfilled without requiring an excessive enlargement of the set Ω_J , which would compromise the locality of the approximation. For this reason, it is not necessary to set a maximum value for controlling the enlargement of the local data set. Considering that the smaller n_{\min} is, the higher is the obtainable level of detail but also the probability of overfitting, in our experiments (which always adopt $d_1 = d_2 = d = 2, 3$), a good low range for its selection has always been $d^2 \leq n_{\min} \leq (d+1)^2$.

Summarizing, the computation of each λ_J , $J \in \Gamma$, is done according to the following algorithm.

Algorithm 1: local smoothing spline approximant

Inputs

- $F \subset \mathbb{R}^3$: scattered data set;
- $X \subset \Omega \subset \mathbb{R}^2$: set of parameter values corresponding to the data in F ;
- \mathcal{G} : uniform tensor-product mesh in Ω (possibly with auxiliary cells);
- V : tensor-product spline space of bi-degree \mathbf{d} associated with \mathcal{G} with B-spline basis \mathcal{B} ;
- $J \in \Gamma$: index of the considered basis function $B_J \in \mathcal{B}$;
- μ : smoothing spline parameter ($\mu > 0$);
- n_{\min} : minimum required number of local data ($3 \leq n_{\min} \ll |F|$);

1. Initialization

- a. initialize $\Omega_J = \text{supp}(B_J)$;
- b. initialize $I_J = \{i : \mathbf{x}_i \in X \cap \Omega_J\}$, $F_J = \{\mathbf{f}_i : i \in I_J\}$ and $X_J = \{\mathbf{x}_i : i \in I_J\}$;

2. while $|F_J| < n_{\min}$
 - a. enlarge Ω_J with the first ring of cells in \mathcal{G} surrounding Ω_J ;
 - b. update $I_J = \{i : \mathbf{x}_i \in X \cap \Omega_J\}$, $F_J = \{\mathbf{f}_i : i \in I_J\}$ and $X_J = \{\mathbf{x}_i : i \in I_J\}$;
3. if the sites in X_J are not collinear, then:
 - a. compute the local approximation $\mathbf{s}_J = \sum_{I \in \Lambda_J} \mathbf{c}_I^{(J)} B_I$ minimizing the objective function in (2) for the data F_J and X_J ;
 - b. set $\lambda_J = \mathbf{c}_J^{(J)}$;

else set $\lambda_J = \frac{1}{|F_J|} \sum_{I \in I_J} \mathbf{f}_I$.

Output

- vector coefficient λ_J .

Note that exploiting a regularized least square approximation and, as a consequence, being able to directly employ the local spline space has significantly simplified the algorithm originally proposed for the first stage in [4, 5], where a variable-degree local polynomial approximation was considered. In particular, the new scheme does not require the selection of a suitable degree for the computation of any coefficient λ_J and eliminates the conversion of the computed approximant from the polynomial to the B-spline basis.

In the following section, after introducing the THB-spline basis, the operator Q is easily extended to hierarchical spline spaces, by also introducing the automatic refinement algorithm here considered. Note that this extension rule ensures that the coefficient associated to a THB-spline basis function remains unchanged on a refined hierarchical mesh if this function remains active on the updated hierarchical configuration.

4 Second stage: THB-spline approximation

Let us consider a sequence $V^0 \subset \dots \subset V^{M-1}$ of M spaces of tensor-product splines of bi-degree $\mathbf{d} := (d_1, d_2)$, $d_i \geq 1, i = 1, 2$, defined on the closed domain Ω , and each one associated with the tensor-product grid \mathcal{G}^ℓ and the basis $\mathcal{B}_{\mathbf{d}}^\ell := \{B_J^\ell\}_{J \in \Gamma_{\mathbf{d}}^\ell}$. Let $\Omega^0 \supseteq \dots \supseteq \Omega^M$ be a sequence of closed domains, with $\Omega^0 := \Omega$ and $\Omega^M := \emptyset$. Each Ω^ℓ , for $\ell = 1, \dots, M-1$ is the union of cells of the tensor-product grid $\mathcal{G}^{\ell-1}$. Let $\mathcal{G}_{\mathcal{H}}$ be the hierarchical mesh defined by

$$\mathcal{G}_{\mathcal{H}} := \{Q \in \mathcal{G}^\ell, 0 \leq \ell \leq M-1\} \quad \text{with} \quad \mathcal{G}^\ell := \{Q \in \mathcal{G}^\ell : Q \subset \Omega^\ell \setminus \Omega^{\ell+1}\},$$

where each \mathcal{G}^ℓ is called the set of *active cells* of level ℓ . The hierarchical B-spline (HB-spline) basis $\mathcal{H}(\mathcal{G}_{\mathcal{H}})$ with respect to the mesh $\mathcal{G}_{\mathcal{H}}$ is defined as

$$\mathcal{H}(\mathcal{G}_{\mathcal{H}}) := \{B_J^\ell : J \in A^\ell, \ell = 0, \dots, M-1\},$$

where

$$A^\ell := \{J \in \Gamma^\ell : \text{supp}(B_J^\ell) \subseteq \Omega^\ell \wedge \text{supp}(B_J^\ell) \not\subseteq \Omega^{\ell+1}\},$$

is the set of *active multi-indices* of level ℓ , and $\text{supp}(B_J^\ell)$ denotes the intersection of the support of B_J^ℓ with Ω^0 . The corresponding hierarchical space is defined as $S_{\mathcal{H}} := \text{span } \mathcal{H}(\mathcal{G}_{\mathcal{H}})$.

For any $s \in V^\ell$, $\ell = 0, \dots, M-2$, let

$$s = \sum_{J \in \Gamma^{\ell+1}} \sigma_J^{\ell+1} B_J^{\ell+1}$$

be its representation in terms of B-splines of the refined space $V^{\ell+1}$. We define the truncation of s with respect to level $\ell+1$ and its (cumulative) truncation with respect to all finer levels as

$$\text{trunc}^{\ell+1}(s) := \sum_{J \in \Gamma^{\ell+1} : \text{supp}(B_J^{\ell+1}) \not\subseteq \Omega^{\ell+1}} \sigma_J^{\ell+1} B_J^{\ell+1},$$

and

$$\text{Trunc}^{\ell+1}(s) := \text{trunc}^{M-1}(\text{trunc}^{M-2}(\dots(\text{trunc}^{\ell+1}(s))\dots)),$$

respectively. For convenience, we also define $\text{Trunc}^M(s) := s$ for $s \in V^{M-1}$. The THB-spline basis $\mathcal{T}(\mathcal{G}_{\mathcal{H}})$ of the hierarchical space $S_{\mathcal{H}}$ was introduced in [13] and can be defined as

$$\mathcal{T}(\mathcal{G}_{\mathcal{H}}) := \{T_J^\ell := \text{Trunc}^{\ell+1}(B_J^\ell) : J \in A^\ell, \ell = 0, \dots, M-1\}.$$

The B-spline B_J^ℓ is called the *mother* B-spline of the truncated basis function T_J^ℓ .

We recall that THB-splines are linearly independent, non-negative, preserve the coefficients of the underlying sequence of B-splines, and form a partition of unity [13, 14]. Besides that, following the general approach introduced in [25], using such basis we can easily construct the vector THB-spline approximation of the whole scattered data set in terms of the hierarchical quasi-interpolant $\mathbf{s} = Q(F, X)$ as follows,

$$\mathbf{s}(\mathbf{x}) := \sum_{\ell=0}^{M-1} \sum_{I \in A^\ell} \lambda_I^\ell(F_I, X_I) T_I^\ell(\mathbf{x}), \quad (3)$$

where each vector coefficient λ_I^ℓ is the one of the mother function B_I^ℓ and is obtained by computing the local regularized B-spline approximation \mathbf{s}_I^ℓ on the data set F_I^ℓ associated to B_I^ℓ as described in Section 3.

In order to define an adaptive approximation scheme, the following algorithm is used to iteratively construct the hierarchical mesh $\mathcal{G}_{\mathcal{H}}$, the corresponding spline space $S_{\mathcal{H}}$ and the final approximating spline \mathbf{s} . As any automatic refinement strategy, the algorithm requires in input some parameters which drive the refinement process. One of them is the error tolerance $\epsilon > 0$ whose value has to be chosen not only considering the accuracy desired for the reconstruction but also taking into account

the level of noise affecting the given points cloud (here assumed without significant outliers). The percentage bound parameter η specifies the number of data points for which the error is required to be within the given tolerance and a value strictly less than 100% is suggested to reduce the influence of outliers with moderate size on the approximation. Another required parameter is the maximum number of levels M_{\max} which has to be chosen considering the maximal level of detail desired for the reconstruction. The choice of the other input parameters, n_{loc} , n_1 and n_2 , is discussed after the algorithm.

Algorithm 2: adaptive hierarchical spline fitting

Inputs

- $F \subset \mathbb{R}^3$: scattered data set with $n = |F|$;
- $X \subset \Omega \subset \mathbb{R}^2$: set of parameter values corresponding to the data in F ;
- \mathcal{G}^0 : initial uniform tensor-product mesh in Ω (possibly with auxiliary cells);
- V^0 : tensor-product spline space of bi-degree \mathbf{d} associated with \mathcal{G}^0 ;
- ϵ : error tolerance;
- η : percentage bound of data points for which the error is required to be within the tolerance ϵ (default: $\eta = 95\%$);
- M_{\max} : integer specifying the maximum number of levels;
- n_{loc} : minimum number of local data required for refinement ($3 \leq n_{\text{loc}} \ll |F|$);
- n_1, n_2 : positive integers specifying the number of uniform horizontal or vertical splittings for the support of a tensor-product B-spline (default: $n_1 = n_2 = 1$).

1. Initialization

- a. set $\mathcal{G}_{\mathcal{H}} = \mathcal{G}^0$ and $S_{\mathcal{H}} = V^0$;
- b. set the current number of level $M = 1$;
- c. use Algorithm 1 to compute the coefficients of the hierarchical QI vector spline $\mathbf{s} \in S_{\mathcal{H}}$ introduced (3).
- d. evaluate the errors at the data sites $\mathbf{x}_i, \in X, i = 1, \dots, n$,

$$e(\mathbf{x}_i) := \|\mathbf{s}(\mathbf{x}_i) - \mathbf{f}_i\|_2, \quad i = 1, \dots, n.$$

2. While $|\{i : e(\mathbf{x}_i) > \epsilon\}|/n > \eta$ and $M \leq M_{\max}$, repeat the following steps:

- a. (*marking*) for each $\ell = 0, \dots, M - 1$, mark the cells of level ℓ in $\mathcal{G}_{\mathcal{H}}$ which are included in the support of a $B_J^\ell, J \in A^\ell$ such that:
 - there exists at least one data site $\mathbf{x}_i \in X \cap \text{supp}(B_J^\ell)$ such that $e(\mathbf{x}_i) > \epsilon$;
 - there are at least $\lceil n_{\text{loc}}/(n_1 n_2) \rceil$ data sites in any of the $n_1 n_2$ subrectangles which uniformly split $\text{supp}(B_J^\ell)$ (note that such splitting is just temporarily considered to check whether this refinement requirement is satisfied);
- b. (*update the hierarchical mesh*) update $\mathcal{G}_{\mathcal{H}}$ by dyadically refining (in the two parametric directions) all the marked cells;
- c. (*update the number of levels*) set M equal to the current number of levels;
- d. (*update the hierarchical space*) update the THB-spline basis of $S_{\mathcal{H}}$ and the sets $A^\ell, \ell = 0, \dots, M - 1$;

- e. (*spline update*) use Algorithm 1 to compute the coefficients of the hierarchical QI spline \mathbf{s} defined in (3) – only coefficients associated with THB-splines having new mother B-splines have to be computed;
- f. (*error update*) evaluate the new errors $e(\mathbf{x}_i), i = 1, \dots, n$ at the data sites;

Outputs

- $\mathcal{G}_{\mathcal{H}}$: hierarchical mesh;
- $\mathcal{T}(\mathcal{G}_{\mathcal{H}})$: THB-spline basis of $S_{\mathcal{H}}$;
- coefficients of the hierarchical spline \mathbf{s} defined in (3).

The refinement criterion has been motivated by the observation that, when the parameter values corresponding to the local data set F_I^ℓ are concentrated in a small part of the support of B_I^ℓ , the quality of the approximation may be affected. For this reason, we consider a splitting of the two sides of $\text{supp}(B_J^{M-1})$ in n_1 and n_2 uniform segments, respectively, and subdivide the support of B_J^{M-1} in the resulting $n_1 n_2$ subregions, where we then check the presence of at least $\lceil n_{\text{loc}} / (n_1 n_2) \rceil$ data points. To simplify the usage of the algorithm by default we set $n_1 = n_2 = 1$ but different values can be chosen if suitable, see e.g. the data set with voids considered in Example 3 of Section 5.

Concerning the parameter n_{loc} , we may observe that the requirement $n_{\text{loc}} \geq n_{\text{min}}$ guarantees that the points needed to compute the coefficients associated with the new functions in the first stage of the next iteration can be found not too far from the support of the functions themselves. Indeed in the algorithm introduced in the previous section, after a few enlargements, Ω_J will surely include the support of a refined function of the previous level intersecting $\text{supp}(B_J)$. As a consequence, analogously to n_{min} , a high value of n_{loc} contributes to the reduction of oscillations deriving from overfitting, but this value should also be low enough to guarantee that the refinement strategy can generate a hierarchical spline space with enough degrees of freedom for satisfying the given tolerance ϵ . For this reason, some tuning is necessary for a good selection of n_{loc} .

The proposed adaptive approximation method also extends to the case, not addressed in the previous works, of surfaces closed in one (or even two) parametric directions. Note that the local nature of the considered approximation approach makes the implementation especially easy, since coefficients associated with a THB present at successive steps of the adaptive refinement procedure (even if possibly further truncated) do not depend on such steps.

5 Examples

We present a selection of tests for the approximation of industrial data obtained by an optical scanning process of four different aircraft engine parts. For each of these surfaces, as a characterizing dimension, we report the length R of the diagonal of the minimal axis-aligned bounding box associated to the given point cloud.

The parameter values are computed in all examples with standard parametrization methods based on a triangulation of the scattered data sets, see e. g., [11, 12]. The bi-degree \mathbf{d} is set equal to $(2, 2)$ in the first considered example and always equal to 3 in the other examples. The results highlight the effects of considering a minimum number of local data points (also) in the first stage of method, as well as the improvements obtained by introducing a regularized B-spline approximation for each local fitting with respect to the scattered data fitting scheme considered in [5]. By combining these two changes, the two-stage approximation algorithm is more stable and unwanted oscillations are further reduced.

Concerning the parameters in input to Algorithm 1, we have always set $\mu = 10^{-6}$, except for Example 2 where it was chosen even smaller. In order to try to produce a very detailed reconstruction, a quite small value has been chosen for n_{\min} , always selecting it between d^2 and $(d + 1)^2$. Concerning the parameter selection for Algorithm 2, in all the presented experiments we have set the maximum number of levels M_{\max} equal to 8, fixing the percentage bound parameter η equal to its default value ($\eta = 95\%$). The error tolerance ϵ has been always set to $5 \cdot 10^{-5}$ m, except for Example 2 where we used the about halved value chosen for the same experiment in [5] (as a reference value, for each example consider the dimension R characterizing the related point cloud). The integer parameters n_1 and n_2 also in input to Algorithm 2 have been always set to their unit default values, except for Example 3 which required a different selection because of the voids present in the considered data set. The only parameter which has required a finer tuning for the reported experiments has been n_{loc} which is anyway always assumed greater than n_{\min} .

Example 1 (Tensile) In this example, we consider THB-spline approximations to reconstruct a part of a tensile from the set of 9281 scattered data shown in Figure 1 (top) which has reference dimension $R = 2.5 \cdot 10^{-2}$ m. We compare the new local scheme based on local B-spline approximations with the algorithm based on local polynomial approximations of variable degree presented in [5], where this test was originally considered. Note that for this example we have never dealt with local sets of collinear points in our experiments.

As the first test, we ran both algorithms with the same setting considered in [5], namely, by starting with a 4×4 tensor-product mesh with $\mathbf{d} = (2, 2)$, tolerance $\epsilon = 5 \cdot 10^{-5}$ m, percentage bound η equal to the default 95% and $n_{\text{loc}} = 20$. The algorithm with local polynomial approximations with the parameter choice considered in [5] ($\sigma = 10^8$) led to an approximation with 2501 degrees of freedom, 96.25% of points below the tolerance and a maximum error of $1.22745 \cdot 10^{-4}$ m. The new scheme based on local B-spline approximations with $n_{\min} = 6$ and $\mu = 10^{-6}$ generated a THB-spline surface with 1855 degrees of freedom that satisfies the required tolerance in 98.88% of points with a maximum error of $8.06007 \cdot 10^{-5}$ m. The number of levels used is $M = 5$, but all the cells of the first two levels are refined.

As the second test, we ran both algorithms by starting with a 16×4 tensor-product mesh with $\mathbf{d} = (2, 2)$, percentage η equal to the default, tolerance $\epsilon = 5 \cdot 10^{-5}$ m, and $n_{\text{loc}} = 15$. The algorithm with local polynomial approximations led to an approximation with 5922 degrees of freedom, 98.18% of points below the tolerance and a maximum error of $1.44222 \cdot 10^{-4}$ m. The surface and the corresponding

hierarchical mesh are shown in Figure 1 (center). This approximation clearly shows strong oscillations on the boundary of the reconstructed surface, due to a lack of available data points for the local fitting in correspondence of high refinement levels. The scheme based on local B-spline approximations, with $n_{\min} = 7$ and $\mu = 10^{-6}$ produced a THB-spline surface with 1960 degrees of freedom that satisfies the required tolerance in the 99.36% of the data points with a maximum error of $8.10814 \cdot 10^{-5}$ m. The surface, free of unwanted oscillations along the boundary, and the corresponding hierarchical mesh are shown in Figure 1 (bottom). The number of levels used is $M = 4$, with all the cells of level 0 refined.

Example 2 (Blade) In this example, we test the second example considered in [5] on the set of 27191 scattered data representing a scanned part of a blade shown in Figure 2 (top) whose reference dimension R is equal to $5 \cdot 10^{-2}$ m. Again, to compare the new local scheme with the algorithm based on local polynomial approximations there considered, we ran both algorithms with the same setting of [5], namely, by starting with a 4×4 tensor-product mesh with $\mathbf{d} = (3, 3)$, tolerance $\epsilon = 2 \cdot 10^{-5}$ m, percentage bound η equal to the default 95% and $n_{\text{loc}} = 60$. The algorithm with local polynomial approximations with the parameter choice considered in [5] ($\sigma = 10^8$) led to an approximation with 12721 degrees of freedom, 97.06% of points below the tolerance and a maximum error of $1.08043 \cdot 10^{-4}$ m. The new scheme based on local B-spline approximations with $n_{\min} = 6$ and $\mu = 10^{-8}$ generated a THB-spline surface with 8314 degrees of freedom that satisfies the required tolerance in 99.94% of points with a maximum error of $1.32976 \cdot 10^{-4}$ m. The surface and the corresponding hierarchical mesh are shown in Figure 2 (bottom). The number of levels used with the new scheme is 7 (the first two are not visible in the mesh because their cells are fully refined), one less than with the old approach. Finally for completeness we precise that the local collinearity check in Algorithm 1 is active only for 5 coefficients.

Example 3 (Endwall) In this example we illustrate the behavior of the adaptive algorithm on data sets with voids by considering the reconstruction of an endwall part from the scattered data set of 43869 points shown in Figure 3 (top) ($R = 5 \cdot 10^{-1}$ m). The figure shows that in this case the data set represents a model with three different holes, where no input data are available. The aim of this reconstruction is to avoid artifacts due to lack of points and obtain a sufficiently regular surface (e.g. by avoiding self-intersections), that can be post-processed with standard geometric software tools to obtain a suitably trimmed model. Consequently, not only the number of points in the local data sets is important to reach this aim, but also their distribution. To properly address this issue, we consider a real density parameter with value between 0 and 1 which determines whether the distribution of the points in the local set is reliable or not for the fitting. The distribution of the local data points is computed as the number of mesh cells of level ℓ inside the support of B^ℓ or its enlargement, which contain at least one point, over the total number of mesh cells, either in the support of B^ℓ or its enlargement. If this ratio is below a density threshold, then more data points are required and the function support is enlarged for the computation of the local approximation in the first stage of the method. The approximation is developed by

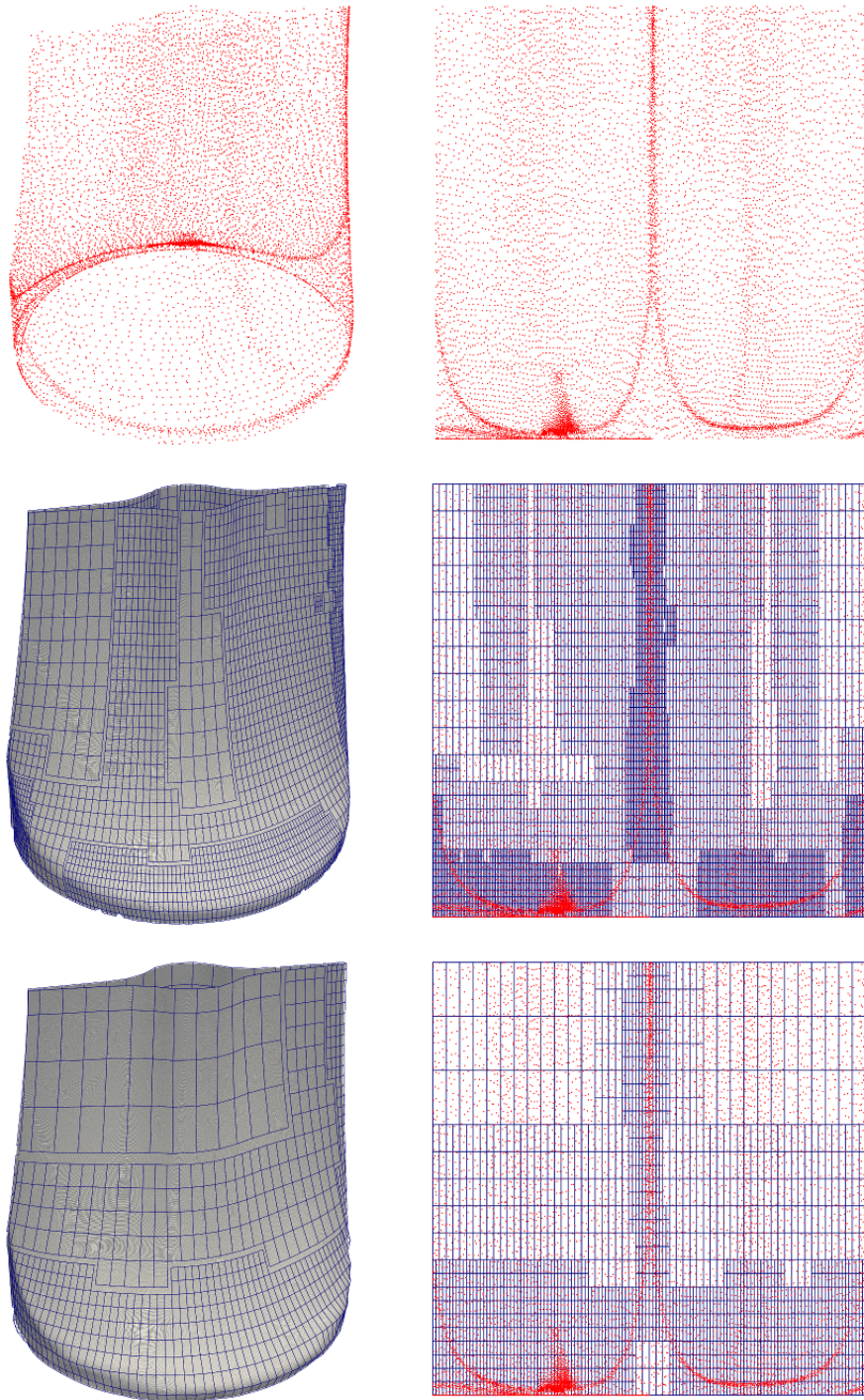


Fig. 1 Example 1: scattered data set corresponding to a critical part of a tensile (top), the reconstructed surfaces and the corresponding hierarchical meshes obtained with the algorithm presented in [5] (center) and the new scheme (bottom).

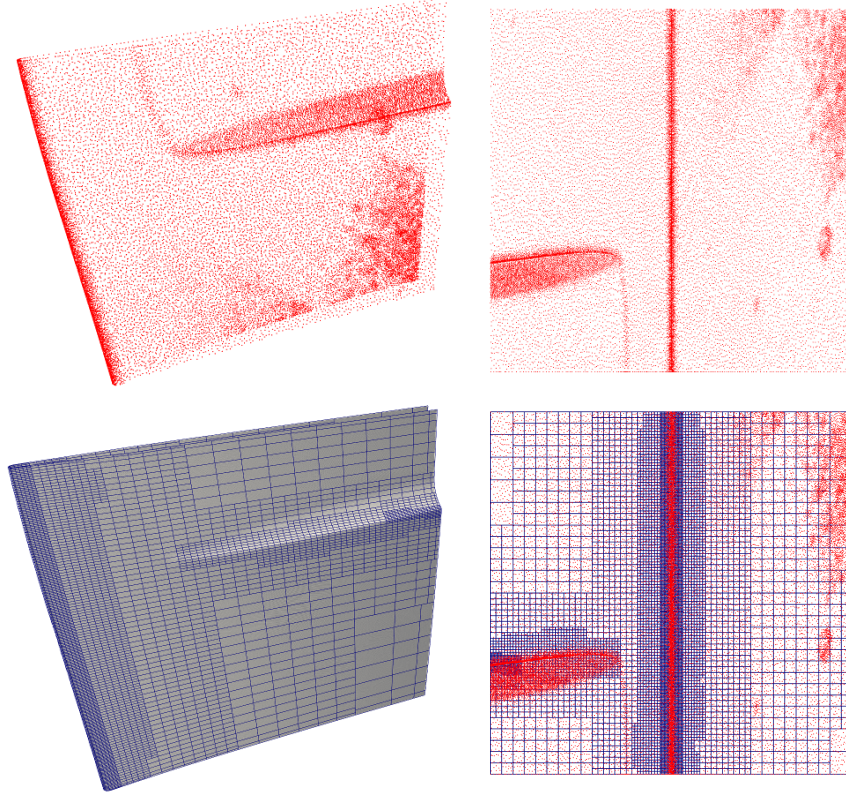


Fig. 2 Example 2: scattered data set corresponding to a critical part of a blade (top), the reconstructed surface and the corresponding hierarchical mesh obtained with the new scheme (bottom).

starting from a 32×32 tensor-product mesh, with $\mathbf{d} = (3, 3)$, $n_{\text{loc}} = 15$, $n_{\text{min}} = 11$, $\mu = 10^{-6}$ and $n_1 = n_2 = 2$. A choice of the density parameter δ equal to 0.3 permits to take care of the difficult distribution of data points in the construction of the approximation. By considering a tolerance $\epsilon = 5 \cdot 10^{-5}$ m and a percentage bound η equal to the default 95%, the refinement generated a THB-spline approximation with $M = 3$ and 11211 degrees of freedom, 98.70% of points below the threshold and a maximum error of $5.68999 \cdot 10^{-4}$ m. The surface and the corresponding hierarchical mesh are shown in Figure 3 (bottom). In this case there are 18 coefficients of the last level and 15 of the last but one (all associated with B-splines whose support intersects a void) such that the related X_J is made up of all collinear points.

Example 4 (Airfoil) This example illustrates the behavior of the new adaptive fitting algorithm with local B-spline approximations for surfaces closed in one parametric direction. We test the scheme to reconstruct a blade airfoil from the set of 19669 scattered data shown in Figure 4 (top) ($R = 10^{-1}$ m). We ran the method by starting with a 32×4 tensor-product mesh with $\mathbf{d} = (3, 3)$, setting $\eta = 95\%$, $\epsilon = 5 \cdot 10^{-5}$ m,

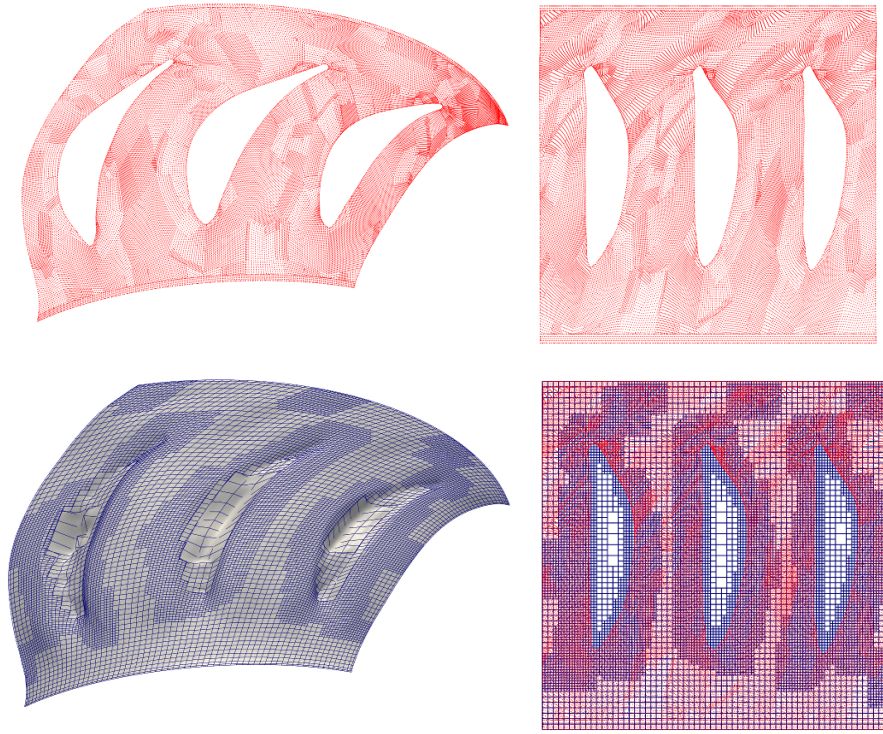


Fig. 3 Example 3: scattered data set corresponding to a critical part of an endwall (top), the reconstructed surface and the corresponding hierarchical mesh obtained with the new scheme (bottom).

and $n_{\text{loc}} = 30$ in Algorithm 2 and using Algorithm 1 with $n_{\text{min}} = 12$ and $\mu = 10^{-6}$. The refinement strategy produced an approximation with $M = 3$ and 1856 degrees of freedom distributed in the last two levels, that satisfies the required tolerance in 95.06% of the data points with maximum error $1.87742 \cdot 10^{-4}$ m (observe also that, as well as for Example 1, at the local stage X_J is never made up of all collinear points). The surface and the corresponding hierarchical mesh are shown in Figure 4 (bottom). By trying to force additional refinement, some oscillations appear. In this case, they are consistent with the data distribution since there are clusters of high density points, due to scan noise. Consequently, they do not represent artifacts caused by regions with very low density of data and cannot be prevented by exploiting the bound for cardinality of the local data sample governed by n_{loc} .

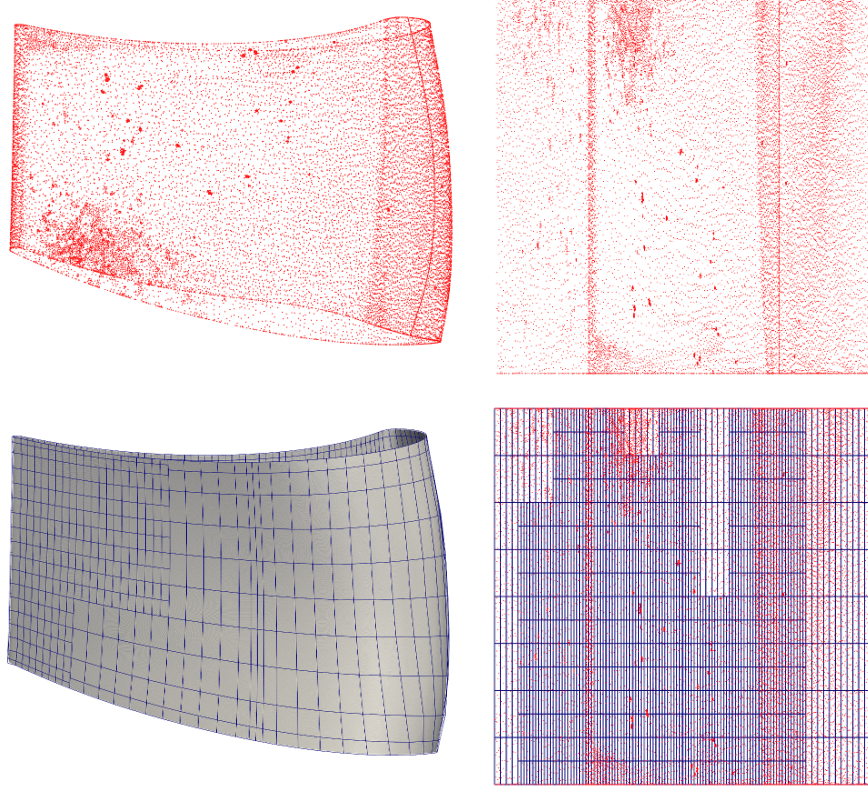


Fig. 4 Example 4: scattered data set corresponding to a critical part of an airfoil (top), the reconstructed surface and the corresponding hierarchical mesh obtained with the new scheme (bottom).

6 Appendix

In this appendix we give the proof that, assuming $\mathbf{d} = (d_1, d_2)$ with $d_i \geq 1, i = 1, 2$, the local vector spline s_J defined in Section 3 exists and is unique, provided that the sites in X_J are not collinear. First of all we observe that the objective function in (2) can be split in the sum of three analogous objective functions, one for each component $s_J^{(k)}, k = 1, 2, 3$, of \mathbf{s}_J ,

$$\sum_{i \in I_J} \|\mathbf{s}_J(\mathbf{x}_i) - \mathbf{f}_i\|_2^2 + \mu E(\mathbf{s}_J) = \sum_{k=1}^3 \left(\sum_{i \in I_J} (s_J^{(k)}(\mathbf{x}_i) - (\mathbf{f}_i)_k)^2 + \mu \rho(s_J^{(k)}) \right),$$

where

$$\rho(s_J^{(k)}) := \int_{\Omega_J} \left(\frac{\partial^2 s_J^{(k)}}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 s_J^{(k)}}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 s_J^{(k)}}{\partial y^2} \right)^2 \mathrm{d}x \mathrm{d}y.$$

Thus the study can be developed in the scalar case and for brevity we remove the subscript or superscript k ranging between 1 and 3. The analysis is developed in the following theorem, where $s_J : \Omega_J \rightarrow \mathbb{R}$ denotes the local spline in \mathcal{S}_J associated to the coefficient vector $\mathbf{c} \in \mathbb{R}^{\ell_J}$, with $\ell_J := |\Lambda_J|$,

$$s_J(\mathbf{x}) = \sum_{I \in \Lambda_J} c_I B_I(\mathbf{x}).$$

Theorem *Let the considered spline bi-degree $\mathbf{d} = (d_1, d_2)$ be such that $d_i \geq 1, i = 1, 2$. When the points $\mathbf{x}_i \in \Omega_J, i \in I_J$ are not collinear, there exists a unique local spline $s_J \in \mathcal{S}_J$ minimizing the following objective function,*

$$\sum_{i \in I_J} (s_J(\mathbf{x}_i) - f_i)^2 + \mu \rho(s_J), \quad (4)$$

where $\mu > 0$. If such points in Ω_J are collinear, then such minimizer does not exist or is not unique.

Proof Let us observe that $\rho(s_J) = \mathbf{c}^T M \mathbf{c}$, where $M \in \mathbb{R}^{\ell_J \times \ell_J}$ is such that

$$M_{i,r} := \int_{\Omega_J} \left(\frac{\partial^2 B_I}{\partial x^2} \right) \left(\frac{\partial^2 B_R}{\partial x^2} \right) + 2 \left(\frac{\partial^2 B_I}{\partial x \partial y} \right) \left(\frac{\partial^2 B_R}{\partial x \partial y} \right) + \left(\frac{\partial^2 B_I}{\partial y^2} \right) \left(\frac{\partial^2 B_R}{\partial y^2} \right) dx dy,$$

where we are assuming that, in the adopted ordering of the B-spline basis elements of \mathcal{S}_J , B_R and B_I are the r -th and the i -th ones. On the other hand it is

$$\sum_{i \in I_J} (s_J(\mathbf{x}_i) - f_i)^2 = \|\mathbf{V} \mathbf{c} - \mathbf{F}\|_2^2 = \mathbf{c}^T A^T A \mathbf{c} - 2\mathbf{F}^T A \mathbf{c} + \mathbf{F}^T \mathbf{F},$$

where $\mathbf{F} \in \mathbb{R}^{|I_J|}$ denotes the vector collecting all the $f_i, i \in I_J$ and A is the $|I_J| \times \ell_J$ collocation matrix of the tensor-product B-spline basis generating \mathcal{S}_J . Thus the objective function in (4) can be written also as the following quadratic function,

$$\mathbf{c}^T (A^T A + \mu M) \mathbf{c} - 2\mathbf{F}^T A \mathbf{c} + \mathbf{F}^T \mathbf{F}.$$

As well known a quadratic function admits a global unique minimizer if and only if the symmetric matrix defining its homogeneous quadratic terms is positive definite and in such case the minimizer is given by its unique stationary point. In our case such matrix is $A^T A + \mu M$ and the stationary points are the solutions of the following linear system of ℓ_J equations in as many unknowns,

$$(A^T A + \mu M) \mathbf{c} = A^T \mathbf{F}.$$

Now, for all positive μ the matrix $A^T A + \mu M$ is symmetric and positive semidefinite since, for any vector $\zeta \in \mathbb{R}^{\ell_J}, \zeta \neq \mathbf{0}$ it is $\zeta^T A^T A \zeta \geq 0$ and $\zeta^T M \zeta \geq 0$, the last inequality descending from the fact that $\zeta^T M \zeta = \rho(s_\zeta)$, where $s_\zeta(\mathbf{x}) = \sum_{I \in \Lambda_J} \zeta_I B_I(\mathbf{x})$. Now if the points $\mathbf{x}_i, i \in I_J$ are distributed in Ω_J along a straight line $ax + by + c = 0$, the proposition proved in Section 3 implies that it is possible

to find $\zeta \in \mathbb{R}^{\ell_J}$, $\zeta \neq \mathbf{0}$ such that $s_\zeta(\mathbf{x}) \equiv ax + by + c, \forall \mathbf{x} \in \Omega_J$. This implies that $s_\zeta(\mathbf{x}_i) = 0, \forall i \in I_J$, that is the vector $A\zeta \in \mathbb{R}^{|I_J|}$ vanishes. On the other hand, clearly it is also $0 = \rho(s_\zeta) = \zeta^T M \zeta$, since $s_\zeta|_{\Omega_J}$ is a linear polynomial. This proves that the symmetric positive semidefinite matrix $(A^T A + \mu M)$ is not positive definite when all the $\mathbf{x}_i, i \in I_J$ are collinear. This is the only possible data distribution associated to a non positive definite matrix. Indeed if the points $\mathbf{x}_i, i \in I_J$ are not collinear, if $\zeta \in \mathbb{R}^{\ell_J}$, $\zeta \neq \mathbf{0}$ is associated to a non vanishing linear polynomial, it is $\zeta^T M \zeta = \rho(s_\zeta) = 0$ but $A\zeta \neq \mathbf{0}$ and so $\zeta^T A^T A \zeta > 0$; on the other hand if $\zeta \in \mathbb{R}^{\ell_J}$, $\zeta \neq \mathbf{0}$ is not associated to a linear polynomial, then $\zeta^T M \zeta = \rho(s_\zeta) > 0$. \square

Acknowledgements

We thank the anonymous reviewer for his/her useful suggestions. Cesare Bracco, Carlotta Giannelli and Alessandra Sestini are members of the INdAM Research group GNCS. The INdAM support through GNCS and Finanziamenti Premiali SUNRISE is also gratefully acknowledged.

References

1. Beirão da Veiga, L., Buffa, A., Sangalli, G., Vazquez, R., Analysis-suitable T-splines of arbitrary degree: definition and properties, *Math. Mod. Meth. Appl. Sci.* **23**, 1979–2003 (2013).
2. Bracco, C., Giannelli, C., Mazzia, F., Sestini, A. Bivariate hierarchical Hermite spline quasi-interpolation. *BIT Numer. Math.* **56**, 1165–1188 (2016).
3. Bracco, C., Lyche, T., Manni, C., Roman, F., Speleers, H., On the dimension of Tchebycheffian spline spaces over planar T-meshes, *Comput. Aided Geom. Des.* **45**, 151–173 (2016).
4. Bracco, C., Giannelli, C., Sestini, A. Adaptive scattered data fitting by extension of local approximations to hierarchical splines. *Comput. Aided Geom. Des.* **52–53**, 90–105 (2017).
5. Bracco, C., Giannelli, C., Großmann, D., Sestini, A. Adaptive fitting with THB-splines: Error analysis and industrial applications. *Comput. Aided Geom. Des.* **62**, 239–252 (2018).
6. Cavoretto, R., De Rossi, A., Perracchione, E., Optimal selection of local approximants in RBF-PU interpolation, *J. Sci. Comput.* **74**, 1–22 (2018).
7. Davydov, O., Morandi, R., Sestini, A., Local Hybrid Approximations for Scattered Data Fitting with Bivariate Splines, *Comput. Aided Geom. Des.* **23**, 703–721 (2006)
8. Davydov, O., Schumaker, L., Interpolation and scattered data fitting on manifolds using projected Powell–Sabin splines, *IMA J. of Numer. Anal.* **28**, 785–805 (2008)
9. Deng, J., Chen, F., Feng, Y., Dimensions of spline spaces over T-meshes, *J. Comput. Appl. Math.* **194**, 267–283 (2006)
10. Dokken, T., Lyche, T., Pettersen, K. F., Polynomial splines over locally refined box-partitions, *Comput. Aided Geom. Des.* **30**, 331–356 (2013).
11. Floater, M.S.: Parametrization and smooth approximation of surface triangulations. *Comput. Aided Geom. Des.* **14**, 231–250 (1997)
12. Floater, M.S., Hormann, K.: Surface parameterization: a tutorial and survey. In: Dodgson N. A., Floater M. S., Sabin M. A. (eds) *Advances in Multiresolution for Geometric Modelling, Mathematics and Visualization*, pp. 157–186. Springer, Berlin, Heidelberg (2005)

13. Giannelli, C., Jüttler, B., Speleers, H. THB-splines: The truncated basis for hierarchical splines. *Comput. Aided Geom. Des.* **29**, 485–498 (2012)
14. Giannelli, C., Jüttler, B., Speleers, H. Strongly stable bases for adaptively refined multilevel spline spaces, *Adv. Comput. Math.* **40** 459–490 (2014)
15. Kiss, G., Giannelli, C., Zore, U., Jüttler, B., Großmann, D., Barner, J. Adaptive CAD model (re-)construction with THB-splines. *Graph. Models* **76**, 273–288 (2014).
16. Kraft, R., Adaptive and Linearly Independent Multilevel B-Splines, in *Surface Fitting and Multiresolution Methods*, Le Méhauté, A. and Rabut, C. and Schumaker, L. L. eds., 209-218, Vanderbilt University Press (1997).
17. Lai M.J., Schumaker L. L., A domain decomposition method for computing bivariate spline fits of scattered data, *SIAM J. of Numerical Analysis* **47**, 911–928 (2009).
18. Lee B.G., Lyche T., Mørken K., Some examples of quasi-interpolants constructed from local spline projectors, in *Mathematical Methods for Curves and Surfaces, Oslo 2000*, T. Lyche and L.L. Schumaker eds., 243–252, Vanderbilt University Press (2001).
19. Lyche T., Schumaker L. L., Local spline approximation methods, *J. of Approximation Theory* **15**, 294–325 (1975).
20. Mourrain, B., On the dimension of spline spaces on planar T-meshes, *Math. Comp.* **83**, 847-871 (2014).
21. Patrizi, F., Dokken, T., Linear dependence of bivariate Minimal Support and Locally Refined B-splines over LR-meshes, *Comput. Aided Geom. Des.* **77** (2020).
22. Prautzsch, H., Boehm, W., Paluszny, M., Bézier and B-spline techniques, Springer, Berlin, 2002.
23. Sederberg, T.W., Zheng, J., Bakenov, A., Nasri, A., T-splines and T-NURCCs, *ACM Trans. Graph.* **22**, 477-484 (2003).
24. Speleers, H. Hierarchical spline spaces: quasi-interpolants and local approximation estimates. *Adv. Comput. Math.* **43**, 235–255 (2017)
25. Speleers, H., Manni, C. Effortless quasi-interpolation in hierarchical spaces. *Numer. Math.* **132**, 155–184 (2016)