Adaptive Finite Element Thin-Plate Spline With Different Data Distributions

Lishan Fang and Linda Stals

Abstract The finite element thin-plate spline fits large scattered data efficiently while retaining the smoothing properties of the thin plate-spline. Its computational cost is reduced by adaptive refinement that only refines sensitive regions identified by an error indicator. Several traditional error indicators of the finite element method were adapted for the finite element thin-plate spline and their performance has been evaluated using a large number of uniformly distributed data. In this article, we build on that work to examine three new data distribution patterns, which are the uniform distribution with missing data, random distribution and random normal distribution. A numerical experiment is conducted to assess the performance of the finite element thin-plate spline and three error indicators with these four data distribution patterns.

1 Introduction

The thin-plate spline is a data fitting technique that possesses many favourable properties like insensitivity to noise [6]. One obstacle of its usage is the high computational cost and memory requirement for large data sets. The finite element thin-plate spline (TPSFEM) was proposed by Roberts, Hegland and Altas [11] to efficiently interpolate large data sets with similar smoothing properties as the thin-plate spline. It uses simple H^1 finite elements resulting in a sparser system of equations as opposed to ones with higher-order finite elements used in [5]. A detailed formulation of the TPSFEM is provided by Stals and Roberts [16] and a brief description is given below similar to the one shown by Fang [8].

Lishan Fang
Mathematical Sciences Institute, Australian National University, e-mail: Lishan.Fang@anu.edu.au

Linda Stals
Mathematical Sciences Institute, Australian National University, e-mail: Linda.Stals@anu.edu.au

Let $\{(\boldsymbol{x}_{(i)}, y_{(i)}) : i = 1, 2, \ldots, n\}$ be the observed data of size n and dimension d on a domain Ω, where $\boldsymbol{x}_{(i)} \in \mathbb{R}^d$ and $y_{(i)} \in \mathbb{R}$ are i-th predictor value and response value, respectively. The TPSFEM s is defined as a combination of piecewise linear basis functions $\boldsymbol{b}$, where $s(\boldsymbol{x}) = \boldsymbol{b}(\boldsymbol{x})^T\boldsymbol{c}$ and $\boldsymbol{c}$ are coeffciients of the basis functions. The TPSFEM s minimises functional

$$J(\boldsymbol{c}, \boldsymbol{g}_1, \ldots, \boldsymbol{g}_d) = \boldsymbol{c}^T A\boldsymbol{c} - 2\boldsymbol{d}^T\boldsymbol{c} + \boldsymbol{y}^T\boldsymbol{y}/n + \alpha \sum_{k=1}^{d} \boldsymbol{g}_k^T L\boldsymbol{g}_k, \tag{1}$$

subject to Constraint $L\boldsymbol{c} = \sum_{k=1}^{d} G_k\boldsymbol{g}_k$, where $\boldsymbol{g}_k$ are coefficients of gradient approximations of s in dimension k, $A = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{b}(\boldsymbol{x}_{(i)})\boldsymbol{b}(\boldsymbol{x}_{(i)})^T$, $\boldsymbol{d} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{b}(\boldsymbol{x}_{(i)})y_{(i)}$, $\boldsymbol{y} = [y_{(1)}, \ldots, y_{(n)}]^T$, L is a discrete approximation to the negative Laplacian and G_k is a discrete approximation to the gradient operator in dimension k.

Smoothing parameter α balances the goodness of fit and smoothness of s. It is estimated iteratively using a stochastic estimator of the generalised cross-validation from Hutchinson [9] and more details are provided in [7]. It may also be calculated using alternate approaches discussed in [5]. Minimiser (1) is solved using Lagrange multipliers and the size of the resulting system of equations is proportional to the number of basis functions. This system is more efficiently solved than that of the thin-plate spline. A comparison using data from Section 4 between the TPSFEM and compactly supported basis functions (CSRBFs) from Wendland [17] with radius 0.5 is shown in Table 1. The TPSFEM achieves similar root mean square error (RMSE) and maximum errors (MAX) as the CSRBFs using a system with significantly fewer nonzero entries. A comprehensive comparison is in progress and will be provided in [15].

Table 1: Computational cost

Technique	No. basis	Dimension	No. nonzero	RMSE	MAX
TPSFEM	900	3603	52,060	0.027	0.20
TPSFEM	1600	6403	93,538	0.014	0.091
CSRBFs	1024	1024	496,274	0.0098	0.157

The remainder of the article is organised as follows. In Section 2, we show adaptive refinement and error indicators of the TPSFEM. In Section 3, four two-dimensional data distribution patterns are displayed and compared regarding their influence on the maximum distance to any data. In Section 4, a numerical experiment is presented to examine the influence of these patterns. In Section 5, we summarise this article and the findings of the experiment.

2 Adaptive Refinement

The accuracy of finite element approximations depends on the mesh size of the finite element grid [10]. The accuracy is improved by adaptive refinement that adapts

the accuracy of the approximation within sensitive regions, like peaks, dynamically during an iterative process. An error indicator marks regions that require finer elements to achieve higher accuracy for refinement. Many error indicators have been developed for approximating partial differential equations but they may not be applicable for the TPSFEM.

The formulation of the TPSFEM is different from that of the traditional finite element method and it may not provide the information required by some error indicators. For example, the TPSFEM uses the observed data instead of given functions of partial differential equations and the data is often perturbed by noise or irregularly distributed. Fang [8] adapted the iterative adaptive refinement process and three error indicators of the finite element method for the TPSFEM. In this article, we will focus on the performance of these three error indicators, which are the auxiliary problem error indicator, recovery-based error indicator and norm-based error indicator.

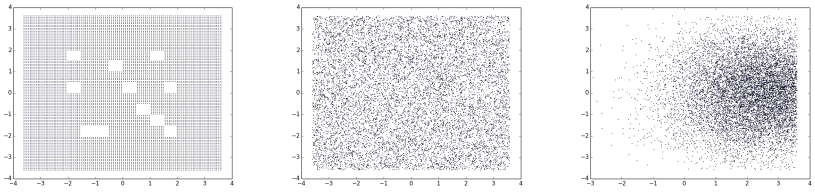
The auxiliary problem error indicator evaluates approximation quality by solving a local approximation, which is the TPSFEM built on a union of elements [1, 8, 10]. It solves Minimiser (1) using a small subset of the observed data within those elements. The local approximations are locally more accurate than the global TPSFEM s and the approximation quality is measured by the difference between them. The recovery-based error indicator estimates errors by post-processing the gradient approximations of the TPSFEM [18]. It improves the discontinuous gradient approximations of s with piecewise linear basis functions and calculates the error as the difference between the two gradient approximations. The norm-based error indicator uses an error bound on the L_∞ norm of the TPSFEM to optimise the approximation [13]. It approximates second-order derivatives of s to identify regions that change rapidly and refine them to improve accuracy.

These three error indicators use different information of the TPSFEM to indicate regions with large errors. The recovery-based error indicator and norm-based error indicator only use $\boldsymbol{c}$ values in Minimiser (1) and they were adapted without major changes. In contrast, the original auxiliary problem error indicator solves boundary value problems and was modified to use data instead of a given function [2, 7]. Consequently, it is more susceptible to changes in the data like noise [8]. When the data distribution pattern changes, these error indicators may behave differently.

As opposed to the finite element method, the error of the TPSFEM may not converge with a smaller mesh size h. Roberts, Hegland and Altas [11] proved that the error convergence of the TPSFEM depends on the smoothing parameter α, maximum distance to data $d_{\boldsymbol{x}}$ and h. The new iterative adaptive refinement process updates the optimal α after the grid is refined and prevents it from dominating the error [8]. Besides, previous studies tested the performance of the error indicators using uniformly distributed data sets of size 1,000,000, which provide sufficiently small $d_{\boldsymbol{x}}$ [8]. When the data is irregularly distributed, we may not have sufficiently small $d_{\boldsymbol{x}}$ over the whole domain and the error convergence will be affected.

3 Data Distribution

The observed data distributes differently depending on the application. For example, data is stored as maps of pixels for digital images or sampled randomly for surface reconstructions [4, 12]. Previous studies on the TPSFEM already deployed several data distribution patterns, including the uniform distribution [8], uniform distribution with missing data [14], random distribution [16] and random normal distribution [11]. We focus on these four data distribution patterns in this article.



(a) **(b)** **(c)**

Fig. 1: Data distribution of 10,000 data points with (a) uniform distribution with data missing; (b) random distribution; and (c) random normal distribution

The uniform distribution places data uniformly on the domain with fixed $d_{\boldsymbol{x}}$, which minimises its influence on the error convergence of the TPSFEM. However, other data distribution patterns may have varied $d_{\boldsymbol{x}}$ across the domain. Data points in certain regions may be missing and the TPSFEM needs to recover these surfaces. An example is shown in Figure 1(a), where data points in eleven square regions are missing and some of them neighbours each other. The error in these regions may not be improved by smaller h since $d_{\boldsymbol{x}}$ is large. Besides, the auxiliary problem error indicator uses data that is not available in those regions and the performance may be affected.

In many applications, the predictor values of data are sampled randomly with equal probabilities instead of a perfect uniform distribution, as shown in Figure 1(b). The random normal distribution places data points using a probability density function defined by a mean and a variance [3]. An example with variance 1.5; and mean 2.5 and 0 for predictor values $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ is shown in Figure 1(c). The density of the predictor values is higher at their mean than the rest of the domain. When a randomly distributed data set contains a large number of data points, the data density will be close across the domain. While it may not have a significant influence on the interpolant s due to sufficiently small $d_{\boldsymbol{x}}$, it may affect the error indicators as some elements may contain few data points. In comparison, a randomly normally distributed data set has different data densities across the domain with varied $d_{\boldsymbol{x}}$. While the error convergence behaviour may not be affected by randomly distributed data, the error indicators were not developed to handle data with various densities and their performance may be affected [5].

4 Numerical Experiment

A numerical experiment was conducted to test the error indicators using these data distribution patterns. The data consists of 10,000 data points limited inside $[-3.6, 3.6]^2$ and is modeled by the peaks function f from MATLAB, where $f(\boldsymbol{x}) = 3(1-x_1)^2 e^{-x_1^2-(x_2+1)^2} - 10(x_1/5 - x_1^3 - x_2^5)e^{-x_1^2-x_2^2} - \frac{1}{3}e^{-(x_1+1)^2-x_2^2}$. It has oscillatory surfaces at the center of the domain and flat surfaces near its boundaries [11]. Gaussian noise with mean 0 and standard deviation 0.01 is also included in some data sets to assess the performance in the presence of noise. The distribution patterns of irregularly distributed data sets have been shown in Figure 1.

We focus on the efficiency of uniform and adaptive grids, which is measured by the error metric versus the number of nodes in the grid. A grid that achieves a low error metric with a smaller number of nodes is considered more efficient. We consider both the root mean square error (RMSE) and approximate error, which measures how closely s fits data and reproduces f, respectively [7, 8]. The approximate error e is defined as $e = \sqrt{\sum_{i=1}^{m} h_i^2 e_i^2}$, where e_i is the difference between s and f at i-th node, h_i is the longest edge connected to i-th node and m is the number of nodes in the grid. The efficiency of final grids are calculated as products of the error metric and the number of nodes and is provided in the legend of each convergence plot.

4.1 Results

The convergence of the RMSE for data sets with the four distribution patterns is shown in Figure 2. Adaptive refinement focuses on refining the oscillatory surfaces at the centre and error convergence rates of all three adaptive grids are higher than that of the uniform grid in Figure 2(a). When the data is uniformly distributed, the three error indicators have similar performance and produce adaptive grids more than twice as efficient as the uniform grid.

Figure 2(c) shows similar error convergence of uniform and adaptive grids with random distribution. The TPSFEM and error indicators are not affected as $d_{\boldsymbol{x}}$ remains sufficiently small within a large number of randomly distributed data points. In contrast, $d_{\boldsymbol{x}}$ is large in some regions where data points are missing or the data is randomly normally distributed shown in figures 1(a) and 1(c), respectively. While this does not markedly affect the TPSFEM, it slightly weakens the performance of the auxiliary problem error indicator, as shown in figures 2(b) and 2(d), respectively. Local approximations of the auxiliary problems are built with different numbers of data points and the accuracy deteriorates.

The convergence of the approximate error for data sets with the four distribution patterns is shown in Figure 3. The error convergence rates of the approximate error with the uniform or random distribution in figures 3(a) and 3(c) are similar to those of the RMSE in figures 2(a) and 2(c). The TPSFEM closely reproduces the original smooth function f when $d_{\boldsymbol{x}}$ is sufficiently small in these two distribution

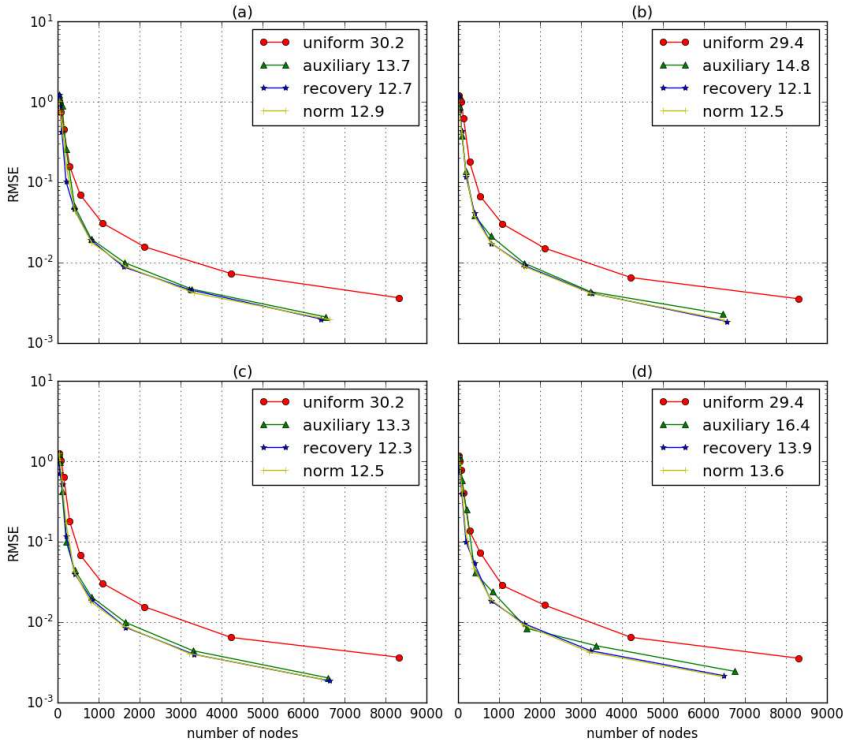


Fig. 2: RMSE for (a) uniform distribution; (b) uniform distribution with missing data; (c) random distribution; and (d) random normal distribution.

patterns. When the data is scarce in some regions, the TPSFEM interpolates smooth surfaces, which may not recover f especially when it is oscillatory. Consequently, the convergence of the approximate errors for the uniform distribution with missing data or random normal distribution slows down in the last few iterations as shown in figures 3(b) and 3(d), respectively. Similarly, the auxiliary problem error indicator underperforms compared to the other two error indicators.

In the presence of noise, the RMSE values of the TPSFEM stop decreasing at some point depending on the noise level of data as demonstrated by Fang [8]. Therefore, we only consider the approximate error here. The convergence of the approximate error for data sets with noise is shown in Figure 4. All error convergence rates are lower than those without noise since the TPSFEM may not reproduce f from noisy data. The error convergence rates with the uniform or random distribution are higher than the others in Figure 3. Elements in these two distribution patterns contain a similar number of data points and the effects of noise are cancelled out when data points are projected on the finite element grid. In comparison, the error with the uniform distribution with missing data and random normal distribution stops decreasing at

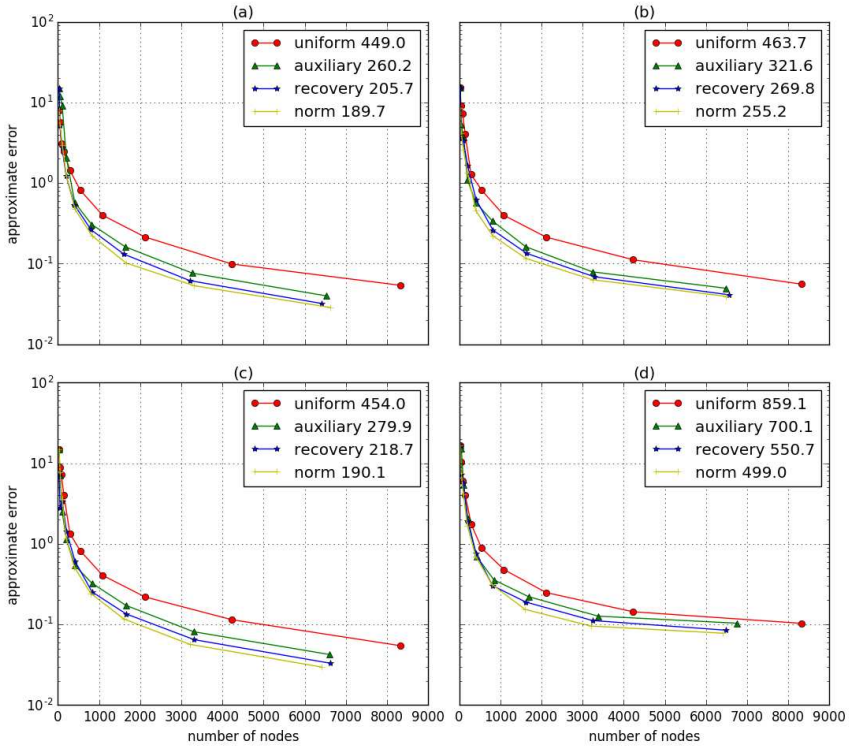


Fig. 3: Approximate error for (a) uniform distribution; (b) uniform distribution with missing data; (c) random distribution; and (d) random normal distribution.

the last two iterations. The error convergence rates with random normal distribution in Figure 3(d) are the lowest of four distribution patterns. Since some elements may contain few data points, it is more sensitive to noise in data, which leads to marked difference to f.

The three error indicators perform differently in the experiment. The performance of the auxiliary problem error indicator worsens when data is perturbed by noise, especially for random distribution in Figure 4(c). Since some elements may contain few data points, the accuracy of the local approximation is more susceptible to noise and may indicate large errors incorrectly. In contrast, the recovery-based error indicator and norm-based error indicator use $\boldsymbol{c}$ values to indicate large errors. Since the effects of the data distribution pattern and noise have been minimised by the TPSFEM, these two error indicators produce efficient adaptive grids for data sets with noise in Figure 4.

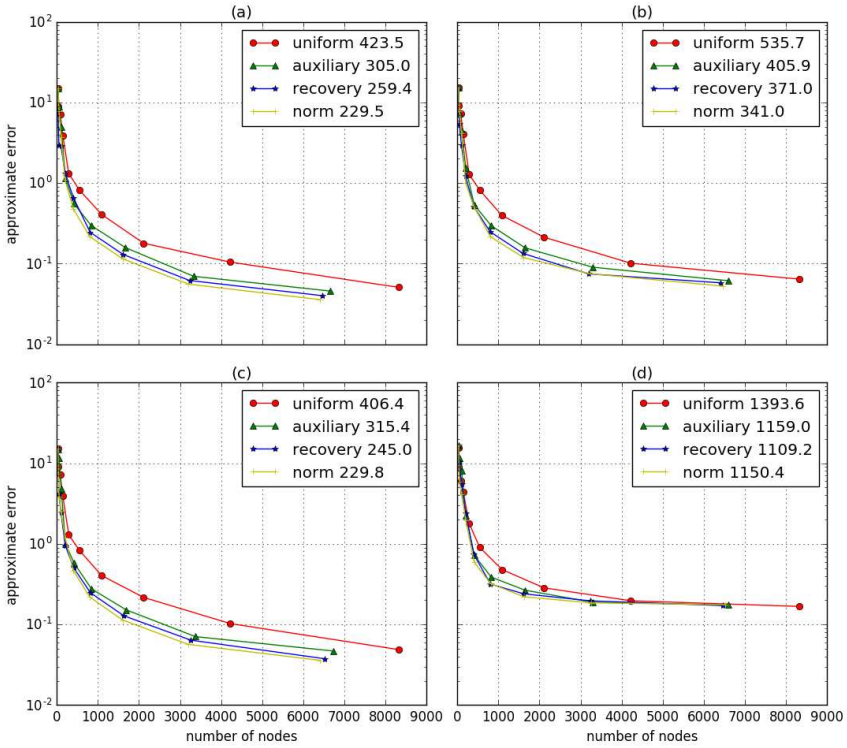


Fig. 4: Approximate error for data sets perturbed by noise with (a) uniform distribution; (b) uniform distribution with missing data; (c) random distribution; and (d) random normal distribution.

5 Conclusion

In this article, we explore four data distribution patterns and investigate their effects on the efficiency of adaptive grids generated using three error indicators. The four data distribution patterns lead to different maximum distances to data and affect the performance of the TPSFEM and its error indicators. While the TPSFEM may not restore the original function in regions with scarce data, it recovers a smooth surface to closely interpolate the data. Besides, the uniform and random distributions have close data densities across the domain and thus have less influence on the TPSFEM than the uniform distribution with missing data and random normal distribution. We also find that all the error indicators significantly improves the efficiency of adaptive grids with all data four distribution patterns. The auxiliary problem error indicator uses data for local approximations and is more vulnerable to changes in the data distribution patterns and noise. In contrast, the recovery-based error indicator and norm-based error indicator only use the information of the TPSFEM and are insensitive to these two factors.

References

1. M. Ainsworth and J. T. Oden. A posteriori error estimation in finite element analysis. *Computer Methods in Applied Mechanics and Engineering*, 142(1-2):1–88, 1997.
2. I. Babuvška and W. C. Rheinboldt. Error estimates for adaptive finite element computations. *SIAM Journal on Numerical Analysis*, 15(4):736–754, 1978.
3. J. S. Bendat and A. Piersol. *Random data: analysis and measurement procedures*, volume 729. 2011.
4. J. C. Carr, R. K. Beatson, J. B. Cherrie, T. J. Mitchell, W. R. Fright, B. C McCallum, and T. R. Evans. Reconstruction and representation of 3d objects with radial basis functions. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 67–76. ACM, 2001.
5. Z. M. Chen, R. Tuo, and W. L. Zhang. Stochastic convergence of a nonconforming finite element method for the thin plate spline smoother for observational data. *SIAM Journal on Numerical Analysis*, 56(2):635–659, 2018.
6. J. Duchon. Splines minimizing rotation-invariant semi-norms in sobolev spaces. pages 85–100, 1977.
7. L. Fang. *Error estimation and adaptive refinement of finite element thin plate spline*. PhD thesis, The Australian National University.
8. L. Fang and L. Stals. Adaptive discrete thin plate spline smoother. *ANZIAM Journal*, (Accepted).
9. M. F. Hutchinson. A stochastic estimator of the trace of the influence matrix for laplacian smoothing splines. *Communications in Statistics-Simulation and Computation*, 19(2):433–450, 1990.
10. W. F. Mitchell. A comparison of adaptive refinement techniques for elliptic problems. *ACM Transactions on Mathematical Software*, 15(4):326–347, 1989.
11. S. Roberts, M. Hegland, and I. Altas. Approximation of a thin plate spline smoother using continuous piecewise polynomial functions. *SIAM Journal on Numerical Analysis*, 41(1):208–234, 2003.
12. R. J. Schalkoff. *Digital image processing and computer vision*, volume 286. Wiley New York, 1989.
13. E. G. Sewell. *Analysis of a finite element method: PDE/PROTRAN*. Springer Science & Business Media, 2012.
14. L. Stals. Efficient solution techniques for a finite element thin plate spline formulation. *Journal of Scientific Computing*, 63(2):374–409, 2015.
15. L. Stals. Radial basis functions: Comparison between compact basis and finite element basis. (In preparation).
16. L. Stals and S. Roberts. Smoothing large data sets using discrete thin plate splines. *Computing and Visualization in Science*, 9(3):185–195, 2006.
17. H. Wendland. Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Advances in Computational Mathematics*, 4(1):389–396, 1995.
18. O. C. Zienkiewicz and J. Z. Zhu. A simple error estimator and adaptive procedure for practical engineerng analysis. *International Journal for Numerical Methods in Engineering*, 24(2):337–357, 1987.