

A direct prediction of the shape parameter in the collocation method of solving Poisson equation

Lin-Tian Luh*

Department of Data Science and Big Data Analytics, Providence University
Shalu, Taichung, Taiwan
Email: ltluh@pu.edu.tw

August 14, 2022

Abstract In this paper we totally discard the traditional trial-and-error algorithms of choosing acceptable shape parameter c in the multiquadrics $-\sqrt{c^2 + \|x\|^2}$ when dealing with differential equations. Instead, we choose c directly by the MN-curve theory and hence avoid the time-consuming steps of solving a linear system required by each trial of the c value in the traditional methods. The quality of the c value thus obtained is supported by the newly born choice theory of the shape parameter. Experiments show that the approximation error of the approximate solution to the differential equation is very close to the best approximation error among all possible choices of c .

Key words: radial basis function; multiquadric; shape parameter; collocation; Poisson equation

MSC: 31A30; 35J05; 35J25; 35J67; 35Q40; 35Q70; 65D05; 65L10; 65N35

1 Introduction

The generalized multiquadrics are defined as

$$\phi(x) := (-1)^{\lceil \beta/2 \rceil} (c^2 + \|x\|^2)^{\beta/2}, \quad \beta \in R \setminus 2N_{\geq 0}, \quad c > 0, \quad x \in R^d, \quad (1)$$

where $\lceil \beta/2 \rceil$ denotes the smallest integer greater than or equal to $\beta/2$ and the constant c is called the shape parameter. These are the most popular radial basis functions (RBFs) and frequently used in the collocation method of solving partial differential equations. In this paper we let $\beta = 1$. In the collocation method an approximate solution to a differential equation is of the form

$$\hat{u}(x) := \sum_{i=1}^N \lambda_i \phi(x - x_i) + p(x), \quad (2)$$

where $p(x) \in P_{m-1}$, the space of polynomials of degree less than or equal to $m-1$ in R^d , and $X = \{x_1, \dots, x_N\}$ is a set of points scattered in the domain. For $m = 0$, $P_{m-1} := \{0\}$. The integer $m := \lceil \beta/2 \rceil$. Since we let $\beta = 1$, here $m = 1$ and $p(x)$ is a constant λ_0 . An unorthodox way even drops λ_0 and lets it be 0. The constants λ_i , $i = 0, \dots, N$, are chosen so that $\hat{u}(x)$ satisfies

*corresponding author

the differential equation (including boundary conditions) at the points x_i , $i = 1, \dots, N$, called the collocation points.

The function $\hat{u}(x)$ originates from the interpolation theory of the radial basis functions, where $\hat{u}(x)$ interpolates a given function $f(x)$ at x_1, \dots, x_N . It is required that $\sum_{i=1}^N \lambda_i p_l(x_i) = 0$ for $l = 1, \dots, Q$, where $\{p_1, \dots, p_Q\}$ is a basis of P_{m-1} . Besides this, the only requirement for $X = \{x_1, \dots, x_N\}$ is that it should be P_{m-1} -unisolvent. Namely, if $p \in P_{m-1}$ and $p(x_i) = 0$ for $i = 1, \dots, N$, then $p(x)$ is a zero polynomial. Further details can be seen in Section 8.5 of Wendland [1]. This origin has to be mentioned because we need it in the design of $\hat{u}(x)$.

This approach of solving differential equation was first introduced by E. Kansa [2, 3]. A huge amount of experiments show that it works very well. The main advantage of this approach, namely the collocation method (or Kansa method), is that the data points are scattered in the domain without meshes. Moreover, its high accuracy is also very attractive. However, the choice of the shape parameter c contained in the multiquadrics is a big problem. Experts in this field only know that it is very influential, but do not know how to choose it. G. Fasshauer [4] pointed out that only trial-and-error algorithms were available. Some people even give up using multiquadrics to solve differential equations due to this problem. This greatly lowers the power of the collocation method. The purpose of this paper is to handle this problem.

2 Materials and Methods

2.1 Sobolev error estimates

We need a space which plays an intermediate role in our approximating work.

Definition 2.1 For any positive number γ ,

$$B_\gamma := \left\{ f \in L^2(R^d) : \hat{f}(\xi) = 0 \text{ if } \|\xi\| > \gamma \right\}$$

where \hat{f} denotes the Fourier transform of f . For each $f \in B_\gamma$, its norm is

$$\|f\|_{B_\gamma} := \left\{ \int |\hat{f}(\xi)|^2 d\xi \right\}^{1/2}.$$

Here, $f \in L^2(R^d)$ means that $f(x)^2$ is integrable. Our main interest is in the Sobolev space $W_2^\tau(R^d)$ because it contains the solutions of a lot of important differential equations. The Sobolev space is defined as follows.

Definition 2.2 For any positive integer τ ,

$$W_2^\tau(R^d) = \left\{ f \in L^2(R^d) : D^k f(x_1, \dots, x_d) \in L^2(R^d) \text{ for } |k| \leq \tau \right\}$$

where $k = (k_1, \dots, k_d)$ and $|k| = \sum_{i=1}^d k_i$. The Sobolev norm is

$$\|f\|_{W_2^\tau(R^d)} := \left\{ \sum_{0 \leq |k| \leq \tau} \int |D^k f(x)|^2 dx \right\}^{1/2}.$$

In the preceding definition the derivatives are more general than the classical derivatives, called distributional derivatives. The vector k can even have negative or noninteger coordinates. Further details can be seen in Yosida [5] and any textbooks of functional analysis.

Before introducing Sobolev error estimates, some necessary ingredients should be defined. Suppose $X = \{x_1, \dots, x_N\} \subseteq \Omega$ is a finite subset of a bounded set Ω in R^d . Then the separation radius is defined by

$$q_X := \frac{1}{2} \min_{i \neq j} \|x_i - x_j\|.$$

Then we have the following core theorem which is just Theorem 3.4 of Narcowich et al. [6].

Theorem 2.3 *Let $\beta, t \in R$ satisfy $\beta > d/2$ and $t \geq 0$. If $f \in W_2^{\beta+t}(R^d)$, then there exists $f_\gamma \in B_\gamma$ such that $f_\gamma|_X = f|_X$ and*

$$\|f - f_\gamma\|_{W_2^\beta(R^d)} \leq 5 \cdot \kappa^{-t} q_X^t \|f\|_{W_2^{\beta+t}(R^d)}, \quad (3)$$

with $\gamma = \kappa/q_X$, where $\kappa \geq 1$ depends only on β and d .

This theorem clearly shows that any function in the Sobolev space, possibly a solution to an important differential equation, can be interpolated by a B_γ function with a good error bound. In the next subsection we will show that any B_γ function can be interpolated by a function of the form (2), also with a good error bound.

2.2 MN-curve theory

We need some basic definitions.

Definition 2.4 *For any positive number σ ,*

$$E_\sigma := \left\{ f \in L^2(R^d) : \int |\hat{f}(\xi)|^2 e^{|\xi|^2/\sigma} d\xi < \infty \right\}$$

where \hat{f} denotes the Fourier transform of f . For each $f \in E_\sigma$, its norm is

$$\|f\|_{E_\sigma} := \left\{ \int |\hat{f}(\xi)|^2 e^{|\xi|^2/\sigma} d\xi \right\}^{1/2}.$$

Obviously, for any $\gamma, \sigma > 0$, $B_\gamma \subseteq E_\sigma$. We are going to show how E_σ functions can be approximated by functions of the form (2).

For any set $\Omega \subseteq R^d$ and any set $X = \{x_1, \dots, x_N\}$ of sample points contained in Ω , the fill distance is defined by

$$\delta(\Omega, X) := \sup_{x \in \Omega} \inf_{i=1, \dots, N} \|x - x_i\|,$$

abbreviated as δ , which measures the spacing of the sample points in Ω . The smaller δ is, the more sample points are needed. In this paper Ω denotes the function domain.

Definition 2.5 *Let d and β be as in (1). The numbers ρ and Δ_0 are defined as follows.*

(a) *Suppose $\beta < d - 3$. Let $s = \lceil (d - \beta - 3)/2 \rceil$. Then*

(i) *if $\beta < 0$, $\rho = (3 + s)/3$ and $\Delta_0 = \frac{(2+s)(1+s) \cdots 3}{\rho^2}$;*

(ii) *if $\beta > 0$, $\rho = 1 + \frac{s}{2\lceil \beta/2 \rceil + 3}$ and $\Delta_0 = \frac{(2m+2+s)(2m+1+s) \cdots (2m+3)}{\rho^{2m+2}}$ where $m = \lceil \beta/2 \rceil$.*

(b) Suppose $d - 3 \leq \beta < d - 1$. Then $\rho = 1$ and $\Delta_0 = 1$.

(c) Suppose $\beta \geq d - 1$. Let $s = -\lceil (d - \beta - 3)/2 \rceil$. Then

$$\rho = 1 \text{ and } \Delta_0 = \frac{1}{(2m+2)(2m+1) \cdots (2m-s+3)} \text{ where } m = \lceil \beta/2 \rceil.$$

For any $f \in E_\sigma$ and $x \in \Omega$, the upper bound of $|f(x) - \hat{u}(x)|$ is a very complicated expression involving both ρ and Δ_0 , as can be seen in Luh [7]. A modified theory for purely scattered data setting can be seen in Luh [8]. In this paper we only need to extract its essential part

$$|f(x) - \hat{u}(x)| \leq \text{MN}(c) \|f\|_{E_\sigma} \quad (4)$$

where c is just the shape parameter defined in (1) and $\text{MN}(c)$ is the MN function to be defined below, as in [8].

In the MN-curve theory, we require that the diameter r of the function domain Ω satisfy $b_0/2 \leq r \leq b_0$ where b_0 is a parameter determined by us. Once b_0 is fixed, there are three cases for the definition of $\text{MN}(c)$.

Case 1 : $\beta < 0$, $|d + \beta| \geq 1$ and $d + \beta + 1 \geq 0$ Let $f \in E_\sigma$ and $\phi(x)$ be as in (1). For any fixed fill distance δ satisfying $0 < \delta \leq b_0/2$, the optimal value of c in $[24\rho\delta, \infty)$ is the number minimizing

$$\text{MN}(c) := \begin{cases} \sqrt{8\rho} c^{(\beta-d-1)/4} \left\{ (\xi^*)^{(d+\beta+1)/2} e^{c\xi^* - (\xi^*)^2/\sigma} \right\}^{1/2} \left(\frac{2}{3}\right)^{c/(24\rho\delta)} & \text{if } c \in [24\rho\delta, 12b_0\rho), \\ \sqrt{\frac{2}{3b_0}} c^{(\beta-d+1)/4} \left\{ (\xi^*)^{(d+\beta+1)/2} e^{c\xi^* - (\xi^*)^2/\sigma} \right\}^{1/2} \left(\frac{2}{3}\right)^{b_0/(2\delta)} & \text{if } c \in [12b_0\rho, \infty), \end{cases}$$

where

$$\xi^* = \frac{c\sigma + \sqrt{c^2\sigma^2 + 4\sigma(d + \beta + 1)}}{4}.$$

Remarks: Note that $\lim_{c \rightarrow 0^+} \text{MN}(c) = \infty$ and $\lim_{c \rightarrow \infty} \text{MN}(c) = \infty$.

Case 2 : $\beta = -1$ and $d = 1$ Let $f \in E_\sigma$ and $\phi(x)$ be as in (1). For any fill distance δ satisfying $0 < \delta \leq b_0/2$, the optimal value of c in $[24\rho\delta, \infty)$ is the number minimizing

$$\text{MN}(c) := \begin{cases} \sqrt{8\rho} c^{(\beta-1)/2} \left\{ \frac{1}{\ln 2} + 2\sqrt{3}\text{M}(c) \right\}^{1/2} \left(\frac{2}{3}\right)^{c/(24\rho\delta)} & \text{if } c \in [24\rho\delta, 12b_0\rho), \\ \sqrt{\frac{2}{3b_0}} c^{\beta/2} \left\{ \frac{1}{\ln 2} + 2\sqrt{3}\text{M}(c) \right\}^{1/2} \left(\frac{2}{3}\right)^{b_0/(2\delta)} & \text{if } c \in [12b_0\rho, \infty), \end{cases}$$

where

$$\text{M}(c) := \begin{cases} e^{1-1/(c^2\sigma)} & \text{if } 0 < c \leq \frac{2}{\sqrt{3\sigma}}, \\ g\left(\frac{c\sigma + \sqrt{c^2\sigma^2 + 4\sigma}}{4}\right) & \text{if } \frac{2}{\sqrt{3\sigma}} < c, \end{cases}$$

g being defined by $g(\xi) := \sqrt{c\xi} e^{c\xi - \xi^2/\sigma}$.

Remarks: The same as Case 1, we have $\lim_{c \rightarrow 0^+} \text{MN}(c) = \infty$ and $\lim_{c \rightarrow \infty} \text{MN}(c) = \infty$.

Case 3 : $\beta > 0$ and $d \geq 1$ Let $f \in E_\sigma$ and $\phi(x)$ be as in (1). For any fixed fill distance δ satisfying $0 < \delta \leq b_0/2$, the optimal value of c in $[24\rho\delta, \infty)$ is the number minimizing

$$\text{MN}(c) := \begin{cases} \sqrt{8\rho} c^{(\beta-d-1)/4} \left\{ \frac{(\xi^*)^{(1+\beta+d)/2} e^{c\xi^*}}{e^{(\xi^*)^2/\sigma}} \right\}^{1/2} \left(\frac{2}{3}\right)^{c/(24\rho\delta)} & \text{if } c \in [24\rho\delta, 12b_0\rho), \\ \sqrt{\frac{2}{3b_0}} c^{(1+\beta-d)/4} \left\{ \frac{(\xi^*)^{(1+\beta+d)/2} e^{c\xi^*}}{e^{(\xi^*)^2/\sigma}} \right\}^{1/2} \left(\frac{2}{3}\right)^{b_0/(2\delta)} & \text{if } c \in [12b_0\rho, \infty), \end{cases}$$

where

$$\xi^* = \frac{c\sigma + \sqrt{c^2\sigma^2 + 4\sigma(1 + \beta + d)}}{4}.$$

Remarks: (a) If $\beta - d - 1 > 0$, $\lim_{c \rightarrow 0^+} \text{MN}(c) = 0$. (b) If $\beta - d - 1 < 0$, $\lim_{c \rightarrow 0^+} \text{MN}(c) = \infty$. (c) If $\beta - d - 1 = 0$, $\lim_{c \rightarrow 0^+} \text{MN}(c)$ is a finite positive number. (d) $\lim_{c \rightarrow \infty} \text{MN}(c) = \infty$. Practically, we never let $\beta - d - 1 \geq 0$. Hence (a) and (c) will not happen.

In all the three cases the requirement $c \geq 24\rho\delta$ is harmless because the value $24\rho\delta$ is usually quite small and a huge amount of experiments show that the optimal choice of c never lies in the interval $(0, 24\rho\delta)$.

The function $\hat{u}(x)$ in (4) in fact interpolates $f(x)$ at the sample points. Hence strictly speaking $\text{MN}(c)$ can only be used to measure the quality of the function interpolation. Nevertheless, collocation is in spirit a kind of interpolation, not just approximation. We require that $\hat{u}(x)$ satisfy the given differential equation at the sample points. Theoretically, the value c minimizing $\text{MN}(c)$ should also be the optimal choice of c in the definition of $\hat{u}(x)$ when dealing with differential equations. Another point is that in (4) it is required that $f \in E_\sigma$ which is mathematically just a subset of the Sobolev space $W_2^\tau(R^d)$. Fortunately, the formula (3) offers a bridge for our approximating $f \in W_2^\tau(R^d)$ with $\hat{u}(x)$. For a fixed set $X = \{x_1, \dots, x_N\}$ of sample points in $\Omega \subseteq R^d$, any $f \in W_2^\tau(R^d)$ can be interpolated by an $f_\gamma \in B_\gamma \subseteq E_\sigma$. Then f_γ can be interpolated by \hat{u} . Both happen in X . The influence of c happens in the latter part of the interpolations. Hence we can directly predict the optimal value of c with the curve of $\text{MN}(c)$. No search is needed.

2.3 Problem setting

We try to handle Poisson equations. A standard 3D Poisson equation is of the form

$$\begin{cases} u_{xx}(x, y, z) + u_{yy}(x, y, z) + u_{zz}(x, y, z) = f(x, y, z) & \text{for } (x, y, z) \in \Omega \setminus \partial\Omega, \\ u(x, y, z) = g(x, y, z) & \text{for } (x, y, z) \in \partial\Omega, \end{cases} \quad (5)$$

where Ω is the domain with boundary $\partial\Omega$, and f, g are given functions. A natural extension to d dimensions can be easily understood by replacing (x, y, z) with (x_1, \dots, x_d) and letting $\Omega \subseteq R^d$.

Our goal is to find an approximate solution $\hat{u}(x)$ of the form (2). Our approach is to find real numbers $\lambda_0, \dots, \lambda_N$ such that $\hat{u}(x) := \sum_{i=1}^N \lambda_i \phi(x - x_i) + \lambda_0$ satisfies (5) in R^d at the sample points $x_1, \dots, x_N \in R^d$, called collocation points. The constant c appearing in ϕ is determined by the MN-curve theory. Of course this process involves solving a system of linear equations with unknowns $\lambda_0, \dots, \lambda_N$, by requiring that

$$\begin{cases} \sum_{j=1}^N \lambda_j \mathcal{L}[\phi(x_i - x_j)] = f(x_i) & \text{for } i = 1, \dots, N_{int}, \\ \sum_{j=1}^N \lambda_j = 0, \\ \sum_{j=1}^N \lambda_j \phi(x_i - x_j) + \lambda_0 = g(x_i) & \text{for } i = N_{int} + 1, \dots, N, \end{cases} \quad (6)$$

where \mathcal{L} denotes the differential operator of the Poisson equation. The sample points $x_i, \dots, x_{N_{int}}$ are located in the interior of the domain cube, and $x_{N_{int}+1}, \dots, x_N$ on the boundary. The requirement $\sum_{j=1}^N \lambda_j = 0$ results from the interpolation theory, as explained in the introduction section. We thus have an $(N + 1) \times (N + 1)$ system of linear equations.

Although its coefficient matrix is not sparse, it can be efficiently solved because its scale is not large, as long as the shape parameter c is well chosen. For a long time, the RBF collocation method has been severely criticized for the full matrix induced by the multiquadrics. Fortunately now we know that the amount of sample points needed can be greatly reduced by choosing the shape parameter c according to our theory. This is exciting. As for the collocation points, they are scattered in the domain and boundary without meshes. Theoretically, if the exact solution $u(x)$ lies in the Sobolev space, the approximate solution $\hat{u}(x)$ thus found should be quite good. Experiments show that the approximation error $|u(x) - \hat{u}(x)|$ is indeed very small.

3 Results

The crux of this approach is the choice of the parameter σ in the definition of $MN(c)$. Once $f_\gamma \in B_\gamma$ in (3) is given, in order to measure $|f_\gamma(x) - \hat{u}(x)|$, there are infinitely many possible choices of σ for the implementation of the inequality (4). If σ is too large, the value of $MN(c)$ will be very large, making the MN curve meaningless. If σ is very small, the value of $MN(c)$ is usually very small. However, $\|f\|_{E_\sigma}$ will become extremely large, also making (4) meaningless. Fortunately, after analyzing the MN curves, we can always find a suitable σ without much effort, as shown in our experiments.

3.1 1D model

Although our main interest is in two and three-dimensional problems, in order to help the reader use our approach, we still include the 1D problem.

Let $u(x) = e^{-x}$ with domain $\Omega = \{x : 0 \leq x \leq 1\}$ be the test function. We are going to solve

$$u_{xx}(x) = e^{-x}$$

for $0 < x < 1$ with $u(0) = 1$ and $u(1) = e^{-1}$.

The approximate solution will be of the form $\hat{u}(x) = \sum_{i=1}^N \phi(x-x_i) + \lambda_0$ where $\phi(x) = -\sqrt{c^2 + x^2}$ and λ_i , $i = 0, \dots, N$, are constants to be determined. The sample points x_i , $i = 1, \dots, N$, are randomly generated and scattered in the unit interval except that two of them are 0 and 1, respectively. The choice of c will be made according to Case 3 of the MN curves.

The MN function value greatly depends on the parameter σ in the definition of $MN(c)$. We present three curves for $\sigma = 10^{-1}$ and three for $\sigma = 10^{-5}$ in Figs. 1-3 and 4-6, respectively. In the figures, δ denotes the fill distance, b_0 denotes the domain diameter, d is the dimension, and β is the parameter in the definition of the multiquadrics (1).

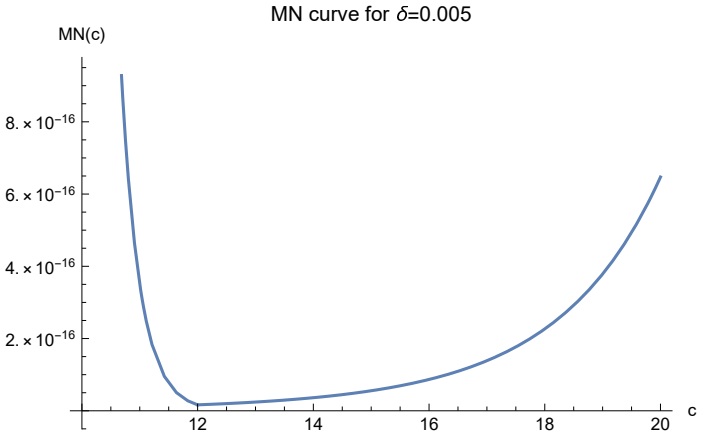


Figure 1: Here $d = 1$, $\beta = 1$, $b_0 = 1$ and $\sigma = 1E-1$.

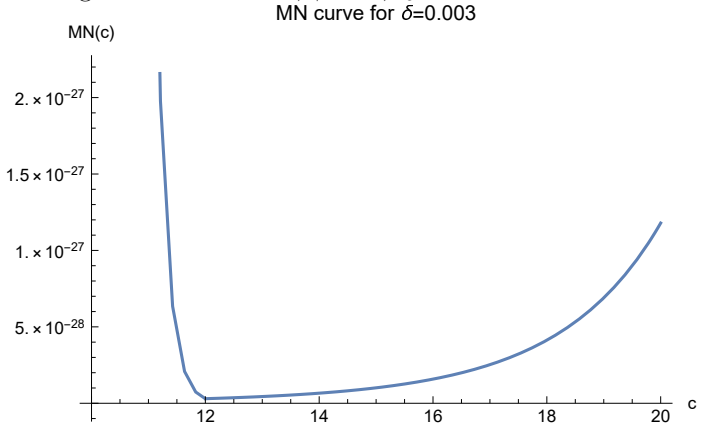


Figure 2: Here $d = 1$, $\beta = 1$, $b_0 = 1$ and $\sigma = 1E-1$.

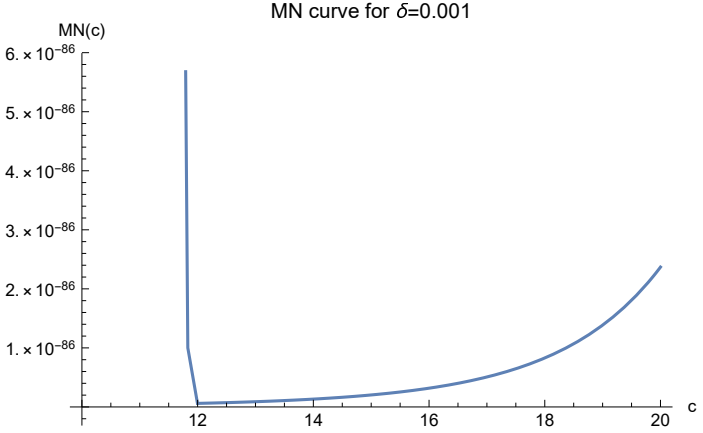


Figure 3: Here $d = 1$, $\beta = 1$, $b_0 = 1$ and $\sigma = 1E-1$.

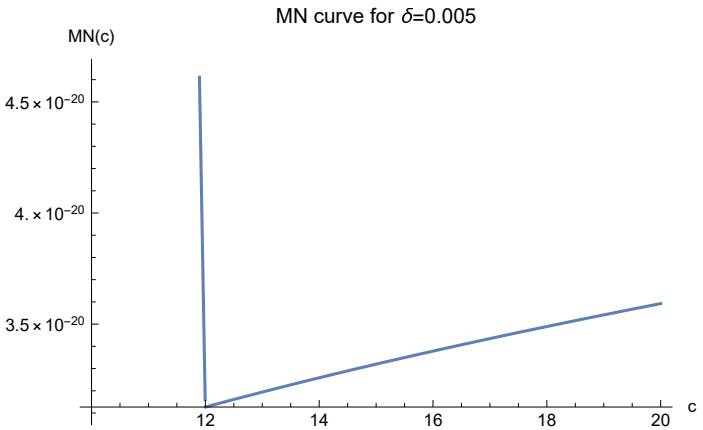


Figure 4: Here $d = 1$, $\beta = 1$, $b_0 = 1$ and $\sigma = 1E-5$.

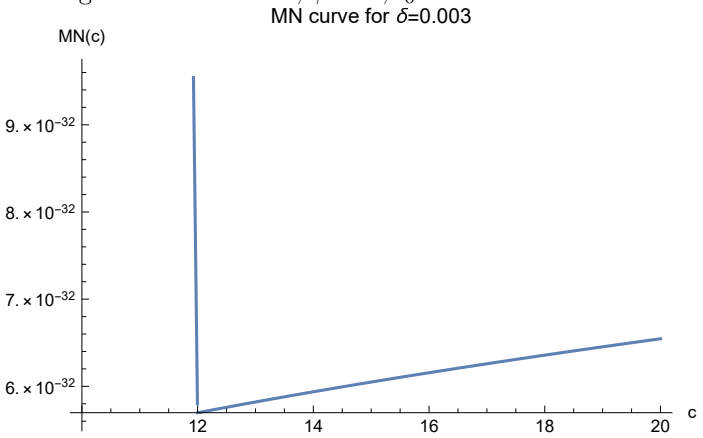


Figure 5: Here $d = 1$, $\beta = 1$, $b_0 = 1$ and $\sigma = 1E-5$.

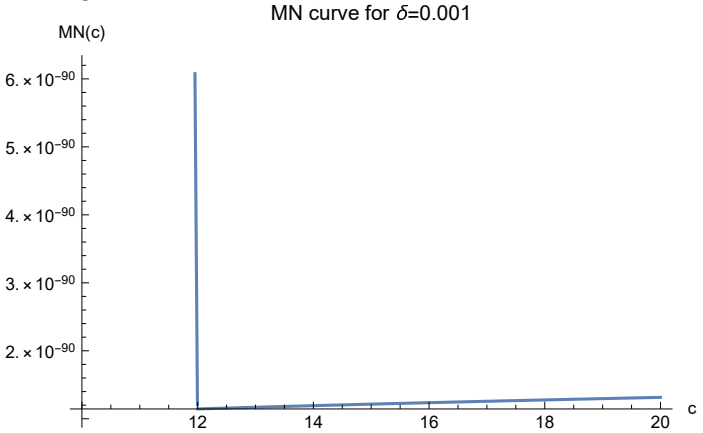


Figure 6: Here $d = 1$, $\beta = 1$, $b_0 = 1$ and $\sigma = 1E-5$.

Note that in these figures $c = 12$ always corresponds to the minimal value of $MN(c)$. It suggests that one should choose $c = 12$ in our approximate solution to the Poisson equation. We present the experimental results in Tab. 1. In the table, N_d denotes the number of data points used, and N_{int} , N_{bdy} denote the number of interior and boundary data points, respectively. We use N_t test points to measure the quality of the approximation

$$\text{RMS} := \left\{ \left(\sum_{i=1}^{N_t} |u(x) - \hat{u}(x_i)|^2 \right) / N_t \right\}^{1/2}.$$

As in our previous papers, b_0 denotes the diameter of the domain, and COND denotes the condition number of the system of the linear equations (6).

In order to cope with the problem of ill-conditioning, enough effective digits were adopted for each step of the calculations. For example, for $N_d = 82$, we adopted 300 effective digits to the right of the decimal point to handle its huge corresponding condition number. Even so, it took only one second for the computer to solve the linear system. All these were achieved in virtue of the computer software Mathematica.

If c is chosen arbitrarily, say $c = 1$, then the RMS will be $4.3 \cdot 10^{-30}$ for $N_d = 82$. As for the comparison with other choices of c values, we are not going to present in this subsection for three reasons. Firstly, the results for $c = 12$ are already quite good. Secondly, in our theory, the prediction is reliable only when enough data points are used, and then the condition number will be extremely large for $d = 1$ and $b_0 = 1$. For example, experimentally, we found that the optimal value of c is 1600 for $N_d = 162$, when $\text{COND} = 1.5 \cdot 10^{1211}$ and $\text{RMS} = 7.96 \cdot 10^{-359}$. In order to get the predicted value $c = 12$, we have to increase the number of data points so that $N_d \gg 162$, as in the experiments for interpolation. Then the condition number will become much larger than $1.5 \cdot 10^{1211}$, although the RMS will be smaller than $7.96 \cdot 10^{-359}$. In practice we don't need such accuracy. Thirdly, such comparisons can be perfectly handled in the 2D and 3D experiments.

Table 1: 1D experiment, $c = 12$, $b_0 = 1$, $N_t = 501$.

N_d	7	12	22	42	82
N_{int}	5	10	20	40	80
N_{bdy}	2	2	2	2	2
RMS	$3.2 \cdot 10^{-6}$	$2.5 \cdot 10^{-11}$	$2.7 \cdot 10^{-24}$	$4.9 \cdot 10^{-52}$	$3.2 \cdot 10^{-127}$
COND	$1.3 \cdot 10^{23}$	$1.4 \cdot 10^{39}$	$6.8 \cdot 10^{69}$	$3.3 \cdot 10^{133}$	$2.0 \cdot 10^{281}$

3.2 2D model

Our test function is now $u(x, y) = e^{-x-y}$ on $\Omega = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1\}$. This function obviously lies in Sobolev space. The Poisson equation is then

$$\begin{cases} u_{xx}(x, y) + u_{yy}(x, y) = 2e^{-x-y} & \text{for } (x, y) \in \Omega \setminus \partial\Omega, \\ u(x, y) = g(x, y) & \text{for } (x, y) \in \partial\Omega, \end{cases} \quad (7)$$

where $g(0, y) = e^{-y}$, $g(x, 0) = e^{-x}$, $g(1, y) = e^{-1-y}$, $g(x, 1) = e^{-x-1}$ for $0 \leq x \leq 1$ and $0 \leq y \leq 1$.

We are trying to find an approximate solution $\hat{u}(x, y) = \sum_{i=1}^N \lambda_i \phi(x - x_i, y - y_i) + \lambda_0$ where $\phi(x, y) = -\sqrt{c^2 + x^2 + y^2}$ and λ_i , $i = 0, \dots, N$, are constants to be determined. The sample points (x_i, y_i) , $i = 1, \dots, N$, are scattered in the domain Ω . Our focus is the choice of c .

Now let us analyze the MN curves. We first let $\sigma = 10^{-1}$. and list five MN curves with different fill distances δ .

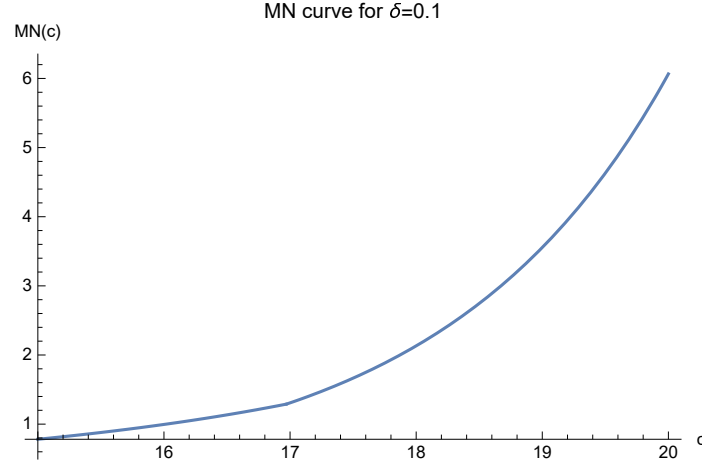


Figure 7: Here $d = 2$, $\beta = 1$, $b_0 = \sqrt{2}$ and $\sigma = 1\text{E}-1$.

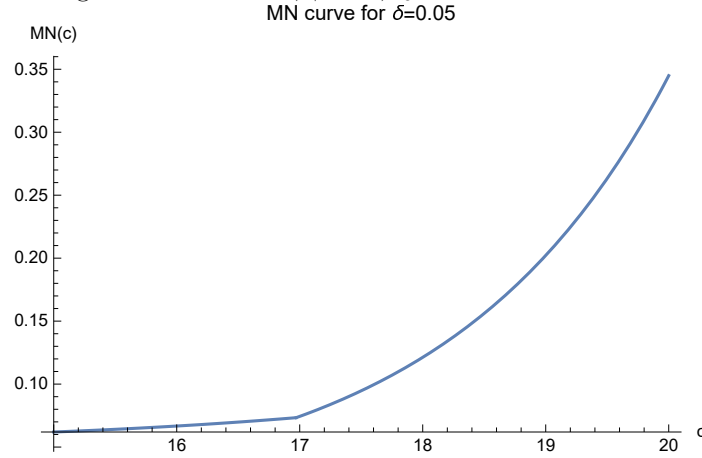


Figure 8: Here $d = 2$, $\beta = 1$, $b_0 = \sqrt{2}$ and $\sigma = 1\text{E}-1$.

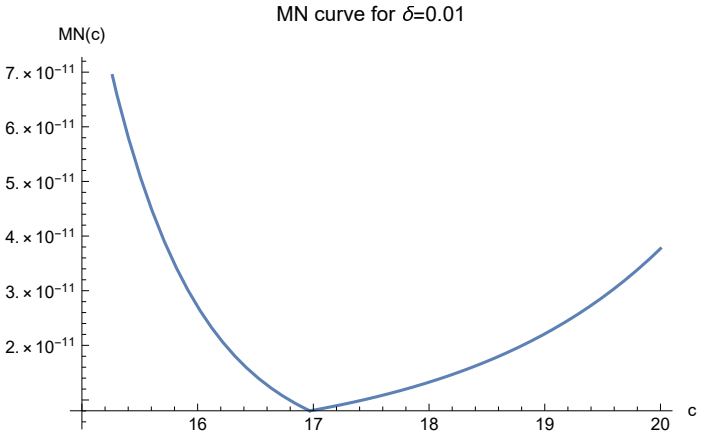


Figure 9: Here $d = 2$, $\beta = 1$, $b_0 = \sqrt{2}$ and $\sigma = 1\text{E}-1$.

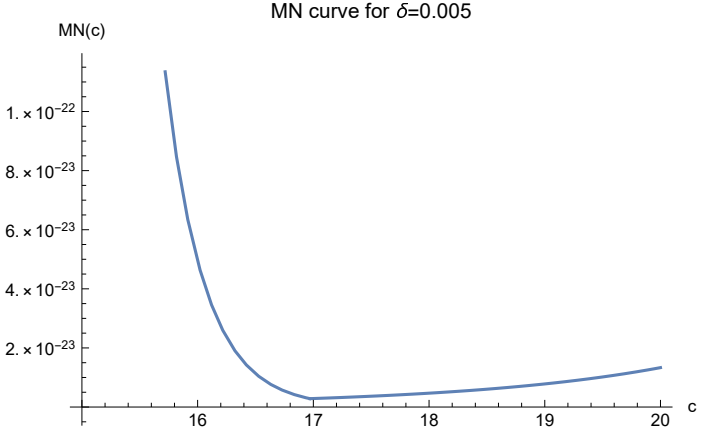


Figure 10: Here $d = 2$, $\beta = 1$, $b_0 = \sqrt{2}$ and $\sigma = 1\text{E}-1$.

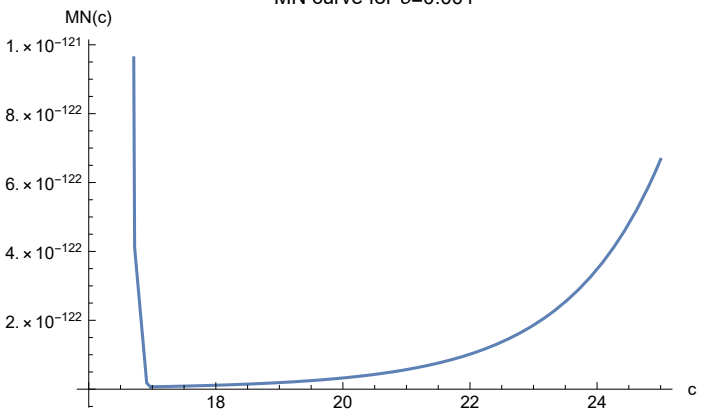


Figure 11: Here $d = 2$, $\beta = 1$, $b_0 = \sqrt{2}$ and $\sigma = 1\text{E}-1$.

In Figs. 7-11 it is clearly seen that as the fill distances decrease, the lowest points of the curves move to a fixed value $c = 17$ which is just $12b_0\rho$ in Case 3 of the definition of $MN(c)$. Now we investigate $\sigma = 10^{-5}$.

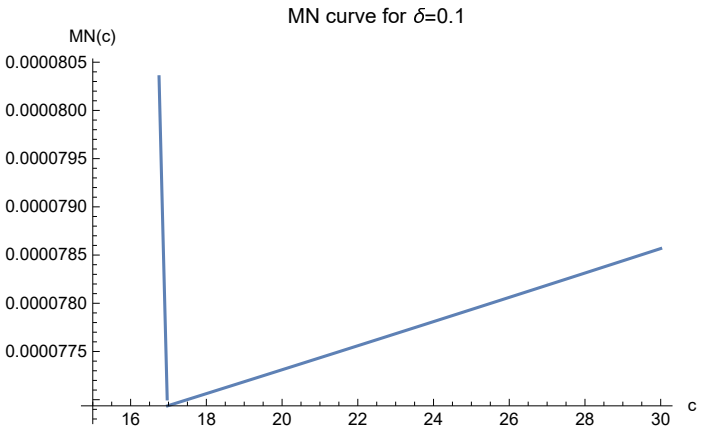


Figure 12: Here $d = 2$, $\beta = 1$, $b_0 = \sqrt{2}$ and $\sigma = 1E-5$.

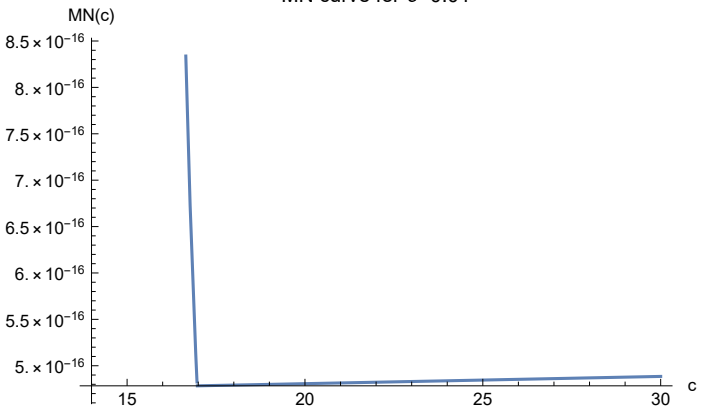


Figure 13: Here $d = 2$, $\beta = 1$, $b_0 = \sqrt{2}$ and $\sigma = 1E-5$.

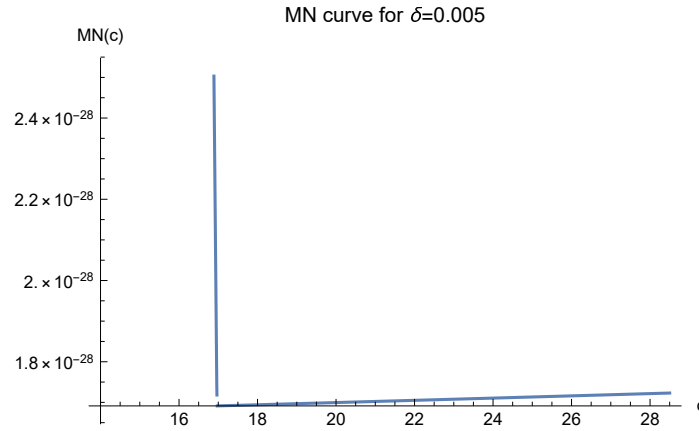


Figure 14: Here $d = 2$, $\beta = 1$, $b_0 = \sqrt{2}$ and $\sigma = 1\text{E}-5$.

Figs. 12-14 also show that the optimal choice of c is 17. In fact, if we test other σ 's, the same result will appear. In order to save space, we do not list them.

All the MN curves strongly suggest that one should choose $c = 17$ whenever enough data points are used. So we investigate its quality and present the results in Tab. 2. Here N_{bdy} and N_{int} denote the numbers of data points located on the boundary and interior of the domain, respectively. Then N_d and N_t denote the total numbers of data points and test points, respectively. The root-mean-square error, used to measure the approximation error, is defined by

$$\text{RMS} := \left\{ \left(\sum_{i=1}^{N_t} |u(x_i, y_i) - \hat{u}(x_i, y_i)|^2 \right) / N_t \right\}^{1/2}.$$

As before, b_0 denotes the diameter of the domain Ω and COND is the condition number of the linear system involved. The most time-consuming work of solving the linear system took only two seconds for 341 data points. Hence we did not put the computer time into the table.

Table 2: 2D experiment, $c = 17$, $b_0 = \sqrt{2}$, $N_t = 961$.

N_d	46	91	141	191	341
N_{int}	5	50	100	150	300
N_{bdy}	41	41	41	41	41
RMS	$2.3 \cdot 10^{-5}$	$2.4 \cdot 10^{-12}$	$2.4 \cdot 10^{-16}$	$7.2 \cdot 10^{-20}$	$9.2 \cdot 10^{-25}$
COND	$1.0 \cdot 10^{40}$	$6.5 \cdot 10^{44}$	$5.4 \cdot 10^{53}$	$1.3 \cdot 10^{64}$	$4.8 \cdot 10^{87}$

For simplicity, the test points were evenly spaced in the domain. The interior data points were purely scattered and generated randomly by Mathematica in the domain, but the boundary

Table 3: 2D experiment, $b_0 = \sqrt{2}$, $N_d = 341$, $N_t = 961$.

c	1	10	17	30	50
RMS	$1.1 \cdot 10^{-8}$	$5.2 \cdot 10^{-22}$	$9.2 \cdot 10^{-25}$	$1.4 \cdot 10^{-24}$	$4.1 \cdot 10^{-24}$
COND	$1.8 \cdot 10^{28}$	$7.0 \cdot 10^{75}$	$4.8 \cdot 10^{87}$	$1.6 \cdot 10^{100}$	$2.0 \cdot 10^{111}$
c	70	90	110	130	150
RMS	$2.2 \cdot 10^{-23}$	$2.4 \cdot 10^{-22}$	$3.2 \cdot 10^{-22}$	$2.8 \cdot 10^{-22}$	$3.6 \cdot 10^{-22}$
COND	$1.2 \cdot 10^{119}$	$1.8 \cdot 10^{125}$	$2.9 \cdot 10^{129}$	$5.9 \cdot 10^{132}$	$5.8 \cdot 10^{135}$

data points were evenly spaced just for ease of programming. The problem of ill-conditioning was overcome by keeping enough effective digits to the right of the decimal point for each step of the computation. For example, when $N_d = 341$, we adopted 200 digits and successfully defeated the large condition number 4.8×10^{87} . In fact, 110 digits are already good enough and will lead to the same result. Even 200 digits, it took only two seconds to solve the linear system. All these were achieved by the help of the arbitrarily precise computer software Mathematica.

Although $c = 17$ leads to satisfactory results, a comparison with other choices of c is also needed. Tab. 3 offers such a comparison. We fix $N_{int} = 300$ and $N_{bdy} = 41$ for all choices of c .

In Tab. 3 it is clear that our theoretically predicted optimal value $c = 17$ coincides exactly with the experimentally optimal value. Also, as depicted by the MN curves in Figs. 12-14, the approximation errors get large very slowly for $c > 17$. This is fully reflected by our experimental results.

The three-dimensional experiment is more challenging and expected to be much more time-consuming. Fortunately it takes only five minutes to compute the linear system, as we shall see in the next subsection.

3.3 3D model

The test function is now $u(x, y, z) = e^{-x-y-z}$ on $\Omega = \{(x, y, z) : 0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq 1\}$. It lies in Sobolev space. The Poisson equation is

$$\begin{cases} u_{xx}(x, y, z) + u_{yy}(x, y, z) + u_{zz}(x, y, z) = 3e^{-x-y-z} & \text{for } (x, y, z) \in \Omega \setminus \partial\Omega, \\ u(x, y, z) = g(x, y, z) & \text{for } (x, y, z) \in \partial\Omega, \end{cases} \quad (8)$$

where $g(x, y, z)$ is just the restriction of $u(x, y, z)$ to the six surfaces of the domain cube. In other words, the Dirichlet condition is adopted.

The approximate solution will be of the form $\hat{u}(x, y, z) = \sum_{i=1}^N \lambda_i \phi(x - x_i, y - y_i, z - z_i) + \lambda_0$ where $\phi(x, y, z) = -\sqrt{c^2 + x^2 + y^2 + z^2}$ and λ_i , $i = 0, \dots, N$, are constants to be determined. The sample points (x_i, y_i, z_i) , $i = 1, \dots, N$, are still scattered in the interior of the domain cube and evenly spaced on the boundary. We are going to find c such that the approximation error $|u(x, y, z) - \hat{u}(x, y, z)|$ is as small as possible.

In order to find a suitable c , one has to analyze the MN curves first. Again, Case 3 of $MN(c)$ applies. However, for 3D MN functions, different σ 's indicate different optimal values of c . For $\sigma = 10^{-1}$, the optimal values of c are shown in Figs. 15-19. These figures show that as long as δ is small enough, the optimal value of c is 20.7846 which is just $12b_0\rho$ in the definition of $MN(c)$.

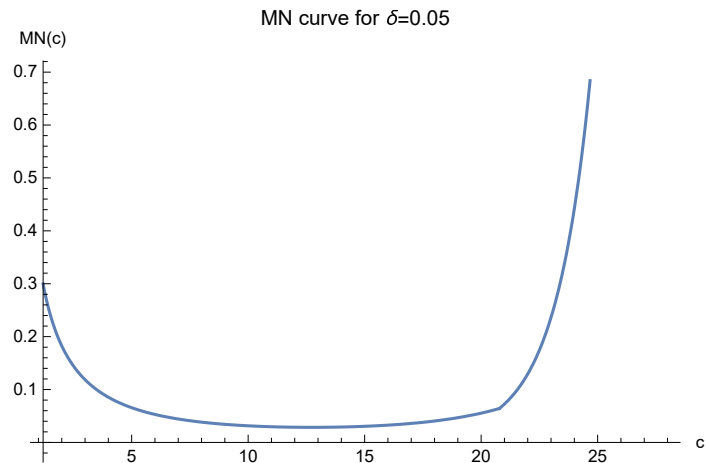


Figure 15: Here $d = 3$, $\beta = 1$, $b_0 = \sqrt{3}$ and $\sigma = 1E-1$.

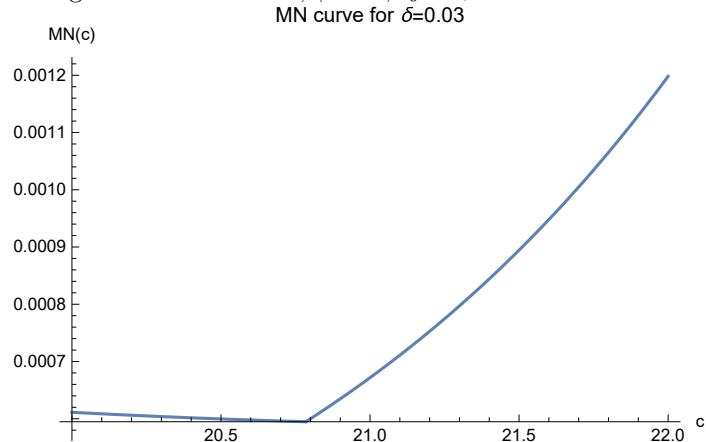


Figure 16: Here $d = 3$, $\beta = 1$, $b_0 = \sqrt{3}$ and $\sigma = 1E-1$.

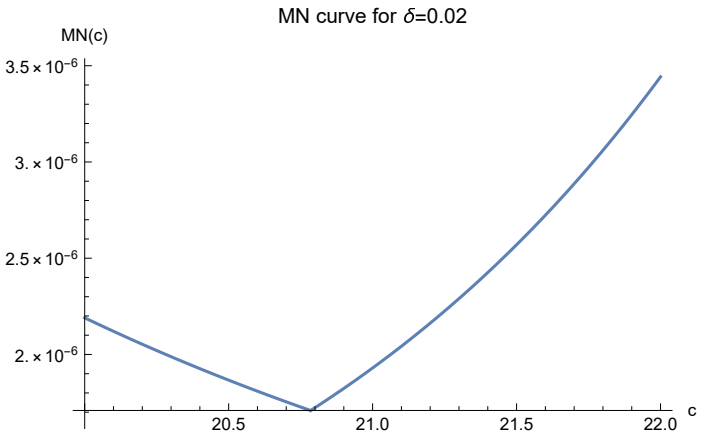


Figure 17: Here $d = 3$, $\beta = 1$, $b_0 = \sqrt{3}$ and $\sigma = 1\text{E}-1$.
MN curve for $\delta=0.01$

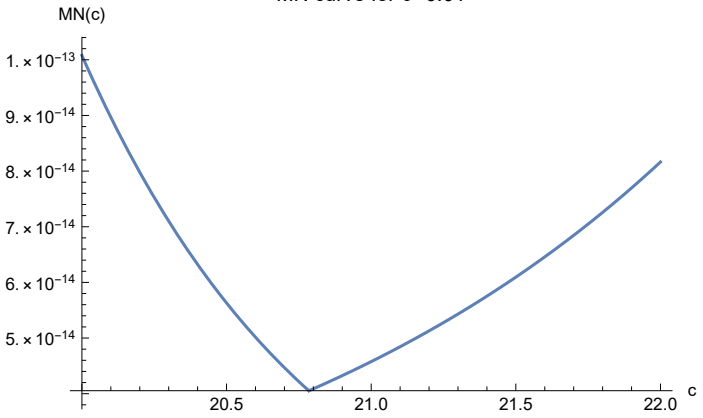


Figure 18: Here $d = 3$, $\beta = 1$, $b_0 = \sqrt{3}$ and $\sigma = 1\text{E}-1$.
MN curve for $\delta=0.001$

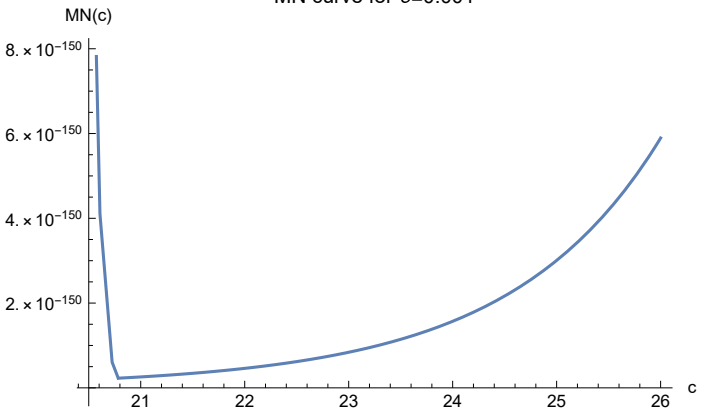


Figure 19: Here $d = 3$, $\beta = 1$, $b_0 = \sqrt{3}$ and $\sigma = 1\text{E}-1$.

Now we test $\sigma = 10^{-5}$. The MN curves are presented in Figs. 20-22. They show that one should choose $c = 129$.

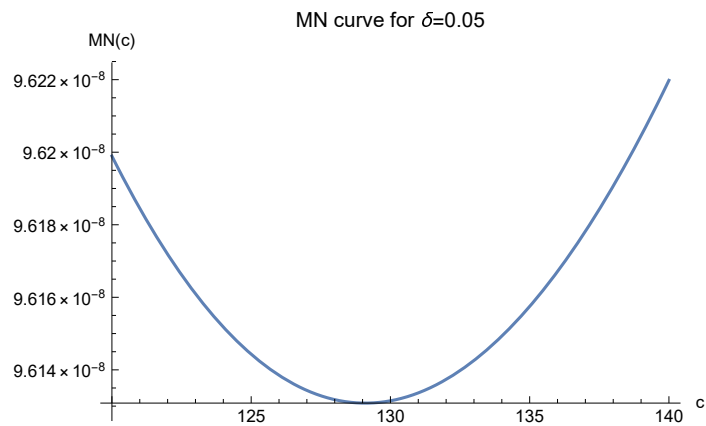


Figure 20: Here $d = 3$, $\beta = 1$, $b_0 = \sqrt{3}$ and $\sigma = 1E-5$.

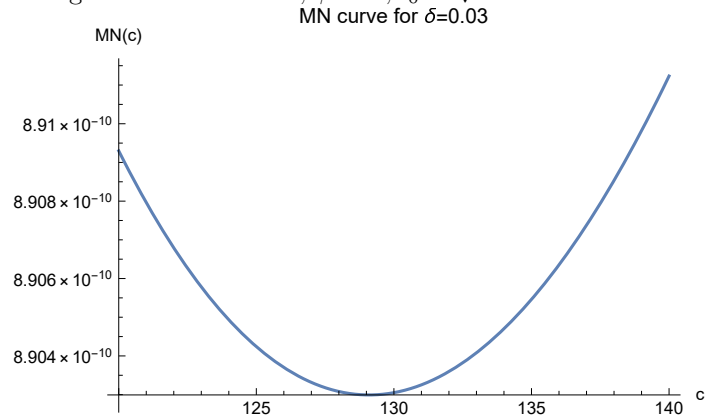


Figure 21: Here $d = 3$, $\beta = 1$, $b_0 = \sqrt{3}$ and $\sigma = 1E-5$.

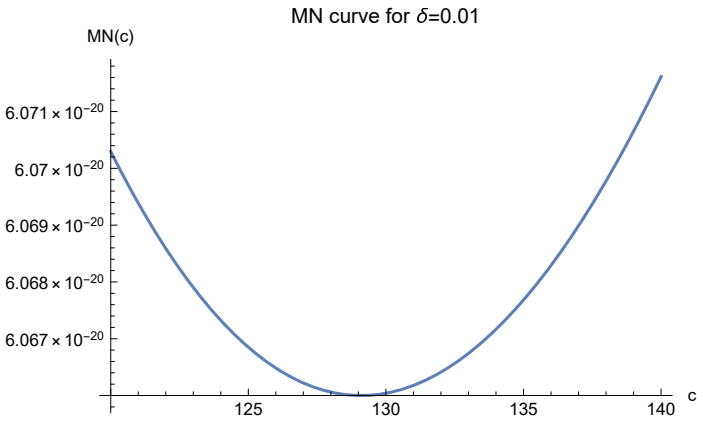


Figure 22: Here $d = 3$, $\beta = 1$, $b_0 = \sqrt{3}$ and $\sigma = 1E-5$.
MN curve for $\delta=0.05$

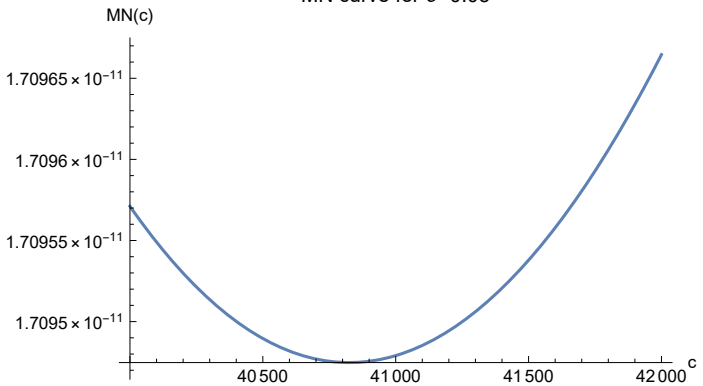


Figure 23: Here $d = 3$, $\beta = 1$, $b_0 = \sqrt{3}$ and $\sigma = 1E-10$.
MN curve for $\delta=0.03$

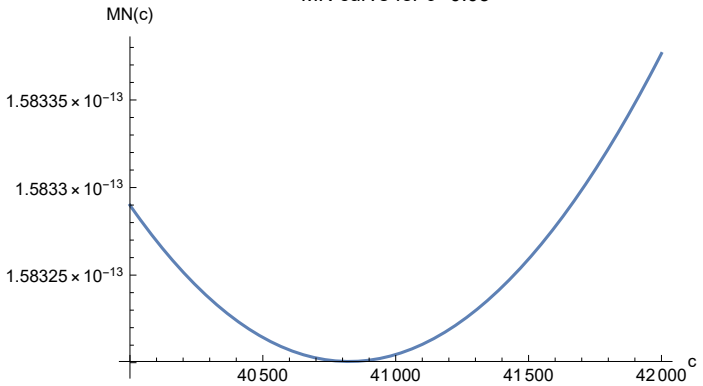


Figure 24: Here $d = 3$, $\beta = 1$, $b_0 = \sqrt{3}$ and $\sigma = 1E-10$.

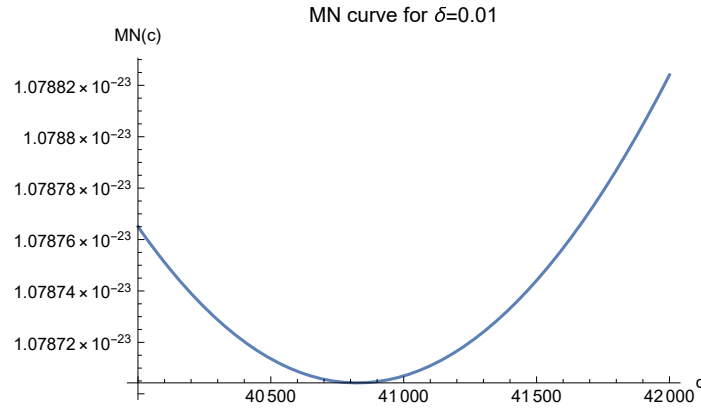


Figure 25: Here $d = 3$, $\beta = 1$, $b_0 = \sqrt{3}$ and $\sigma = 1\text{E}-10$.

The MN curves for $\sigma = 10^{-10}$ are as in Figs. 23-25. They indicate that one should let $c = 40800$.

The three different optimal values of c are all logically correct. However, when σ is very small, $\|f\|_{E_\sigma}$ in (4) will be extremely large, making (4) meaningless. Hence we should choose $c = 20.7846$ according to Figs. 15-19, where $\sigma = 10^{-1}$ is larger and the $\text{MN}(c)$ values are reasonably small.

For $c = 20.7846$, we compare different numbers of data points. The results are presented in Tab. 4.

Table 4: 3D experiment, $c = 20.7846$, $b_0 = \sqrt{3}$.

N_d	616	666	766	966	1366
N_{int}	50	100	200	400	800
N_{bdy}	566	566	566	566	566
N_t	1200	1200	1200	1200	1800
RMS	$6.7 \cdot 10^{-9}$	$1.5 \cdot 10^{-11}$	$3.1 \cdot 10^{-13}$	$1.8 \cdot 10^{-15}$	$7.0 \cdot 10^{-19}$
COND	$2.4 \cdot 10^{71}$	$2.6 \cdot 10^{71}$	$3.0 \cdot 10^{71}$	$3.7 \cdot 10^{71}$	$5.5 \cdot 10^{71}$

The computation is very efficient. Even though we adopted 200 effective digits for each step of the calculations, the most time-consuming work of solving the linear system took only two seconds for $N_d = 616$ and five minutes for $N_d = 1366$. In fact, keeping only 100 effective digits would have obtained the same RMS's and COND's. We stopped adding more data points at $N_d = 1366$ because the RMS is already good enough.

The comparison among different values of c is presented in Tab. 5. We fixed $N_d = 1366$, $N_{int} = 800$ and $N_{bdy} = 566$. In order to cope with the problem of ill-conditioning, enough effective digits were adopted for the calculations. For $c = 1$, we used 50 digits, and for $c = 220$, 140 digits were used. The most time-consuming work of solving the linear system took less than six minutes'

computer time always.

Note that Figs. 15-19 show that if δ is small enough, or equivalently the number of data points is large enough, the values of $MN(c)$ get large very slowly for $c > 20.7846$. Numerical investigations also show this. Figs. 20-25 even tend to move the optimal value of c to the right. All these are supported by the RMS's in Tab. 5. Although for $c > 20.7846$, the choice of c does not influence the approximation error much, one still has to choose $c = 20.7846$ because its corresponding condition number is smaller. In other words, the theoretically predicted optimal value of c coincides with the experimentally optimal one.

Table 5: 3D experiment, $b_0 = \sqrt{3}$, $N_d = 1366$, $N_t = 1800$.

c	1	15	20.7846	30	40
RMS	$8.7 \cdot 10^{-7}$	$2.6 \cdot 10^{-7}$	$7.0 \cdot 10^{-19}$	$7.6 \cdot 10^{-18}$	$1.0 \cdot 10^{-19}$
COND	$4.6 \cdot 10^{21}$	$6.4 \cdot 10^{65}$	$5.5 \cdot 10^{71}$	$2.5 \cdot 10^{78}$	$4.7 \cdot 10^{83}$
c	60	80	100	120	140
RMS	$3.4 \cdot 10^{-19}$	$8.7 \cdot 10^{-19}$	$6.8 \cdot 10^{-19}$	$1.4 \cdot 10^{-19}$	$3.8 \cdot 10^{-19}$
COND	$1.1 \cdot 10^{91}$	$1.9 \cdot 10^{96}$	$2.5 \cdot 10^{100}$	$5.1 \cdot 10^{103}$	$3.2 \cdot 10^{106}$
c	180	220			
RMS	$1.7 \cdot 10^{-18}$	$3.7 \cdot 10^{-19}$			
COND	$1.2 \cdot 10^{111}$	$5.6 \cdot 10^{114}$			

4 Discussion

In our experiments the exact solution function $u(x_1, \dots, x_d)$ is a natural function contained in the Sobolev space $W_2^\tau(R^d)$ for any $\tau \geq 0$. An approximate solution function $\hat{u}(x_1, \dots, x_d)$ could always be efficiently found with very small approximation error. The optimality of our choice of the shape parameter c has also been corroborated. It means that any differential equation may be effectively handled by our approach, as long as its solution belongs to Sobolev space. This is exciting. However, how to apply our approach to solving various important but hard differential equations is still very challenging, especially when the solution does not lie in Sobolev space.

Funding: This work was supported by the Taiwanese Ministry of Science and Technology [MOST project 109-2115-M-126-003-].

Conflicts of Interest: The author declares no conflict of interest.

References

- [1] Wendland, H. *Scattered Data Approximation*; Cambridge University Press; 2005.
- [2] Kansa, E.J. *Multiquadrics-a scattered data approximation scheme with applications to computational fluid dynamics I: Surface approximations and partial derivative estimates*. Comput. Math. Applic. 1990, V. 19 (8/9): 127-145.
- [3] Kansa, E.J. *Multiquadrics-a scattered data approximation scheme with applications to computational fluid dynamics II: Solutions to parabolic, hyperbolic, and elliptic partial differential equations*. Comput. Math. Applic. 1990, V. 19 (8/9): 147-161.
- [4] Fasshauer, G. *Meshfree approximation methods with MATLAB*; World Scientific Publishers, Singapore, 2007.
- [5] Yosida, K. *Functional Analysis*; Springer, 1980.
- [6] Narcowich, F.J.; Ward, J.D.; Wendland, H. *Sobolev error estimates and a Bernstein inequality for scattered data interpolation via radial basis functions*. Constr. Approx. 2006, 20:175-186.
- [7] Luh, L-T. *The mystery of the shape parameter IV*. Eng. Anal. Bound. Elem. 2014, 48, 24-31.
- [8] Luh, L-T. *The choice of the shape parameter-a friendly approach*. Eng. Anal. Bound. Elem. 2019, 98, 103-109.