Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2024-0028

Vol. **37**, No. 1, pp. 220-249
January 2025

# A Local Model Reduction Method Based on $k$-Nearest-Neighbors for Parametrized Nonlocal Problems

Caixia Nan[1], Qiuqi Li[2] and Huailing Song[2,3,*]

[1] *Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong.*
[2] *School of Mathematics, Hunan University, Changsha 410082, China.*
[3] *Greater Bay Area Institute for Innovation, Hunan University, Guangzhou, China.*

Received 5 February 2024; Accepted (in revised version) 26 August 2024

**Abstract.** In this paper, the model reduction method based on $k$-nearest-neighbors is provided for the parametrized nonlocal partial differential equations (PDEs). In comparison to standard local PDEs, the stiffness matrix of the corresponding nonlocal model loses sparsity due to the nonlocal interaction parameter $\delta$. Specially the nonlocal model contains uncertain parameters, enhancing the complexity of computation. In order to improve the computation efficiency, we combine the $k$-nearest-neighbors with the model reduction method to construct the efficient surrogate models of the parametrized nonlocal problems. This method is an offline-online mechanism. In the offline phase, we develop the full-order model by using the quadratic finite element method (FEM) to generate snapshots and employ the model reduction method to process the snapshots and extract their key characters. In the online phase, we utilize $k$-nearest-neighbors regression to construct the surrogate model. In the numerical experiments, we first verify the convergence rate when applying quadratic FEM to the nonlocal problems. Subsequently, for the linear and nonlinear nonlocal problems with random inputs, the numerical results illustrate the efficiency and accuracy of the surrogate models.

**AMS subject classifications**: 45A05, 45G10, 45P05, 65C30, 65R20, 65R99

**Key words**: Parametrized nonlocal PDEs, surrogate model, quadratic finite element method, proper orthogonal decomposition, dynamic mode decomposition, $k$-nearest-neighbors.

## 1 Introduction

Nonlocal phenomena, which have appeared observed in various fields such as physics, materials, biology, and social sciences, are ubiquitous in nature [11]. It is well-known

*Corresponding author. *Email addresses:* caixia.nan@polyu.edu.hk (C. Nan), qli28@hnu.edu.cn (Q. Li), shling@hnu.edu.cn (H. Song)

that the nonlocal parabolic model is a generalization of the classic local parabolic model, incorporating the horizon parameter $\delta$ to measure the range of nonlocal interactions. The specific nonlocal parabolic equation is given by:

$$\frac{\partial u}{\partial t} = \mathcal{L}_\delta u + f.$$

As $\delta \to 0$, the nonlocal model converges to the corresponding classic local model. Essentially, the discretization of the local Laplacian operator only requires computation of the discrete node values, while the discretization of the nonlocal diffusion operator necessitates information on feature interactions occurring between spatial points separated by a finite distance when $\delta > 0$ [14]. Consequently, the nonlocal model has significant advantages in exploring defects, for example, the nucleation and propagation of cracks [2].

In recent years, the exploration of theoretical and numerical analysis of nonlocal models has received much attention. Tian and Du [38] conducted a comprehensive study of piecewise constant finite element method (FEM) and piecewise linear FEM schemes for nonlocal diffusion and linear peridynamic equations, providing detailed discussions of singular kernels and analyzing fundamental theoretical and numerical scheme issues. Additionally, Li et al. [12, 15] investigated the exponential time differencing method for semilinear parabolic equations, such as the nonlocal Allen-Cahn equation, to explore its maximum bound principle. They also considered the nonlocal Cahn-Hilliard equation, rigorously establishing energy stability and convergence analysis [10, 25]. In [13], the authors adopted the implicit Runge-Kutta method and discontinuous Galerkin method for nonlocal diffusion problems, where the stability and error estimates of the fully discrete numerical schemes were presented. Furthermore, Nan and Song [31] applied integrating factor Runge-Kutta method and finite difference method (FDM) to solve the nonlocal Allen-Cahn equation, successfully demonstrating the maximum bound principle.

It is well known that nonlocal interactions lead to a denser discrete system compared to standard local models due to the horizon $\delta$, whether applying the FEM or the FDM. This increased density poses a significant challenge for the computation of the nonlocal model, and the inclusion of stochastic input further effects computational efficiency [24, 40]. Consequently, constructing a cheaper and simplified surrogate model for parametrized nonlocal model has become a significant subject of research. Surrogate model is an approximate model in the low dimensional subspace of the solution space. The success of these model reduction methods relies on the assumption that the solution manifold can be embedded in a low dimensional space [3]. But, the important class of problems arising from parametric dynamical systems typically induce a rough solution manifold with slowly decaying Kolmogorov n-widths. This suggests that traditional model-order reduction (MOR) methods [5, 22, 26, 28] are generally ineffective. In recent years, there has been a growing interest in the development of MOR techniques for parametric dynamical systems to overcome the limitations of linear global approximations. A large class of methods consider the dynamical low approximation, enabling both the deterministic and stochastic basis functions to evolve in time [30, 32, 45]. Other strategies,

based on deep learning algorithms, were proposed in [16, 29, 33, 46] to construct efficient surrogate models for time-dependent parametrized PDEs. In this context, we aim to employ classical methods to construct an efficient and reliable approximation of the input-output relationship for parametric dynamical systems to efficiently construct surrogate models. This includes the proper orthogonal decomposition (POD) method, the dynamic mode decomposition method (DMD), the $k$-nearest-neighbors (KNN) method, and the discrete empirical interpolation method (DEIM). These MOR methods are capable of handling problems of any spatial dimension.

Currently, there is limited discussion regarding the reduction for the nonlocal problems, and we present prior relevant research on this subject. The first research to explore the MOR for nonlocal model is presented in [42]. The authors considered MOR for nonlocal diffusion problems with uncertain parameters, where the piecewise FEM was used for spatial discretization and POD was used to generate the simplified surrogate models. Guan et al. [19] developed the reduced-basis approximations for nonlocal diffusion equations with affine random coefficients and illustrated the efficiency of the proposed model through numerical results. Lu and Nie [27] focused on the implementations of the meshfree-based ROMs for time-dependent nonlocal models with inhomogeneous volume constraints. They introduced a mixed reproducing approximation with nodal interpolation property to develop a meshfree collocation method, with results confirming the efficiency of their approach.

The KNN method is known for its ease of implementation and excellent performance in modeling physical systems with a large number of parameters [21]. Typically, it classifies training points based on the Euclidean metric from the observation $x$, defining it as the neighborhood of $x$. The size of the neighborhood is usually determined by the predefined $k$. Ghosh [18] introduced an adaptive choice nearest neighbor classification technique to select $k$, and the numerical examples illustrated the utility of the proposed method. Belkin et al. [1] provided theoretical foundations for interpolated classifiers by analyzing local interpolating schemes. Xing et al. [44] explored a class of interpolation weighting schemes in the nearest neighbors algorithm to enforce zero training error and discussed the universality of their results. Gao et al. [17] proposed a new surrogate model based on KNN regression and DMD to solve various nonlinear parameterized PDEs, demonstrating the strong predictive ability of KNN-DMD within the training time region.

POD is a widely-used dimensionality reduction method that extracts dominant features of a system to form a reduced basis, with Galerkin method projecting dynamic equations onto this reduced basis for reduced-order modeling of high-dimensional systems. It has been successfully applied to optimization involving PDEs or feedback control laws. Many studies have explored its applications, such as Rathinam and Petzold [34] investigated the basic properties of POD, and provided error analysis for reduced-order models, establishing the POD basis based on data ensembles with small perturbations. Sachs and Volkwein [35] applied POD to optimization problems with PDEs and studied posteriori error estimates. Schmidt et al. [36] applied POD for parabolic PDEs with

parameters and provided posteriori error estimate.

The DMD is a data-driven method used for analyzing and extracting dynamic information from time-series data, which decomposes a dynamical system into a set of spatiotemporal coherent structures represented as the DMD modes from a set of time series data, i.e., snapshots. This method can generate a low-dimensional eigenfunction-based approximation of the underlying Koopman operator. Initially developed in the field of fluid dynamics, then has been applied in various domain, including engineering, physics, and biology. Numerous studies have been devoted to this method. Drmac et al. [7] provided a robust numerical linear algebra framework to solve structured least squares problems based on DMD, with potential extensions to various computational tasks, such as multistatic antenna array processing. Colbrook [6] investigated a measure-preserving extended DMD (mpDMD), which is flexible and easy to use with any preexisting DMD-type method, and with different type of data. Their work also demonstrated the robustness of the mpDMD to the noise compared to other DMD-type methods. Li et al. [23] developed an adaptive method based on DMD, employing the Taylor expansion at each step, to construct an efficient and reliable surrogate model.

In this paper, we focus on the exploration of MOR for parametrized nonlocal PDEs. There are two main difficulties: firstly, due to the horizon $\delta$, the matrix generated by the nonlocal model is more complex than the corresponding local model. Secondly, the model inputs to the parametrized PDEs often contain uncertainties. Estimating the amount of uncertain parameters and quantifying their effects usually requires a large number of repetitive work [20]. To overcome these difficulties, we introduce the MOR methods to generate surrogate models of the parametrized nonlocal PDEs. The process of generating surrogate model requires an offline-online computational decomposition to improve efficiency. In the offline phase, we calculate the parametrized nonlocal PDEs based on the full-order models to produce the snapshots and obtain the data matrices. Singular value decomposition (SVD) is used to obtain the reduced operator matrices. In the online phase, we utilize the KNN method to select the $k$ nearest neighbor samples and calculate the weights of the samples to construct the surrogate models. The computation cost of the online phase is completely independent of the spatial discretization, which leads to high computational efficiency. In the numerical examples, we first demonstrate the convergence rate of the quadratic FEM for the nonlocal linear PDEs. Then, we consider the nonlocal PDEs with stochastic inputs, including linear and nonlinear cases. We compare the predicted solution produced by the proposed surrogate models and the reference solution based on the full-order equation. We calculate the relative error (RE) and mean error over the parameter sets at time $t$ (MTE), which indicate the potential of these MOR methods in predicting the parametrized nonlocal model solutions.

The rest of the paper is organized as follows. In Section 2, we introduce the parametrized nonlocal model. In Section 3, the piecewise quadrature FEM is proposed for spatial discretization and the weak formulation is derived. Section 4 and Section 5 correspond to the online phase and offline phase of the MOR methods for the surrogate models. Several numerical examples are presented in Section 6 to demonstrate the per-

formance and effectiveness of the introduced MOR methods. We make some conclusions in the last Section.

# 2 Preliminaries of the parametrized model and nonlocal operator

In this section, we begin by describing the parametrized PDEs, then we introduce the nonlocal operator to the model. Following that, we provide some notations for the model.

Let us consider the following general parametrized time-dependent PDEs with the initial condition,

$$\begin{cases} \dfrac{\partial u}{\partial t}+\mathcal{N}[u(x,t;\boldsymbol{\mu})]=f(x,t;\boldsymbol{\mu}), \ \ (x,t;\boldsymbol{\mu})\subset\Omega\times\mathcal{T}\times\mathcal{P}, \\ u(x,0;\boldsymbol{\mu})=u^0(x;\boldsymbol{\mu}), \end{cases} \tag{2.1}$$

where $\mathcal{P}\in\mathbb{R}^v$ and $v$ is the dimension of parameter vector, $\mathcal{T}$ represents the time domain ($\mathcal{T}$=(0, T], $T>0$), and $\Omega\in\mathbb{R}$ is the space domain. The notation $\mathcal{N}[u(x,t;\boldsymbol{\mu})]$ represents the linear or nonlinear operator associated with the spatial $x$, time $t$, and kernel (in nonlocal case), characterized by the parameter vector $\boldsymbol{\mu}$. The function $u(x,t;\boldsymbol{\mu})$ denotes the solution field that we aim to approximate, while $f(x,t;\boldsymbol{\mu})$ is the linear or nonlinear term.

## 2.1 Nonlocal operator

Considering the general parametrized PDEs (2.1), we present the nonlocal operator in one dimension,

$$\mathcal{N}[u(x,t;\boldsymbol{\mu})]=-\mathcal{L}_\delta u(x,t;\boldsymbol{\mu}), \tag{2.2}$$

where $\mathcal{L}_\delta$ is a nonlocal operator. The value of $\mathcal{L}_\delta u$ at point $x$ requires information about $u$ at points $y\neq x$, in addition to $x$. Define the nonlocal operator $\mathcal{L}_\delta$ by

$$\mathcal{L}_\delta u(x)=\int_{B_\delta(x)}(u(y)-u(x))\rho_\delta(x,y)dy, \quad \forall x\in I, \tag{2.3}$$

where $I=(\alpha,\beta)\subset\Omega\in\mathbb{R}$. $B_\delta(x)=\{y\in\Omega:|y-x|<\delta\}$ denotes a neighborhood centered at $x$ with radius $\delta$, and $\rho_\delta(x,y)$: $I\times I\to R$ is a nonnegative symmetric kernel function, i.e., $\rho_\delta(x,y)=\rho_\delta(y,x)\geq 0$. We consider kernel as the radial type, i.e., $\rho_\delta(x,y)=\rho_\delta(|x-y|)$. The kernel function $\rho_\delta(x,y)$ contains a parameter $z$, which belongs to parameter vector $\boldsymbol{\mu}$, and its specific expression in difference problem is given in Section 6. In addition, the second-order moment of $\rho_\delta(s)$ (s=$|x-y|$) is defined to be positive and finite,

$$0<C_\delta=\int_0^\delta s^2\rho_\delta(s)ds<\infty. \tag{2.4}$$

Eq. (2.4) allows for more general kernels, which may or may not be locally integrable.

Substituting the nonlocal operator (2.2) into the general parametrized Eq. (2.1), it can be transformed into:

$$\frac{\partial u}{\partial t}-\mathcal{L}_\delta u(x,t;\boldsymbol{\mu})=f(x,t;\boldsymbol{\mu}),\quad (x,t,\boldsymbol{\mu})\subset I\times\mathcal{T}\times\mathcal{P}. \tag{2.5}$$

Subsequently, we provide some notations for the spatial definition, boundary condition, and bilinear formulation of the nonlocal Eq. (2.5).

## 2.2 Notations

Let $I$ be a bounded line segment in the space $\Omega$, and let $N_x$ denote the number of grid points on $I$ except the endpoints, which are uniformly distributed. Here, $h=\frac{I}{N_x+1}$ represents the mesh size, and $I_n=[x_n,x_{n+1}]$, where $x_n=nh$ for $n=0,1,\cdots,N_x$. The value of $\mathcal{L}_\delta u$ at point $x$ interacts with the point $y$ in the interval $(x-\delta,x+\delta)$, which has a nonzeros length. Consequently, the domain $I$ is imposed the interval $\delta$ due to the nonlocality of interactions, where $\delta$ denotes the volume constraint, and we refer to the extended domain as $I'$. To clarify this, the domain is represented as follows,

$$I=(\alpha,\beta),\quad I'=(\alpha-\delta,\beta+\delta),\ \Gamma=\bar{I}'\setminus I=[\alpha-\delta,\alpha]\cup[\beta,\beta+\delta].$$

The boundary condition of Eq. (2.5) is set as,

$$f(x,t;\boldsymbol{\mu})=g(x,t;\boldsymbol{\mu}),\quad \forall(x,t;\boldsymbol{\mu})\subset\Gamma\times\mathcal{T}\times\mathcal{P},$$

where $g$ is the given function. In most cases, we define $g=u$, except in special circumstances.

Next, we give the definition for the energy space and the constrained energy space according to [8],

$$S(I')=\Big\{v\in L^2(I')|\int_I\int_{B_{\delta(x)}}(v(y)-v(x))^2\rho_\delta(|x-y|)dydx<\infty\Big\},$$

and $S_0(I')=\{w\in S(I')|w=0 \text{ on } \Gamma\}$. In order to compute Eq. (2.5) on $\Gamma$, we define the affine space [42],

$$S_g(I')=\{w\in S(I'):w=g\quad \text{on } \Gamma\}.$$

The bilinear form of the nonlocal linear operator on $L^2(I)\times L^2(I)$ is defined as

$$(\mathcal{L}_\delta u,v):=\frac{1}{2}\int_I\int_{-\delta}^{\delta}(u(y)-u(x))(v(y)-v(x))\rho_\delta(|x-y|)dydx,$$

which can be referred to [39]. The inner product on $L^2(I)$ can be represented as

$$(u,v)_I:=\int_I u(x)v(x)dx.$$

Multiplying the test function $v(x)$ by (2.5) and integrating over $I$ yields

$$\Big(\frac{\partial u(x,t;\boldsymbol{\mu})}{\partial t},v(x)\Big)_I-\big(\mathcal{L}_\delta u(x,t;\boldsymbol{\mu}),v(x)\big)_I=\big(f(x,t;\boldsymbol{\mu}),v(x)\big)_I. \tag{2.6}$$

It is important to note that the test function $v(x)$ must be zero on the boundary domain $\Gamma$.

Applying the nonlocal Green's first identity established in [9], we can derive the spatial discretization equation of (2.6), that is

$$\begin{aligned}&\int_I\frac{\partial u(x,t;\boldsymbol{\mu})}{\partial t}v(x)dx+\frac{1}{2}\int_I\int_{-\delta}^{\delta}(u(y,t;\boldsymbol{\mu})-u(x,t;\boldsymbol{\mu}))(v(y)-v(x))\rho_\delta(|x-y|)dydx\\&=\int_I f(x,t;\boldsymbol{\mu})v(x)dx,\end{aligned} \tag{2.7}$$

where $u\in S(I')$ and $v\in S_0(I')$. The detailed expression of the discrete scheme for Eq. (2.7) will be discussed in the following section.

## 3 Quadratic FEM for parameterized nonlocal model

In this section, we use the quadratic FEM for the parametrized nonlocal PDEs. The detailed explanation of quadratic FEM for the nonlocal model is provided, considering the complexity of the matrix caused by the horizon $\delta$.

We begin by uniformly dividing the domain $I'$ into intervals,

$$\alpha-\delta=x_{-p}<x_{-p+1}<\cdots<x_0<x_1<\cdots<x_{N_x}<x_{N_x+1}<\cdots<x_{N_x+1+p}=\beta+\delta, \tag{3.1}$$

where $x_0=\alpha$, $x_{N_x+1}=\beta$, and $p$ is a positive integer. The spaces are defined as $S^h(I')\subset S(I')$, $S_0^h(I')\subset S_0(I')$ (i.e., $S_0^h(w)=\{w^h\in S(I'):w^h=0 \text{ on } \Gamma\}$), and $S_g^h=\{w^h\in S^h(I'):w^h=g^h \text{ on } \Gamma\}$. $S^h(I')$ is the space of continuous piecewise quadrature polynomials.

For any $t\in\mathcal{T}$, we find the numerical solution $u_h(x,t)\in S^h(I')$ and present the weak spatial discretization formula for (2.7) as follows,

$$\begin{aligned}&\int_I\frac{\partial u_h(x,t;\mu_\xi)}{\partial t}v_h(x)dx+\int_I\int_0^{\delta}(u_h(y,t;\mu_\xi)-u_h(x,t;\mu_\xi))(v_h(y)-v_h(x))\rho_\delta(|x-y|)dydx\\&=\int_I f(x,t;\mu_\xi)v_h(x)dx,\qquad\forall\, v_h(x)\subset S_0^h(I').\end{aligned} \tag{3.2}$$

Here, $\mu_\xi$ represents the fixed parameter for the deterministic problem. In Eq. (3.2), we employ piecewise quadratic polynomials from $S^h(I')$ for the space discretization, with

the specific formats of the basis functions given by

$$\phi_{n,1}(x)=2\Big(\frac{x-x_n}{h}\Big)^2-3\frac{x-x_n}{h}+1,$$
$$\phi_{n,2}(x)=2\Big(\frac{x-x_n}{h}\Big)^2-\frac{x-x_n}{h},$$
$$\phi_{n,3}(x)=-4\Big(\frac{x-x_n}{h}\Big)^2+4\frac{x-x_n}{h}.$$

It is easy to see that the nodes with respect to quadratic FEM basis functions on the domain $I'$ are

$$(x'_{-2p},\, x'_{-2p+1},\, \cdots,\, x'_0,\, x'_1,\, \cdots,\, x'_{2N_x+1},\, x'_{2N_x+2},\, \cdots,\, x'_{2N_x+2+2p}), \tag{3.3}$$

with $x'_{-2p}=\alpha-\delta$ and $x'_{2N_x+2+2p}=\beta+\delta$. The relationships between $x_n$ and $x'_n$ are

$$x'_{2n}=x_n,\quad \text{for } n=-p,\cdots,N_x+1+p,$$

and

$$x'_{2n+1}=\frac{x_n+x_{n+1}}{2},\quad \text{for } n=-p,\cdots,N_x+p.$$

In other words, based on original partition results (3.1), we add nodes in the center of each $I'_n$ for $n=-p,\cdots,N_x+p$.

To provide a clearer understanding of the basis function, we present the relationship between $\{\psi_j\}_{j=-2p}^{2N_x+2+2p}$ and $\{\phi_j\}_{j=-p}^{N_x+1+p}$ on each node.

For $n=-p$, we have

$$\psi_{2n}=\phi_{n,1},\quad x_n<x<x_{n+1}.$$

For $n=-p,\cdots,N_x+p-1$, there are

$$\psi_{2n+1}=\phi_{n,3},\quad x_n<x<x_{n+1},$$

and

$$\psi_{2(n+1)}=\begin{cases}\phi_{n,2}, & x_n<x<x_{n+1},\\ \phi_{(n+1),1}, & x_{n+1}<x<x_{n+2}.\end{cases}$$

For $n=N_x+p$, we obtain

$$\psi_{2n+1}=\phi_{n,3},\quad x_n<x<x_{n+1}\quad\text{and}\quad \psi_{2n+2}=\phi_{n,2},\quad x_n<x<x_{n+1}.$$

Clearly, the basis function $\{\phi_j\}_{j=-p}^{N_x+1+p}$ is defined at node $x_j$, while the basis function $\{\psi_j\}_{j=-2p}^{2N_x+2+2p}$ is defined at node $x'_j$.

Then, we partition the given time $T\in\mathcal{T}$ as

$$0=t^0<t^1<t^2<\cdots<t^{N_t}=T,$$

with $\Delta t=\frac{T}{N_t}$. The finite element approximation $u_h(x,t;\mu_\xi)$ approximately equals to $u(x,t;\mu_\xi)$, which can be expressed in the form of

$$u_h(x,t^n;\mu_\xi)=\sum_{j=-2p}^{2N_x+2+2p}c_j^n\psi_j(x).$$

By applying the backward Euler method to the time discretization, we obtain the fully discrete formulation of (2.5),

$$\begin{aligned}&\sum_{j=-2p}^{2N_x+2+2p}c_j^{n+1}\int_I\psi_j(x)\psi_i(x)dx\\&\quad+\Delta t\sum_{j=-2p}^{2N_x+2+2p}c_j^{n+1}\int_I\int_0^\delta(\psi_j(y)-\psi_j(x))(\psi_i(y)-\psi_i(x))\rho_\delta(|x-y|)dydx\\=&\Delta t\int_I f(x,t^{n+1};\mu_\xi)\psi_i(x)dx+\sum_{j=-2p}^{2N_x+2+2p}c_j^n\int_I\psi_j(x)\psi_i(x)dx,\\&\qquad\text{for } i=1,2,\cdots,2N_x+1.\end{aligned}\tag{3.4}$$

Based on the initial condition and the volume constraints, more specific fully discrete equation can be provided. Before presenting it, we give the following three notations to clarify Eq. (3.4).

First, since the test function $v_h$ is zero on $\Gamma$, we have

$$\psi_i(x)=0,\quad\text{on } i=\{-2p,\cdots,0\}\cup\{2N_x+2,\cdots,2N_x+2+2p\}.$$

Second, for the node on $\Gamma$, we represent it as

$$c_j^n=g^h(x_j',t^n;\mu_\xi),\quad\text{for } n=0,\cdots,N_t,\text{ and } j=\{-2p,\cdots,0\}\cup\{2N_x+2,\cdots,2N_x+2+2p\}.$$

Third, for the value of $c_j$ at $t^0$, there exists

$$c_j^0=u^0(x_j';\mu_\xi),\quad\text{for } j=1,\cdots,2N_x+1.$$

Taking into account the above notations, the unknown coefficients $c_j$ ($j=1,\cdots,2N_x+1$) are

determined by solving the following linear system,

$$\begin{aligned}
&\sum_{j=1}^{2N_x+1} c_j^{n+1}\int_I \psi_j(x)\psi_i(x)dx \\
&\quad+\Delta t\sum_{j=1}^{2N_x+1} c_j^{n+1}\int_I\int_0^\delta (\psi_j(y)-\psi_j(x))(\psi_i(y)-\psi_i(x))\rho_\delta(|x-y|)dydx \\
=&\Delta t\int_I f(x,t^{n+1};\mu_{\tilde{\xi}})\psi_i(x)dx+\sum_{j=0}^{2N_x+2} c_j^n\int_I \psi_j(x)\psi_i(x)dx \\
&-\Delta t\sum_{j\in\mathbb{N}} g^h(x_j',t^{n+1})\int_I\int_0^\delta (\psi_j(y)-\psi_j(x))(\psi_i(y)-\psi_i(x))\rho_\delta(|x-y|)dydx \\
&-g^h(x_0',t^{n+1})\int_I \psi_0(x)\psi_i(x)dx-g^h(x_{2N_x+2}',t^{n+1})\int_I \psi_{2N_x+2}(x)\psi_i(x)dx, \qquad (3.5)
\end{aligned}$$

for $i=1,\cdots,2N_x+1$ and $\mathbb{N}=\{-2p,\cdots,0\}\cup\{2N_x+2,\cdots,2N_x+2+2p\}$. In the case of local PDEs, the stiffness matrix is sparse when using the quadrature FEM to space discretization. However, for the nonlocal model, the stiffness matrix loses its sparsity. More specifically, compared to the local PDEs setting, this results in a denser stiffness matrix and increased assembly for obtaining the results of the nonlocal parametrized PDEs.

Given the full-order discrete scheme (3.5) of the nonlocal model, we present the equivalent matrix formulation as

$$\begin{aligned}
\boldsymbol{M}_1\boldsymbol{C}_1^{n+1}+\Delta t\boldsymbol{B}_1\boldsymbol{C}_1^{n+1}=&\Delta t\boldsymbol{F}^{n+1}+\boldsymbol{M}_2\boldsymbol{C}_2^n-\Delta t\boldsymbol{B}_2(g_1^h)^{n+1} \\
&-\Delta t\boldsymbol{B}_3(g_2^h)^{n+1}-\boldsymbol{M}_3 g^h(x_0',t^{n+1})-\boldsymbol{M}_4 g^h(x_{2N_x+2}',t^{n+1}), \qquad (3.6)
\end{aligned}$$

where for $i=1,\cdots,2N_x+1$, there are

$$\begin{aligned}
&\boldsymbol{M}_1(i,j)=\int_I \psi_j(x)\psi_i(x)dx, \quad j=1,\cdots,2N_x+1, \\
&\boldsymbol{M}_2(i,j)=\int_I \psi_j(x)\psi_i(x)dx, \quad j=0,\cdots,2N_x+2, \\
&\boldsymbol{M}_3(i)=\int_I \psi_0(x)\psi_i(x)dx, \quad \boldsymbol{M}_4(i)=\int_I \psi_{2N_x+2}(x)\psi_i(x)dx,
\end{aligned}$$

and

$$\begin{aligned}
&\boldsymbol{B}_1(i,j)=\int_I\int_0^\delta (\psi_j(y)-\psi_j(x))(\psi_i(y)-\psi_i(x))\rho_\delta(|x-y|)dydx, \quad j=1,\cdots,2N_x+1, \\
&\boldsymbol{B}_2(i,j)=\int_I\int_0^\delta (\psi_j(y)-\psi_j(x))(\psi_i(y)-\psi_i(x))\rho_\delta(|x-y|)dydx, \quad j=-2p,\cdots,0, \\
&\boldsymbol{B}_3(i,j)=\int_I\int_0^\delta (\psi_j(y)-\psi_j(x))(\psi_i(y)-\psi_i(x))\rho_\delta(|x-y|)dydx, \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad j=2N_x+2,\cdots,2N_x+2+2p,
\end{aligned}$$

and

$$g_1^h = g^h(x_j', t), \quad j = -2p, \cdots, 0, \quad g_2^h = g^h(x_j', t), \quad j = 2N_x + 2, \cdots, 2N_x + 2 + 2p,$$

and

$$\boldsymbol{F}(i) = \int_I f(x,t;\mu_{\tilde{\xi}})\psi_i(x)dx, \quad \boldsymbol{C}_1 = [c_1, c_2, \cdots, c_{2N_x+1}]', \quad \boldsymbol{C}_2 = [c_0, c_1, \cdots, c_{2N_x+2}]'.$$

The detailed procedure of employing the quadratic FEM for parametrized nonlocal problems is provided. Following that, we present the surrogate model for the full-order model.

## 4 Surrogate model based on KNN regression

As previously mentioned, it is typically computationally expensive when dealing with the multi-parameter inputs, multi-physics and time-dependent problems of standard PDEs, whether using the FDM or the FEM. Not to mention nonlocal problems, which involve the nonlocal parameter $\delta$. To overcome these challenges, researchers have focused on constructing the inexpensive and simplified surrogate model to reduce computational complexity and improves efficiency. In this section, we introduce the KNN method, which is employed in the online phase for generating surrogate model.

Assume that the solution $\boldsymbol{u}$ ($\boldsymbol{u}$ is a matrix for space $x$ and time $t$) is a random field, possessing enough regularity [47]. Using a group of random samples $\{\mu_i\}_{i=1}^{\check{M}}$, i.e., training set $\Xi = \{\mu_i\}_{i=1}^{\check{M}}$, we construct the surrogate model using the KNN regression based on MOR methods, where $\check{M}$ is the number of training parameter sets. The offline-online computational decomposition is obtained to enhance efficiency. The main idea of the offline is to solve the full-order PDEs and subsequently utilize MOR methods to process the snapshots and extract their main features. While the online phase is to apply the KNN method to determine the surrogate model by searching the $k$ nearest samples in training parameter set based on the test parameters features. Here, we provide a detailed introduction to the KNN method for the sake of completeness.

The KNN method is efficient for stochastic input problems due to its simplicity and popularity. The principle is to select the predefined $k$ nearest neighbors (from the training cases) of the sample points based on the defined metric measure. Then, the weights of the $k$ selected points are calculated using the weighted average. The predicted solution is obtained as a linear combination of the weights and the corresponding vectors. In other words, the predicted solution $\boldsymbol{u}(\mu_{\tilde{\xi}})$ with the new parameter $\mu_{\tilde{\xi}}$ can be expressed as

$$\boldsymbol{u}(\mu_{\tilde{\xi}}) = \frac{\sum_{j=1}^{k} \omega_j \boldsymbol{u}(\mu_j)}{\sum_{j=1}^{k} \omega_j}, \tag{4.1}$$

where $\omega_j$ denotes the weight assigned to the corresponding $\boldsymbol{u}(\mu_j)$, defined by

$$\omega_j = \frac{1}{|\mu_\xi - \mu_j|}, \tag{4.2}$$

and $|\cdot|$ represents the Euclidean distance. The selection of an value for $k$ is crucial in the KNN method. Generally, a small $k$ implies that the training domain is small around the object parameter used for prediction, which increases the sensitivity of prediction with respect to the selected parameter (overfitting). Conversely, a large $k$ means that the training instances are located in a big domain around the parameter values, which may overlook small but critical relationships between the parameter and the prediction, resulting in reduced similarity for the training instances and biased prediction (underfitting) [18].

# 5 Model reduction methods for the offline stage

In this section, we provide some MOR methods employed in the offline phase to construct the surrogate model based on the training set $\Xi$. Specifically, the POD method, a well-known approach for reducing the system's complexity, has been extensively utilized for dimensional reduction. The DMD method, a data-driven technique for model reduction, is also becoming increasingly popular. In the following, we provide a brief review of the POD method and DMD method separately.

## 5.1 POD method

Let $V$ be a Hilbert space equipped with the inner product $(\cdot,\cdot)_V$ and the norm $\|\cdot\|_V$. The snapshots are defined as the solutions of (3.5) at a selected time $t$, denoted as $\boldsymbol{u}^1,\boldsymbol{u}^2,\cdots,\boldsymbol{u}^r \in V$ ($\boldsymbol{u}^r = \boldsymbol{u}(t^r)$ is a vector for $r$th time). The purpose of the POD method is to find a set of orthogonal basis functions $\{\gamma_i\}_{i=1}^{\hat{r}}$ that optimally approximate the snapshots $\{\boldsymbol{u}^i\}_{i=1}^{r}$. The basis set $\{\gamma_i\}_{i=1}^{\hat{r}}$ solves the following optimal problem,

$$\begin{aligned} &\min_{\{\gamma_j\}_{j=1}^{\hat{r}}} \sum_{l=1}^{r} \|\boldsymbol{u}^l - \Sigma_{j=1}^{\hat{r}} (\boldsymbol{u}^l,\gamma_j)_V \gamma_j\|_V^2, \\ &\text{s.t. } (\gamma_i,\gamma_j)_V = \tilde{\delta}_{ij}, \quad i,j = 1,\cdots,\hat{r}, \end{aligned} \tag{5.1}$$

where $\tilde{\delta}$ is a kronecker function. To get the solution of (5.1), one use the snapshots method [37] to calculate the correlation matrix $\boldsymbol{K}$ corresponding to the snapshots,

$$\boldsymbol{K}_{i,j} = (\boldsymbol{u}^i,\boldsymbol{u}^j)_V, \quad i,j = 1,\cdots,r,$$

where $\boldsymbol{K}$ is positive semi-definite. We define $\Lambda = diag(\sigma_1,\cdots,\sigma_r)$, where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ are the eigenvalues of matrix $\boldsymbol{K}$ and $\boldsymbol{\varphi}_1,\boldsymbol{\varphi}_2,\cdots,\boldsymbol{\varphi}_r$ are the associate eigenvectors. The POD basis can be derived by

$$\gamma_j = \frac{1}{\sqrt{\sigma_j}} \Sigma_{i=1}^{r} (\boldsymbol{\varphi}_j)_i \boldsymbol{u}^i,$$

where $(\boldsymbol{\varphi}_j)_i$ is the $i$th element of $\boldsymbol{\varphi}_j$ with $j=1,\cdots,\hat{r}$. Naturally, the number of the POD basis functions (i.e., $\hat{r}$) is fewer than the rank number of $\boldsymbol{K}$ (i.e., $r$), which can be chosen based on the decay speed of the eigenvalues of the correlation matrix [43]. The related reduced basis model is located in Appendix A.

### 5.2 DMD method

In order to further improve the computational efficiency and achieve the prediction at any time, we introduce the DMD method, which is a more recent numerical technique and a powerful and versatile technique for analyzing time series data. To begin with it, we obtain the field data, i.e., temporal snapshots, from the experiments or numerical simulations, and use the sampling time step to separate the original temporal snapshots into two subsequent snapshot sequences,

$$\boldsymbol{W}_1=[\boldsymbol{u}^0,\,\boldsymbol{u}^1,\,\cdots,\boldsymbol{u}^{q-1}]\quad\text{and}\quad\boldsymbol{W}_2=[\boldsymbol{u}^1,\,\boldsymbol{u}^2,\,\cdots,\boldsymbol{u}^q],\tag{5.2}$$

where $\boldsymbol{u}^i\in\mathbb{R}^l$ denotes the $i$th snapshot, $q$ is the number of snapshots and $l$ represents the number of measurement nodes per snapshots. $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ are assumed to be connected using the following linear dynamics,

$$\boldsymbol{W}_2=\boldsymbol{A}\boldsymbol{W}_1,$$

where $\boldsymbol{A}$ is a matrix, and a best-fit linear operator relating $\boldsymbol{W}_2$ to $\boldsymbol{W}_1$ with $\boldsymbol{A}=\boldsymbol{W}_2(\boldsymbol{W}_1)^\dagger$. "†" represents the Moore-Penrose pseudoinverse. When applying the DMD to solve a nonlinear PDEs, $\boldsymbol{A}$ is a global linear approximation to the nonlinear dynamics.

Calculating the SVD of $\boldsymbol{W}_1$, we obtain

$$\boldsymbol{W}_1\approx\boldsymbol{X}\Lambda\boldsymbol{Y}^*,$$

where $\boldsymbol{X}\in\mathbb{C}^{l\times\zeta}$, $\Lambda\in\mathbb{C}^{\zeta\times\zeta}$, $\boldsymbol{Y}\in\mathbb{C}^{q\times\zeta}$, "$*$" denotes the transpose of a matrix or vector, and $\zeta$ denotes the rank of $\Lambda$.

Then, we compute the $\zeta\times\zeta$ projection of $\boldsymbol{A}$ on $\boldsymbol{X}$, yields

$$\tilde{\boldsymbol{A}}=\boldsymbol{X}^*\boldsymbol{A}\boldsymbol{X}=\boldsymbol{X}^*\boldsymbol{W}_2\boldsymbol{Y}\Lambda^{-1},$$

where $\tilde{\boldsymbol{A}}$ is a low rank matrix approximating to $\boldsymbol{A}$. Compute the eigendecomposition of $\tilde{\boldsymbol{A}}$, we have

$$\tilde{\boldsymbol{A}}\boldsymbol{P}=\boldsymbol{P}\boldsymbol{\Pi},$$

where $\boldsymbol{\Pi}=diag(a_1,\cdots,a_\zeta)$ is eigenvalue matrix, and the columns of $\boldsymbol{P}$ are the corresponding eigenvectors.

Following that, the DMD mode corresponding to a particular DMD eigenvalue in $\boldsymbol{\Pi}$ is given, represented as

$$\Psi=\boldsymbol{W}_2\boldsymbol{Y}\Lambda^{-1}\boldsymbol{P}.$$

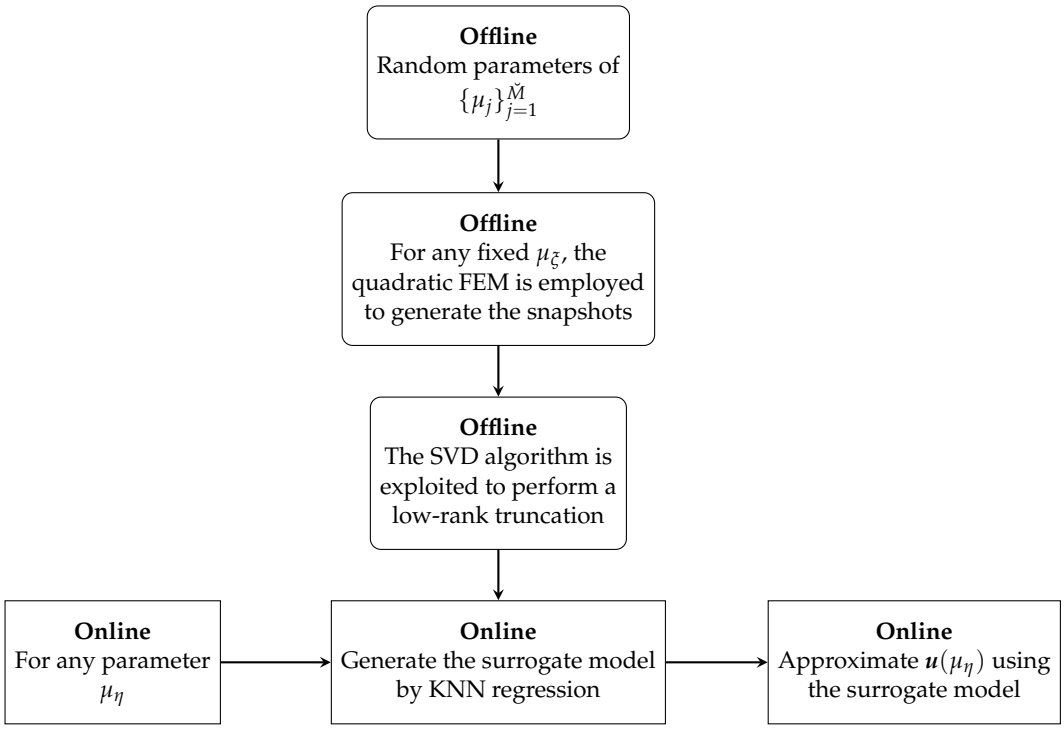



Figure 1: Flow chart of the online-offline for ROMs.

Finally, we employ the above approximated eigenvalues and corresponding eigenvectors of $\boldsymbol{A}$ to get the solutions for all time in the future,

$$\boldsymbol{u}^{n+1} \approx \boldsymbol{u}_{DMD}^{n+1} = \Psi \boldsymbol{\Pi}^{n+1} \boldsymbol{b}, \quad n \geq q, \quad \text{and} \quad \boldsymbol{b} = \Psi^{\dagger} \boldsymbol{u}^{0}.$$

More details about DMD please see [7, 17, 41].

More importantly, it is necessary to note that the accuracy of the proposed surrogate models may be influenced by the rank of reduced SVD in the offline phase. In all numerical examples, we will explicitly specify the values of these numbers.

Fig. 1 illustrates the flow chart outlining the offline and online phases of surrogate models. In the offline phase, a set of training parameters $\{\mu_j\}_{j=1}^{\check{M}}$ is initially provided, distributed randomly over the parameter range, and full-order models are calculated to generate snapshots. Subsequently, employing the POD or DMD method, the main feature are extracted using the SVD algorithm. In the online phase, for any given parameter $\mu_\eta$ within the parameter range, we use the KNN regression to construct an appropriate surrogate model. This surrogate model is then used to predict a solution that approximates $\boldsymbol{u}(\mu_\eta)$.

# 6 Numerical experiments

In this section, we first verify the spatial accuracy of the quadratic FEM for the nonlocal PDEs with a smooth solution, as shown in Section 6.1. Two kernel functions are introduced. The first kernel is a constant [38], denoted as

$$\rho_\delta(s) = 3\delta^{-3}. \tag{6.1}$$

The second kernel function is defined by

$$\rho_\delta(s) = \frac{2-2z}{\delta^{2-2z}} s^{-1-2z}, \quad z \in [-0.5,1), \tag{6.2}$$

which is parametrized by $z$, as presented in [39]. These two kernels satisfy Eq. (2.4) with $C_\delta = 1$ in one dimension.

Next, we present a few numerical examples to verify the accuracy and efficiency of the surrogate models for the parametrized nonlocal PDE with the parametrized kernel function (6.2). The domain $I = [0,1]$ and the horizon $\delta = 4h$ are used for the parametrized nonlocal PDEs. The result of full-order equation is taken as the reference solution. In Section 6.2, we apply the proposed POD-KNN method to solve the nonlocal linear equation, considering two parameters located in the kernel function and the coefficient of the nonlocal operator. Within this subsection, we present the Gaussian process regression (GPR) approach with POD method for nonlocal problems to make the comparison. In Section 6.3, we consider the nonlocal PDEs with parametrized boundaries to demonstrate efficiency and accuracy of the DMD-KNN and POD-KNN methods. Finally, in Section 6.4, we adopt the nonlocal Allen-Cahn equation as an example of the nonlinear PDE, where the parameters lie in the coefficients of the nonlocal operator term and the nonlinear term, and we adopt POD-DEIM-KNN method to generate the surrogate model. More importantly, we calculate the RE values between the reference solution and the predicted solution by fixing parameters, as well as MTE values, which are important metrics for evaluating surrogate models.

For the given parameter $\mu_{\tilde{\xi}}$, RE at time $t_j$ is defined as follows,

$$\varepsilon_{RE}(x,t^j;\mu_{\tilde{\xi}}) = \frac{\|u_{pred}(x,t^j;\mu_{\tilde{\xi}}) - u_{ref}(x,t^j;\mu_{\tilde{\xi}})\|_{L^2}}{\|u_{ref}(x,t^j;\mu_{\tilde{\xi}})\|_{L^2}}, \quad t^j \in \mathcal{T},$$

where $u_{pred}(x,t^j;\mu_{\tilde{\xi}})$ and $u_{ref}(x,t^j;\mu_{\tilde{\xi}})$ represent the predicted solution and reference solution. The MTE at time $t^j$ is defined by

$$\varepsilon_{MTE}(t^j) = \frac{1}{N_\mu} \sum_{i=1}^{N_\mu} \varepsilon_{RE}(x,t^j;\mu_i),$$

with $N_\mu$ representing the size of the testing parameter sets.

### 6.1 Convergence test

Let us begin with the following simple nonlocal linear time-dependent PDEs and its specific expression is

$$u_t - \mathcal{L}_\delta u = f, \tag{6.3}$$

with the smooth solution

$$u(x,t) = x^2(1-x^2)\sin(t).$$

Setting the final time as $T=1$, we use the quadratic FEM for the spatial discretization. To compute the errors, we fix the spacial mesh size as $h = 1/2^4, 1/2^5, 1/2^6, 1/2^7$ and the time step size as $\Delta t = h^3$. The $L^\infty$-norm and $L^2$-norm of the errors, along with their corresponding convergence rates, are shown in Table 1 and Table 2, where the expected spatial convergence rates are observed. It is clear that the numerical solutions are almost independent of $\delta$ and $z$.

For the constant kernel (6.1), the source term of Eq. (6.3) is represented by

$$f_\delta(x,t) = x^2(1-x^2)\cos(t) + \left(12x^2 - 2 + \frac{6}{5}\delta^2\right)\sin(t).$$

By setting the horizon parameter as $\delta = h, 4h, 6h$, we test the convergence of Eq. (6.3), and the corresponding results are presented in Table 1.

Table 1: Convergence test of the full-order nonlocal problem (6.3) with $\delta = h, 4h, 6h$.

| | $h$ | $\|u-u_h\|_{L^\infty}$ | Rate | $\|u-u_h\|_{L^2}$ | Rate |
|---|---|---|---|---|---|
| $\delta = h$ | $1/2^4$ | $7.2344e-06$ | – | $7.1715e-06$ | – |
| | $1/2^5$ | $5.0879e-07$ | 3.8297 | $5.7173e-07$ | 3.6489 |
| | $1/2^6$ | $4.1539e-08$ | 3.6145 | $5.1244e-08$ | 3.4799 |
| | $1/2^7$ | $4.2111e-09$ | 3.3022 | $5.1448e-09$ | 3.1362 |
| $\delta = 4h$ | $1/2^4$ | $7.4167e-05$ | – | $6.9081e-05$ | – |
| | $1/2^5$ | $4.7971e-06$ | 3.9505 | $4.4353e-06$ | 3.9612 |
| | $1/2^6$ | $3.0824e-07$ | 3.9600 | $2.9250e-07$ | 3.9225 |
| | $1/2^7$ | $2.0067e-08$ | 3.9412 | $2.0210e-08$ | 3.8553 |
| $\delta = 6h$ | $1/2^4$ | $1.6406e-04$ | – | $1.5281e-04$ | – |
| | $1/2^5$ | $1.0573e-05$ | 3.9558 | $9.6711e-06$ | 3.9819 |
| | $1/2^6$ | $6.7370e-07$ | 3.9721 | $6.1945e-07$ | 3.9646 |
| | $1/2^7$ | $4.2965e-08$ | 3.9709 | $4.0661e-08$ | 3.9293 |

Table 2: Convergence test of the full-order nonlocal parameterized problem (6.3) with $z=-\frac{1}{2},-\frac{1}{4},-\frac{1}{8}$.

| | $h$ | $\|u-u_h\|_{L^\infty}$ | Rate | $\|u-u_h\|_{L^2}$ | Rate |
|---|---|---|---|---|---|
| $z=-\frac{1}{2}$ | $1/2^4$ | 8.8749$e$−06 | – | 8.4355$e$−06 | – |
| | $1/2^5$ | 5.6991$e$−07 | 3.9609 | 5.2506$e$−07 | 4.0037 |
| | $1/2^6$ | 3.6362$e$−08 | 3.9702 | 3.3854$e$−08 | 3.9551 |
| | $1/2^7$ | 2.8174$e$−09 | 3.6900 | 3.2712$e$−09 | 3.3714 |
| $z=-\frac{1}{4}$ | $1/2^4$ | 6.1480$e$−06 | – | 5.8210$e$−06 | – |
| | $1/2^5$ | 3.9366$e$−07 | 3.9651 | 3.6261$e$−07 | 4.0048 |
| | $1/2^6$ | 2.5189$e$−08 | 3.9661 | 2.3724$e$−08 | 3.9340 |
| | $1/2^7$ | 2.1896$e$−09 | 3.5241 | 2.6387$e$−09 | 3.1684 |
| $z=-\frac{1}{8}$ | $1/2^4$ | 4.8079$e$−06 | – | 4.5422$e$−06 | – |
| | $1/2^5$ | 3.0729$e$−07 | 3.9677 | 2.8313$e$−07 | 4.0039 |
| | $1/2^6$ | 1.9721$e$−08 | 3.9618 | 1.8769$e$−08 | 3.9150 |
| | $1/2^7$ | 1.9031$e$−09 | 3.3733 | 2.3305$e$−09 | 3.0096 |

For the parametrized kernel function (6.2), the source term of Eq. (6.3) is

$$f_\delta(x,t)=x^2(1-x^2)\cos(t)+\left(12x^2-2+\frac{2z-2}{z-2}\delta^2\right)\sin(t).$$

By fixing horizon $\delta=4h$, and taking the parameter in the kernel function as $z=-\frac{1}{2},-\frac{1}{4},-\frac{1}{8}$, respectively, we calculate the errors and convergence rates of Eq. (6.3). The results are presented in Table 2.

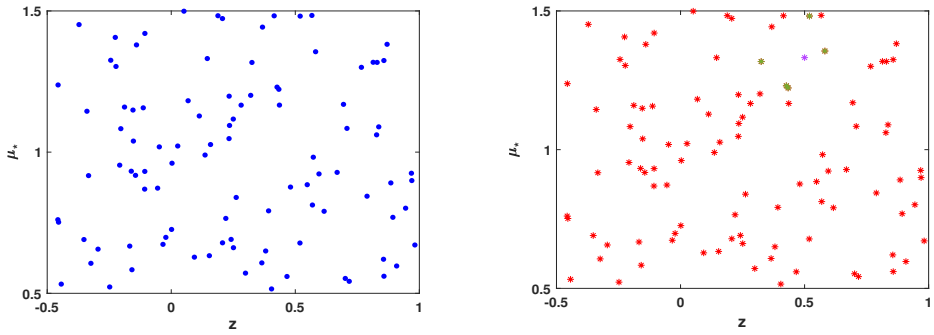
From Table 1 and Table 2, we observe that the errors of $L^\infty$-norm and $L^2$-norm decay and the convergence rate approximates to third-order as the spatial grid grows exponentially. Generally, the expected order is observed using the quadratic FEM for the nonlocal model.

## 6.2 Linear parametrized nonlocal models

In this part, we focus on the parametrized nonlocal linear PDEs with the parametrized kernel function (6.2),

$$u_t-\mu_*\mathcal{L}_\delta u=f. \tag{6.4}$$

There have two parameters in (6.4), one exists in the kernel function (6.2) with $z$, and another is $\mu_*$, i.e., $\boldsymbol{\mu}=(z,\mu_*)$. The value of the parameter is produced by "rand" and the range of parameter is set as $\mathcal{P}(z,\mu_*)=[-0.5,1)\times(0.5,1.5)$. Fixing $\delta=4h$, we calculate the problem with $h=1/2^7$, $\Delta t$ = 1E-4 and $T=1$.




(a) Random parameters (b) Samples of random parameters

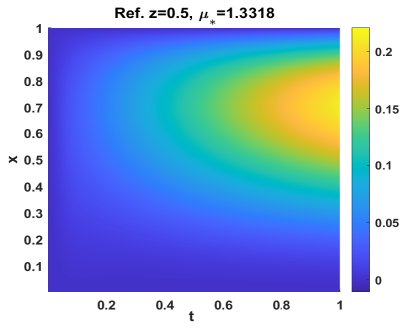
Figure 2: Left: Training sample with 'rand'; Right: Training samples (red), selected samples (green) by KNN, and the target parameter (purple).

In Fig. 2, we generate a size of $\breve{M}=100$ samples for $z$ and $\mu_*$ as the training sets. The left figure presents the training samples, and the right side figure contains the training samples (red), $k$ selected samples (green) by KNN, and the chosen sample (purple). Here, the chosen parameter is $(z,\mu_*)=(0.5,1.3318)$.

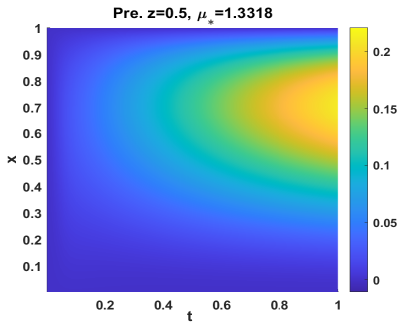
Fig. 3 gives the surrogate model solutions using the POD-KNN and POD-GPR methods with 100 training parameters ($\breve{M}=100$) and the reference solutions using the quadratic FEM to the full-order PDEs with parameters $z=0.5$ and $\mu_*=1.3318$. Specially, $\hat{r}=6$ and $k=5$ are adopted in the POD-KNN method. From these figures, we see that the reference solutions and the predicted solutions change along with time. The predicted solutions and reference solutions have no clear difference with respect to time. It illustrates that both the POD-KNN method and POD-GPR method can generate a parameter-independent surrogate model with respected to $z$ and $\mu_*$.

To provide a clear comparison between the different prediction solutions and the reference solutions, we plot the RE between them and select solutions in several time layers, please see Fig. 4. Fig. 4(a) illustrates the RE between the reference solutions and the predictions from the POD-KNN and POD-GPR models. The RE for the reference solutions and POD-KNN predicted solutions ranges from 0 to 2E-3, while the RE for the reference solutions and POD-GPR predicted solutions ranges from 0 to 7E-3. Fig. 4(b) and Fig. 4(c) plot the reference solutions and the predicted solutions at time $t=0.04,0.2,0.6$. We zoom in on the solutions at $t=0.04$ and $t=0.2$, and place them in corresponding figures, respectively. Although POD-GPR method can approximate the reference solutions well as time increases, the POD-KNN approach exhibits more stability, and its RE value remains within an acceptable range.

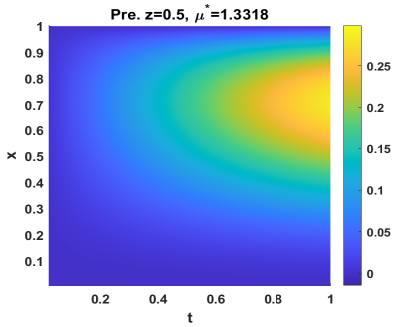
Given the computationally-intensive nature of GPR for nonlocal problems, which can generate dense matrices, the simpler and more efficient KNN approach may be preferred for these types of computations. KNN is a flexible and computationally-lightweight method, making it a suitable choice when the problem size or density of the matrix poses




(a) Reference solutions




(b) Predicted solutions using POD-KNN




(c) Predicted solutions using POD-GPR

Figure 3: (a): Reference solutions (Ref.) at $(z,\mu_*)=(0.5,1.3318)$; (b) and (c): Predicted solutions (Pre.) at $(z,\mu_*)=(0.5,1.3318)$ using POD-KNN and POD-GPR, respectively.

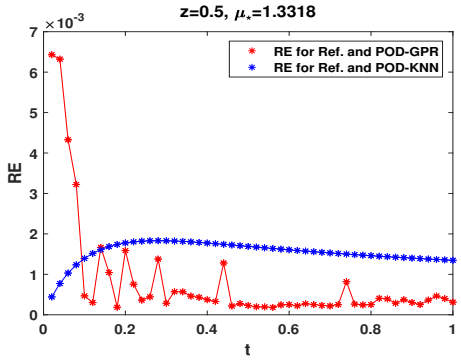
challenges for more complex regression techniques like GPR. Furthermore, we combine the KNN and DMD methods to provide accurate predictions for any time layer.

Then, we plot the MTE values using the POD-KNN method with 450 group test parameters ($N_\mu=450$), and results are shown in Fig. 5. It can be observed that the results of MTE are ranging from 0 to $2E-3$ along with time during the time interval $(0,1]$. This implies that the POD-KNN method has the potential to generate an accurate surrogate model of the parametrized nonlocal PDEs.

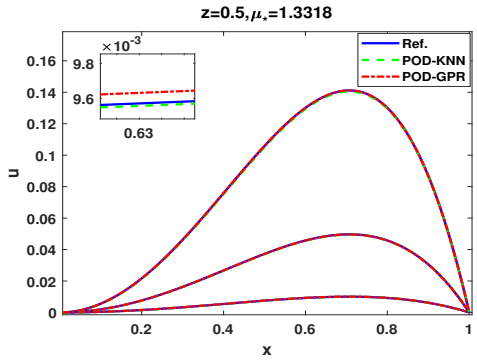
Furthermore, based on the reference and the POD-KNN methods, fixing the final time, Fig. 6 presents the probability density estimate of $u(x,t;\boldsymbol{\mu})$ at the single measurement location, where the variances of $u(\bar{x},\bar{t};\boldsymbol{\mu})$ is maximal (left) or minimal (right) for all $x\in I$. These figures illustrate that the POD-KNN method renders the same probability density as the reference solution.

### 6.3 Linear nonlocal PDEs with the parametrized boundary

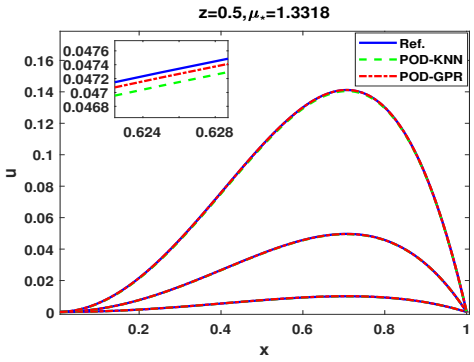
In this example, we consider the nonlocal equation with parametrized boundary using the DMD-KNN and POD-KNN methods. The special expression of the nonlocal model




(a) RE between reference solutions and POD-KNN, POD-GPR predictions




(b) Comparison of the reference solutions and POD-KNN, POD-GPR predictions




(c) Comparison of the reference solutions and POD-KNN, POD-GPR predictions

Figure 4: (a): RE between the predicted solutions and the reference solutions at $(z,\mu_*)=(0.5,1.3318)$; (b) and (c): The predicted solutions and the reference solutions at $t$=0.04, 0.2, 0.6.

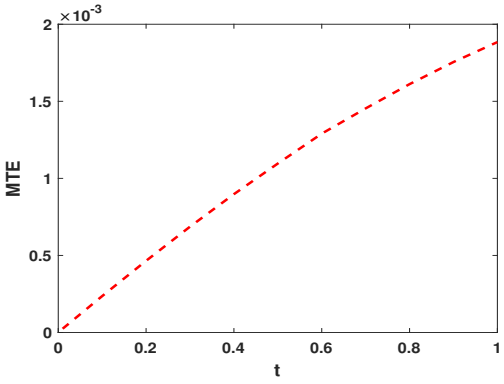



Figure 5: MTE in the time interval $[0,1]$.

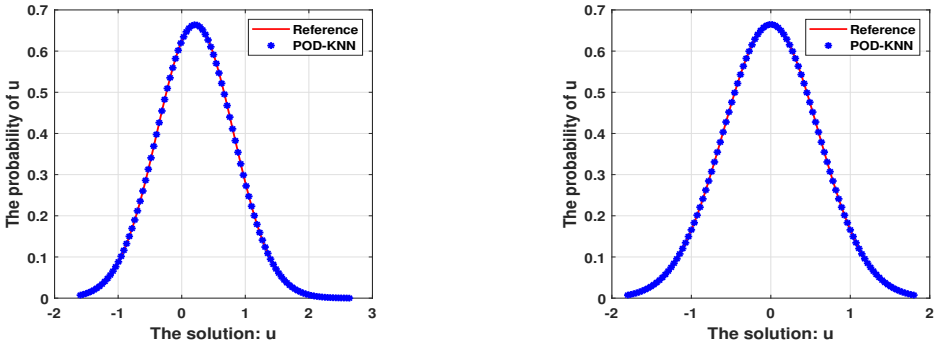



Figure 6: Probability density of $u(\bar{x},\bar{t};\boldsymbol{\mu})$ for reference and the POD-KNN methods, where the variance of $u(\bar{x},\bar{t};\boldsymbol{\mu})$ is maximal (left) and minimal (right) for all $x\in I$.

is

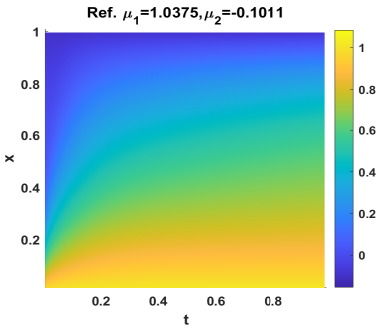
$$u_t-\mathcal{L}_\delta u=f, \tag{6.5}$$

with

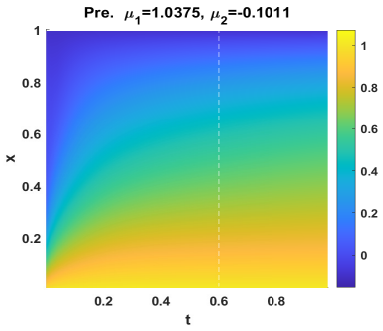
$$u=\begin{cases} u+\mu_1, & (x,\mu_1)\in(-\delta,0]\times(0.9,1.1),\\ u, & x\in(0,1),\\ u+\mu_2, & (x,\mu_2)\in[1,1+\delta)\times(-0.2,0). \end{cases}$$

It is clear from the boundary condition that the parameter is $\boldsymbol{\mu}=(\mu_1,\mu_2)$. By selecting $z=0.5$ in the kernel function (6.2) and setting $h=1/2^7$, $T=1$, and $\Delta t=1E-4$, we solve Eq. (6.5). Here, a set of 60 parameters ($\breve{M}=60$) is randomly generated as the training samples to product the surrogate models using both the DMD-KNN and POD-KNN methods. We first present the reference solution with the chosen parameter $(\mu_1,\mu_2)=(1.0375,-0.1011)$ until $T=1$, as shown in Fig. 7(a). Subsequently, we select the solutions at time $t\in(0,0.6]$ to generate snapshots, and apply DMD-KNN to produce the predicted solution in $t\in(0,1)$ with $\lambda=7$ ($\lambda$ denotes the rank of SVD algorithm in DMD) and $k=5$, as depicted in Fig. 7(b). Comparing Fig. 7(a) and Fig. 7(b), there is no clear difference, indicating the predicted potential of the DMD method outside the time interval. When POD-KNN method is adopted, the snapshots are uniformly distributed in the time domain $(0,1)$, and we obtain the predicted solution with $\hat{r}=7$ and $k$=5, as seen in Fig. 7(c). The same variation over time are observed in Fig. 7(a) and Fig. 7(c).

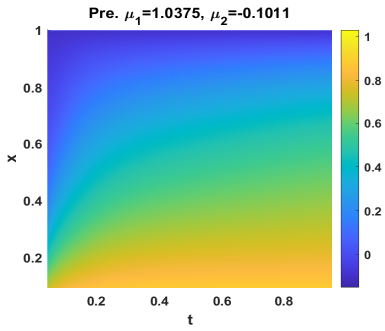
To make the figures in Fig. 7 more clear, we calculate the corresponding RE values, and plot the reference solutions and predicted solutions at several time levels, as shown in Fig. 8. Fig. 8(a) presents the RE values between the reference solutions and the DMD-KNN solutions. Initially, the values increase along with time, then decrease, and finally, it increase again after $t=0.6$. Fig. 8(b) plots the reference solutions and DMD-KNN solutions for several selected time layers from Fig. 7, and it is obvious that the predicted




(a) Reference solutions




(b) Predicted solutions using DMD-KNN



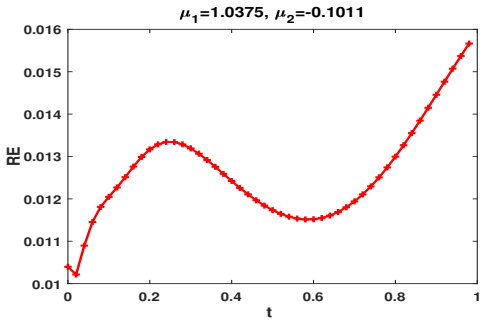

(c) Predicted solutions using the POD-KNN

Figure 7: (a): Reference solutions (Ref.) at $(\mu_1,\mu_2)=(1.0375,-0.1011)$; (b) and (c): Predicted solutions (Pre.) at $(\mu_1,\mu_2)=(1.0375,-0.1011)$.

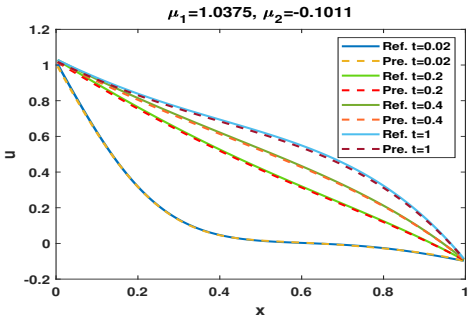
solution by DMD-KNN closely matches the reference solution. Fig. 8(c) depicts the RE values between the reference solution and the POD-KNN solution, and the values first increase, and decrease soon. Fig. 8(d) displays the reference and predicted solutions using POD-KNN for selected time layers from Fig. 7, and it is clear that the predicted solutions using the POD-KNN method also matches well to the reference solutions.

Then, the CPU times are provided for the following two cases: firstly, we present the time costs for the full-order quadratic FEM, and surrogate models produced by the DMD-KNN and POD-KNN MOR methods; secondly, we provide the online phase costs for a single sample in the surrogate model, and the results are displayed in Table 3. In the first case, adding the online phase and offline phase time, then dividing by the total training parameters to derive the average CPU costs. The average CPU costs decrease along with the increase number of the parameters. In the second case, the online phase costs are separately extracted to highlight the surrogate model's efficiency. It is evident that both two surrogate models produce nearly instantaneous predicted solutions.

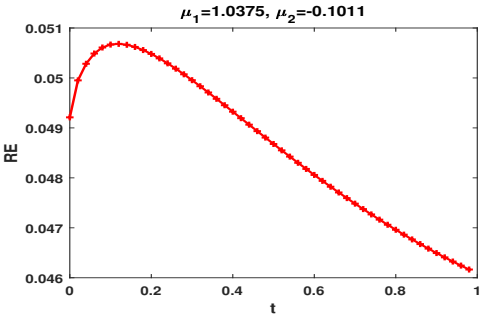
Additionally, we compute the MTE values by full-order equation and surrogate models with same 300 test samples ($N_\mu=300$) during the temporal range $(0,1)$, as shown in Fig. 9. The MTE values produced by surrogate model using DMD-KNN and full-order



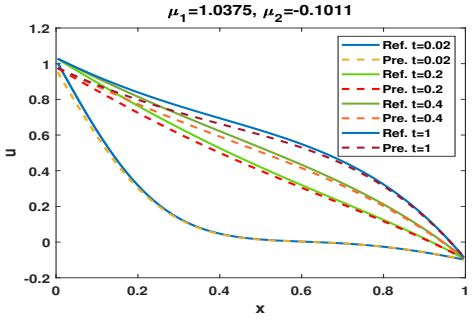

(a) RE between reference solutions and DMD-KNN predictions




(b) Comparison of the reference solutions and DMD-KNN predictions




(c) RE between reference solutions and POD-KNN predictions




(d) Comparison of the reference solutions and POD-KNN predictions

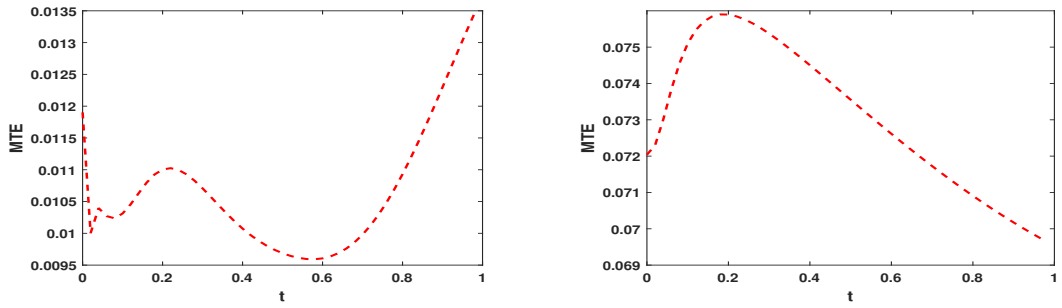
Figure 8: (a) and (c): RE between the predicted solutions and the reference solutions at $(\mu_1,\mu_2)=(1.0375,-0.1011)$; (b) and (d): The predicted solutions and the reference solutions at $t=0.02$, 0.2, 0.4, 1.

Table 3: CPU time costs (s) for solving the parametrized nonlocal problem (6.5).

| Methods | Full-order Quadratic FEM | DMD-KNN | POD-KNN |
|---|---|---|---|
| Average CPU costs | 5.534 | 4.087 | 4.536 |
| Online CPU costs | – | 0.078 | 0.082 |

equation range from 0.0095 to 0.0135. Meanwhile, the MTE values generated by surrogate model using POD-KNN and full-order equation range from 0.069 to 0.076. These two figures demonstrate that both the DMD-KNN and POD-KNN methods can produce the good surrogate models of the parametrized nonlocal PDEs.

In the end of this example, we use the DMD-KNN method and reference method to generate the probability density estimates of $u(x,t;\boldsymbol{\mu})$ at the single measurement location, as shown in Fig. 10, where the variance of $u(\bar{x},\bar{t};\boldsymbol{\mu})$ is maximal (left) or minimal (right) for all $x\in I$ and fixed final time. These figures demonstrate that the DMD-KNN method raises the same probability density as the reference solution.




(a) MTE in the time interval [0,1) (b) MTE in the time interval [0,1)

Figure 9: Left: MTE of the parameters $\mu_1$, $\mu_2$ using the DMD-KNN; Right: MTE of the parameters $\mu_1$, $\mu_2$ using the POD-KNN.

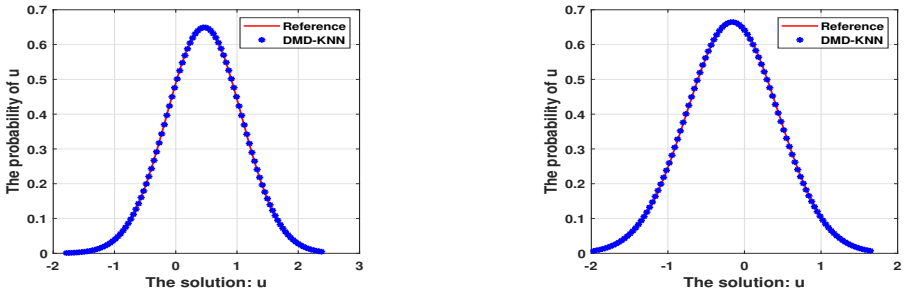



Figure 10: Probability density of $u(\bar{x},\bar{t};\boldsymbol{\mu})$ for reference and the DMD-KNN methods with parameter $\mu_1$ and $\mu_2$, where the variance of $u(\bar{x},\bar{t};\boldsymbol{\mu})$ is maximal (left) and minimal (right) for all $x\in I$.

## 6.4 Test for the parametrized nonlocal Allen-Cahn equation

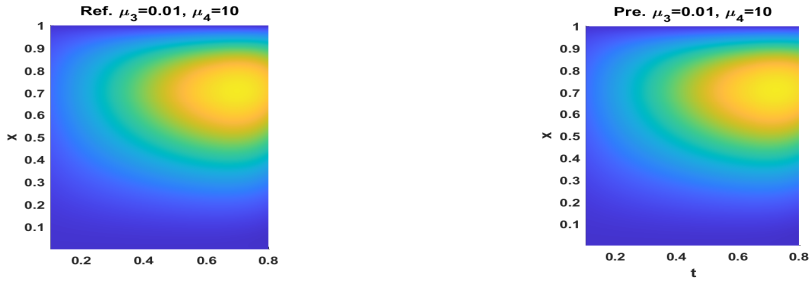
In this subsection, we aim to develop a surrogate model for a nonlocal nonlinear parametrized PDE, using the Allen-Cahn equation as an example, with the following specific expression:

$$u_t=\mu_3\mathcal{L}_\delta u+f(u), \tag{6.6}$$

where parameter $\mu_3$ is usually a very small constant, and we set $\mu_3\in[1E-4,1E-1]$. $f(u)$ is the nonlinear term, expressed by the cubic function,

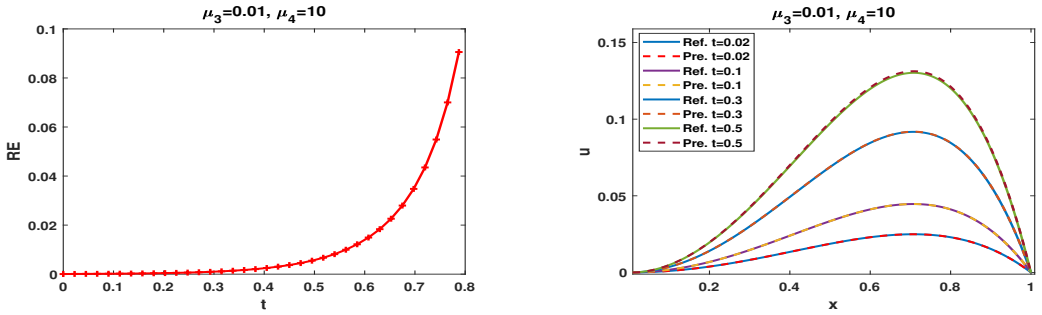
$$f(u)=\mu_4(u-u^3), \tag{6.7}$$

where $\mu_4$ is another parameter, varying in the range $[9.5,10.5]$. The parameter set is $\boldsymbol{\mu}=(\mu_3,\mu_4)$. Fixing the horizon as $\delta=4h$, $h=1/2^7$, $\Delta t=1E-4$, $T=0.8$, and $z=-0.5$ in kernel function (6.2), we obtain the reference solution for the full-order Allen-Cahn equation (6.6).




(a) Reference solutions (b) Predicted solutions

Figure 11: Left: Reference solutions (Ref.) at $(\mu_3,\mu_4)=(0.01,10)$; Right: Predicted solutions (Pre.) at $(\mu_3,\mu_4)=(0.01,10)$.




(a) RE between reference solutions and POD-DEIM-KNN predictions

(b) Comparison of the reference solutions and POD-DEIM-KNN predictions

Figure 12: Left: RE between the reference solution and the predicted solution at $(\mu_3,\mu_4)=(0.01,10)$; Right: The predicted solutions and the reference solutions at $t=0.02$, 0.1, 0.3, 0.5.

Using a set of 80 parameters ($\check{M}=80$) as the training samples to generate the surrogate model. The predicted solution is obtained employing the POD-DEIM-KNN method with $\hat{r}=6$, $m=5$ and $k=5$. The specific surrogate model produced by POD-DEIM is presented in Appendix B. The parameter range is $\mathcal{P}(\mu_3,\mu_4)=[1E-4,1E-1]\times[9.5,10.5]$. In Fig. 11, the reference solution and predicted solution at $T\in(0,0.8]$ are shown with the fixed parameter $(\mu_3,\mu_4)=(0.01,10)$. It is evident from the figures in Fig. 11 that the reference and predicted solutions vary over time.

Following this, we analyze the RE values between the reference solutions and predicted solutions of Fig. 11 and choose several time levels to plot, as shown in Fig. 12. Clearly, the predicted solution closely approximates the reference solution.

Finally, choosing the training samples as $N_\mu=200$, fixing the final time and considering the reference and POD-DEIM-KNN methods, Fig. 13 shows the probability density estimates of $u(x,t;\boldsymbol{\mu})$ at the single measurement, where the variance of $u(\bar{x},\bar{t};\boldsymbol{\mu})$ is maximal (left) or minimal (right) for all $x\in I$. It illustrates that the POD-DEIM-KNN method yields the same probability density as the reference solution.

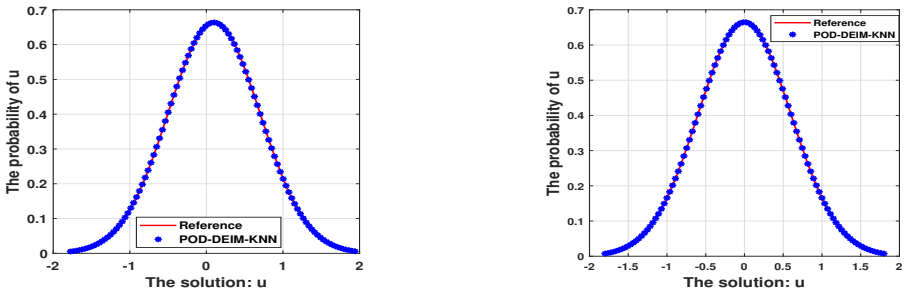



Figure 13: Probability density of $u(\bar{x},\bar{t};\boldsymbol{\mu})$ for reference and the POD-DEIM-KNN methods with parameter $\mu_3$ and $\mu_4$, where the variance of $u(\bar{x},\bar{t};\boldsymbol{\mu})$ is maximal (left) and minimal (right) for all $x\in I$.

# 7 Conclusions

In this paper, we explored the parametrized nonlocal PDEs using the quadratic FEM for spatial discretization in one dimension. More importantly, we developed surrogate models for different parametrized nonlocal model, aiming to achieve the efficient and reliable approximations of the full-order model. The entire calculation process of the surrogate model follows the offline-online decomposition. During the offline phase, we computed the full-order PDEs to generate the snapshots, and used MOR methods to extract the main features of the snapshots. In the online phase, the KNN regression was used to select $k$ nearest samples, and the surrogate models were produced.

Some numerical examples were carried out to test the accuracy of the full-order scheme and the efficiency of the surrogate models. Firstly, we verified that the quadratic FEM achieves a third-order accuracy for the nonlocal model. Subsequently, we focused on the construction of the surrogate models for the parametrized nonlocal PDEs. Three cases are considered based on the location of the parameters, including (1) the parameters were appeared in the kernel function and the linear operator; (2) the parameters existed in the boundaries; (3) the parameters located in the linear and nonlinear terms. We compared the predicted solution and reference solution, and found no significant difference. More specially, we calculated the RE and the MTE values, and the results demonstrated that the surrogate models maintain a high level of accuracy. Furthermore, probability density estimates were provided, and it can be observed that the surrogate model offers a same maximal and minimal variances with the full-order model at the single measurement location. The numerical examples demonstrated the accuracy of our proposed surrogate models.

Despite the sparse loss caused by the quadratic FEM used in the nonlocal model with stochastic parameters, the proposed MOR methods perform well and show the potential for predictions. Naturally, the horizon $\delta$ also can be seen as a parameter in the nonlocal model, which is expected to make no affected for the corresponding surrogate model, and this surrogate model produce predicted solutions approximating the reference solution

very well. Moreover, we intend to explore new approaches for high-order dimensional parametrized nonlocal models to alleviate computational costs.

## Acknowledgments

This work was supported by National Key R&D Program of China (No. 2021YFA1001300), Natural Science Foundation of Hunan Province (Nos. 2021JJ30084, 2022JJ40030), National Natural Science Foundation of China (Nos. 12271150, 12101216), Hong Kong Research Grants Council grant 15303121, and the Hong Kong Polytechnic University Postdoctoral Research Fund 1-W261, Guangdong Basic and Applied Basic Research Foundation (No. 2024A1515012548).

## Appendix A: Using the POD-Galerkin method to generate ROMs

After obtaining the POD basis $\{\gamma_j\}_{j=1}^{\hat{r}}$, we denote $\boldsymbol{U}_{\hat{r}}^1$ ($(2N_x+1)\times\hat{r}$) as the POD basis matrix of the model without boundary, and $\boldsymbol{U}_{\hat{r}}^2$ ($(2N_x+3)\times\hat{r}$) as the POD basis matrix of the model with boundary. This yields the reduced system of Eq. (3.6), represented as

$$\begin{aligned}\bar{\boldsymbol{M}}_1\bar{\boldsymbol{C}}_1^{n+1}+\Delta t\bar{\boldsymbol{B}}_1\bar{\boldsymbol{C}}_1^{n+1}=&\Delta t\bar{\boldsymbol{F}}^{n+1}+\bar{\boldsymbol{M}}_2\bar{\boldsymbol{C}}_2^{n}-\Delta t\bar{\boldsymbol{B}}_2(g_1^h)^{n+1}-\Delta t\bar{\boldsymbol{B}}_3(g_2^h)^{n+1}\\&-\bar{\boldsymbol{M}}_3g^h(x_0',t^{n+1})-\bar{\boldsymbol{M}}_4g^h(x_{2N_x+2}',t^{n+1}),\end{aligned}\tag{A.1}$$

with

$$\begin{aligned}&\bar{\boldsymbol{M}}_1=(\boldsymbol{U}_{\hat{r}}^1)^T\boldsymbol{M}_1\boldsymbol{U}_{\hat{r}}^1,\quad \bar{\boldsymbol{M}}_2=(\boldsymbol{U}_{\hat{r}}^1)^T\boldsymbol{M}_2\boldsymbol{U}_{\hat{r}}^2,\ v\ \bar{\boldsymbol{M}}_3=(\boldsymbol{U}_{\hat{r}}^1)^T\boldsymbol{M}_3,\quad \bar{\boldsymbol{M}}_4=(\boldsymbol{U}_{k}^1)^T\boldsymbol{M}_4,\\&\bar{\boldsymbol{B}}_1=(\boldsymbol{U}_{\hat{r}}^1)^T\boldsymbol{B}_1\gamma_{\hat{r}}^1,\quad \bar{\boldsymbol{B}}_2=(\boldsymbol{U}_{\hat{r}}^1)^T\boldsymbol{B}_2,\quad \bar{\boldsymbol{B}}_3=(\boldsymbol{U}_{\hat{r}}^1)^T\boldsymbol{B}_3,\\&\bar{\boldsymbol{C}}^1=(\boldsymbol{U}_{\hat{r}}^1)^T\boldsymbol{C}_1,\quad \bar{\boldsymbol{C}}^2=(\boldsymbol{U}_{\hat{r}}^2)^T\boldsymbol{C}_2,\quad \bar{\boldsymbol{F}}=(\boldsymbol{U}_{\hat{r}}^1)^T\boldsymbol{F}.\end{aligned}$$

Comparing Eqs. (3.6) and (A.1), the matrix is reduced from $(2N_x+1)\times(2N_x+1)$ dimensions to $\hat{r}\times\hat{r}$ dimensions, where $\hat{r}\ll 2N_x+1$.

## Appendix B: Using POD-DEIM-Galerkin to solve the nonlinear term

The main idea of DEIM is to deal with the nonlinear term of PDEs to construct the surrogate models. Specifically, DEIM is used to approximate a nonlinear function by projecting it onto a low dimensional subspace, where the subspace approximates the space produced by the nonlinear function and is spanned by a basis of dimension $m\ll 2N_x+1$. In this context, the function $f(x,t;\mu_{\xi})$ in (2.5) is considered as the nonlinear term with the fixed parameter $\mu_{\xi}$ for Eq. (2.5), represented by $\tilde{\boldsymbol{F}}(\varsigma)$ with $\varsigma=x,t$ or $\mu_{\xi}$. By projecting $\tilde{\boldsymbol{F}}(\varsigma)$ onto a subspace spanned by $\{\boldsymbol{u}^1,\cdots,\boldsymbol{u}^m\}\subset\mathbb{R}^{2N_x+1}$, we obtain

$$\tilde{\boldsymbol{F}}(\varsigma)\approx\boldsymbol{U}\boldsymbol{c}(\varsigma),$$

where $\boldsymbol{U}=[\boldsymbol{u}^1,\cdots,\boldsymbol{u}^m]\in\mathbb{R}^{(2N_x+1)\times m}$ and $\boldsymbol{c}(\zeta)$ represents the corresponding coefficient vector. The vector $\boldsymbol{c}(\zeta)$ can be uniquely determined by

$$\boldsymbol{P}^T\tilde{\boldsymbol{F}}(\varsigma)=(\boldsymbol{P}^T\boldsymbol{U})\boldsymbol{c}(\varsigma),$$

where $\boldsymbol{P}=[e_{\rho_1},\cdots,e_{\rho_m}]\in\mathbb{R}^{(2N_x+1)\times m}$ and the vector $e_{\rho_i}=[0,\cdots,0,1,0,\cdots,0]^T\in\mathbb{R}^{2N_x+1}$ is the $\rho_i$th column of the identity matrix for $i=1,\cdots,m$. The final approximation of the nonlinear term is presented by

$$\hat{\boldsymbol{F}}(\varsigma)\approx\boldsymbol{U}\boldsymbol{c}(\varsigma)=\boldsymbol{U}(\boldsymbol{P}^T\boldsymbol{U})^{-1}\boldsymbol{P}^T\tilde{\boldsymbol{F}}(\varsigma).$$

For more information about DEIM, please refer to [4]. The detailed process of reducing the order of the nonlinear parametrized nonlocal PDEs is omitted, and the surrogate model produced by the POD-DEIM-Galerkin method is presented as follows,

$$\begin{aligned}&\bar{\boldsymbol{M}}_1\bar{\boldsymbol{C}}_1^{n+1}+\Delta t\bar{\boldsymbol{B}}_1\bar{\boldsymbol{C}}_2^{n+1}\\=&\Delta t\cdot\hat{\boldsymbol{F}}^{n+1}+\bar{\boldsymbol{M}}_2\bar{\boldsymbol{C}}_1^{n}-\Delta t\bar{\boldsymbol{B}}_2(g_1^h)^{n+1}-\Delta t\bar{\boldsymbol{B}}_3(g_2^h)^{n+1}\\&-\bar{\boldsymbol{M}}_3g^h(x_0',t^{n+1})-\bar{\boldsymbol{M}}_4g^h(x_{2N_x+2}',t^{n+1}).\end{aligned}$$

**References**


[1] M. Belkin, D.J. Hsu and P. Mitra, Overfitting or perfect fitting? risk bounds for classification and regression rules that interpolate, Adv. Neural Inf. Processing Syst., 31 (2018), 2306-2317.

[2] M.S. Breitenfeld, P.H. Geubelle, O. Weckner and S.A. Silling, Non-ordinary state-based peridynamic analysis of stationary crack problems, Comput. Methods Appl. Mech. Engrg., 272 (2014), 233-250.

[3] A. Buhr and K. Smetana, Randomized local model order reduction, SIAM J. Sci. Comput. 40 (2018) A2120-A2151.

[4] S. Chaturantabut and D.C. Sorensen, A state space error estimate for POD-DEIM nonlinear model reduction, SIAM J. Numer. Anal., 50 (2012), 46-63.

[5] P. Chen, A. Quarteroni and G. Rozza, A weighted reduced basis method for elliptic partial differential equations with random input data, SIAM J. Numer. Anal., 51 (2013), 3163-3185.

[6] M.J. Colbrook, The mpDMD algorithm for data-driven computations of measure-preserving dynamical systems, SIAM J. Numer. Anal., 61 (2023), 1585-1608.

[7] Z. Drmač, I. Mezić and R. Mohr, On least squares problems with certain vandermonde-khatri-rao structure with applications to DMD, SIAM J. Sci. Comput., 42 (2020), A3250-A3284.

[8] Q. Du, M. Gunzburger, R.B. Lehoucq and K. Zhou, Analysis and approximation of nonlocal diffusion problems with volume constraints, SIAM Rev., 54 (2012), 667-696.

[9] Q. Du, M. Gunzburger, R.B. Lehoucq and K. Zhou, A nonlocal vector calculus, nonlocal volume-constrained problems, and nonlocal balance laws, Math. Models Methods Appl. Sci., 23 (2013), 493-540.

[10] Q. Du, L. Ju, X. Li and Z. Qiao, Stabilized linear semi-implicit schemes for the nonlocal Cahn-Hilliard equation, J. Comput. Phys., 363 (2018), 39-54.

[11] Q. Du, Nonlocal Modeling, Analysis, and Computation, CBMS-NSF Regional Conf. Ser. in Appl. Math. 94, SIAM, Philadelphia, 2019.

[12] Q. Du, L. Ju, X. Li and Z. Qiao, Maximum principle preserving exponential time differencing schemes for the nonlocal Allen-Cahn equation, SIAM J. Numer. Anal., 57 (2019), 875-898.
[13] Q. Du, L. Ju and J. Lu, Analysis of fully discrete approximations for dissipative systems and application to time-dependent nonlocal diffusion problems, J. Sci. Comput., 78 (2019), 1438-1466.
[14] Q. Du, Y. Tao, X. Tian and J. Yang, Asymptotically compatible discretization of multidimensional nonlocal diffusion models and approximation of nonlocal Green's functions, IMA J. Numer. Anal., 39 (2019), 607-625.
[15] Q. Du, L. Ju, X. Li and Z. Qiao, Maximum bound principles for a class of semilinear parabolic equations and exponential time-differencing schemes, SIAM Rev., 63 (2021), 317-359.
[16] S. Fresca and A. Manzoni, POD-DL-ROM: Enhancing deep learning-based reduced order models for nonlinear parametrized PDEs by proper orthogonal decomposition, Comput. Methods Appl. Mech. Engrg., 388 (2022), 114181.
[17] Z. Gao, Y. Lin, X. Sun and X. Zeng, A reduced order method for nonlinear parameterized partial differential equations using dynamic mode decomposition coupled with *k*-nearest-neighbors regression, J. Comput. Phy., 452 (2022), 110907.
[18] A.K. Ghosh, On nearest neighbor classification using adaptive choice of k, J. Comput. Graph. Stat., 16 (2007), 482-502.
[19] Q. Guan, M. Gunzburger. C.G. Webster and G. Zhang, Reduced basis methods for nonlocal diffusion problems with random input data, Comput. Methods Appl. Mech. Engrg., 317 (2017), 746-770.
[20] L. Jiang and Q. Li, Model's sparse representation based on reduced mixed GMsFE basis methods, J. Comput. Phys., 338 (2017), 285-312.
[21] P. Kumari, B. Bhadriraju, Q. Wang and J.S.-Il Kwon, Development of parametric reduced-order model for consequence estimation of rare events, Chem. Engrg. Res. Des., 169 (2021), 142-152.
[22] Q. Li and L. Jiang, A novel variable-separation method based on sparse and low rank representation for stochastic partial differential equations, SIAM J. Sci. Comput., 39 (2017), A2879-A2910.
[23] Q. Li, C. Liu, M. Li and P. Zhang, An adaptive method based on local dynamic mode decomposition for parametric dynamical systems, Commun. Comput. Phys., 35 (2024), 38-69.
[24] X. Li, Z. Qiao and H. Zhang, An conditionally energy stable finite difference scheme for a stochastic Cahn-Hilliard equation, Sci. China. Math. 59 (2016), 1815–1834.
[25] X. Li, Z. Qiao and C. Wang, Convergence analysis for a stabilized linear semi-implicit numerical scheme for the nonlocal Cahn–Hilliard equation, Math. Comp., 90 (2021), 171-188.
[26] Q. Li and P. Zhang, A variable-separation method for nonlinear partial differential equations with random inputs, SIAM J. Sci. Comput., 42 (2020), A723-A750.
[27] J. Lu and Y. Nie, A reduced-order fast reproducing kernel collocation method for nonlocal models with inhomogeneous volume constraints, Comput. Math. Appl., 121 (2022), 52-61.
[28] A.J. Majda and D. Qi, Strategies for reduced-order models for predicting the statistical responses and uncertainty quantification in complex turbulent dynamical systems, SIAM Rev., 60 (2018), 491-549.
[29] L. Ma, R. Li, F. Zeng, L. Guo and G.E. Karniadakis, Bi-Orthogonal fPINN: a physics-informed neural network method for solving time-dependent stochastic fractional PDEs, Commun. Comput. Phys., 34 (2023), 1133-1176.
[30] E. Musharbash, F. Nobile and T. Zhou, Error analysis of the dynamically orthogonal approximation of time dependent random PDEs, SIAM J. Sci. Comput., 37 (2015), A776–A810.

[31] C. Nan and H. Song, The high-order maximum-principle-preserving integrating factor Runge-Kutta methods for nonlocal Allen-Cahn equation, J. Comput. Phys., 456 (2022), 111028.
[32] S. Pagani, A. Manzoni, A. Quarteroni, Numerical approximation of parametrized problems in cardiac electrophysiology by a local reduced basis method, Comput. Methods Appl. Mech. Engrg., 340 (2018), 530-558.
[33] A.F. Psaros, X. Meng, Z. Zou, L. Guo and G.E. Karniadakis, Uncertainty quantification in scientific machine learning: methods, metrics, and comparisons, J. Comput. Phys., 477 (2023), 111902.
[34] M. Rathinam and L.R. Petzold, A new look at proper orthogonal decomposition, SIAM J. Numer. Anal., 41 (2003), 1893-1925.
[35] E.W. Sachs and S. Volkwein, POD Galerkin approximations in PDE-constrained optimizationz, GAMM Mitt. Ge. Math. Mech. Angew., 33 (2010), 194-208.
[36] A. Schmidt, A. Potschka, S. Körkel and H.G. Bock, Derivative-extended POD reducted-order modeling for parameter estimation, SIAM J. Sci. Comput., 35 (2013), A2696-A2717.
[37] L. Sirovich, Turbulence and the dynamics of coherent structures, Part 1: Coherent structures, Quart. Appl. Math., 45 (1987), 561-571.
[38] X. Tian and Q. Du, Analysis and comparison of different approximations to nonlocal diffusion and linear peridynamic equations, SIAM J. Numer. Anal., 51 (2013), 3458-3482.
[39] X. Tian and Q. Du, Asymptotically compatible schemes and applications to robust discretization of nonlocal models, SIAM J. Numer. Anal., 52 (2014), 1641-1665.
[40] T. Tang and T. Zhou, Recent developments in high order numerical methods for uncertainty quantification, Sci. china. Math. 45 (2015), 891–928.
[41] M.O. Williams, P.J. Schmid and J.N. Kutz, Hybrid reduced-order integration with proper orthogonal decomposition and dynamic mode decomposition, Multiscale Model. Simul., 11 (2013), 522-544.
[42] D.R. Witman, M. Gunzburger and J. Peterson, Reduced-order modeling for nonlocal diffusion problems, Int. J. Numer. Meth. Fluids, 83 (2017), 307-327.
[43] Z. Wu and Z. Zhang, An iterative algorithm for POD basis adaptation in solving parameteric convection-diffusion equation, Comput. Mehtods Appl. Mech. Engrg., 391 (2022), 114498.
[44] Y. Xing, Q. Song and G. Cheng, Benefit of interpolation in nearest neighbor algorithms, SIAM J. Math. Data Sci., 4 (2022), 935-956.
[45] Y. Zhao, Z. Mao, L. Guo,Y. Tang and G. E. Karniadakis, A spectral method for stochastic fractional PDEs using dynamically-orthogonal/bi-orthogonal decomposition, J. Comput. Phys., 461(2022), 111213.
[46] D. Zhang, L. Lu, L. Guo and G.E. Karniadakis, Quantifying total uncertainty in physics-informed neural networks for solving forward and inverse stochastic problems, J. Comput. Phys., 397 (2019), 108850.
[47] Z. Zou, D. Kouri and W. Aquino, An adaptive local reduced basis method for solving PDEs with uncertain inputs and evaluating risk, Comput. Mehtods Appl. Mech. Engrg., 345 (2019), 302-322.