UC San Diego
UC San Diego Previously Published Works

Title

A physics-constrained data-driven approach based on locally convex reconstruction for noisy database

Permalink

https://escholarship.org/uc/item/0xs1002k

Authors

He, Qizhi
Chen, Jiun-Shyan

Publication Date

2020-05-01

DOI

10.1016/j.cma.2019.112791

Peer reviewed

A Physics-Constrained Data-Driven Approach Based on Locally Convex Reconstruction for Noisy Database

Qizhi He, Jiun-Shyan Chen[*]

Department of Structural Engineering, University of California, San Diego, La Jolla, CA, 92093, USA

Abstract

Physics-constrained data-driven computing is an emerging hybrid approach that integrates universal physical laws with data-driven models of experimental data for scientific computing. A new data-driven simulation approach coupled with a locally convex reconstruction, termed the local convexity data-driven (LCDD) computing, is proposed to enhance accuracy and robustness against noise and outliers in data sets in the data-driven computing. In this approach, for a given state obtained by the physical simulation, the corresponding optimum experimental solution is sought by projecting the state onto the associated local convex manifold reconstructed based on the nearest experimental data. This learning process of local data structure is less sensitive to noisy data and consequently yields better accuracy. A penalty relaxation is also introduced to recast the local learning solver in the context of non-negative least squares that can be solved effectively. The reproducing kernel approximation with stabilized nodal integration is employed for the solution of the physical manifold to allow reduced stress-strain data at the discrete points for enhanced effectiveness in the LCDD learning solver. Due to the inherent manifold learning properties, LCDD performs well for high-dimensional data sets that are relatively sparse in real-world engineering applications. Numerical tests demonstrated that LCDD enhances nearly one order of accuracy compared to the standard distance-minimization data-driven scheme when dealing with noisy database, and a linear exactness is achieved when local stress-strain relation is linear.

Keywords: data-driven computing; locally convex reconstruction; manifold learning; noisy data; local convexity data-driven (LCDD); reproducing kernel (RK) approximation

---

[*] Correspondence to: Jiun-Shyan Chen, Structural Engineering Department, University of California, San Diego (UCSD), La Jolla, CA 92093, USA. E-mail: js-chen@ucsd.edu

1. Introduction

With the proliferation of high-resolution datasets and the significant advances in numerical algorithms, the emerging idea by utilizing both data-driven models and physical models simultaneously to enhance traditional scientific computing and engineering design procedures [1,2] has attracted increasingly attentions. This general approach is usually termed as data-driven modeling [3] or data-driven engineering science. Data-driven modeling has a close connection with the various areas such as statistics, data mining, and machine learning, which allow the extraction of insightful information or the hidden structures from large volumes of data [4] for enhanced scientific computing. The data-driven approaches, such as machine learning [5,6], have been widely applied to computational biological [7] and medical diagnosis [8], material informatics [9,10], and other predictive physics problems [4,11].

Recently, these approaches have been extended to the field of engineering mechanics, such as learning constitutive models in solid mechanics [12–14], surrogate models in fluid mechanics [15–17] and physical models or governing equations purely extracted from the collected data [18–20]. In conjunction with machine learning techniques such as manifold learning [21] or neural networks [22], the recent studies [23–25] offer a new paradigm for data-driven computing for various applications such as design of materials [26]. There is a vast body of literature devoted to these subjects, including the recent developments based on nonlinear dimensionality reduction [24], nonlinear regression, deep learning [27–29], among others.

Nevertheless, pure data-driven methodology in the area of simulation-based engineering sciences (SBES) [30] is ineffective since in many physical systems well-accepted physical laws exist while useful data in SBES are very expensive to acquire [20,31]. Thus, it is imperative to develop data-driven simulation approaches that can leverage the physical principles with limited data for highly complex systems. A solution to develop effective predictive models for complex real-world problems is to combine physics-based models with data-driven techniques under a hybrid computational framework. There are three types of hybrid physics-data approaches, depending on the roles of physics laws and data play in the hybrid model. The first approach enforces known physical constraints into data-driven models [32,33], which can be considered as a data-fit type surrogate model. In the second approach, on the contrary, the existing physical models are enriched by the information learned from data. This general framework can be used

for obtaining data-enhanced physical models [34,35], online updating dynamical system in a manner similar to data assimilation [36], or constructing reduced-order model [25,37–40]. The third approach is to apply data-driven models and physical models separately to approximate different aspects of the physical system and be connected consistently to perform numerical simulation.

The third class of methods is particularly attractive for computational mechanics because it preserves the well-accepted physical laws under the variational framework and the prerequisite of big data is much relaxed. In contrast, most other data-driven methods for solid or fluid simulation directly construct machine learning surrogates to link the input-output relation with approximated physics laws [17, 24, 33] or with the constitutive models replaced by supervised learning models such as neural networks [12-14, 27], which could over-parameterize the material relation and lead to numerical instability. In addition, the setting of training and architecture hyperparameters for neural networks is not straightforward.

Under the framework of the third approach mentioned above, Kirchdoerfer and Ortiz [41–43] have proposed a material model-free data-driven method, so called distance-minimizing data-driven computing (DMDD), for modeling elasticity problems. This data-driven method enforces equilibrium and compatibility and directly utilizes the material database, e.g. stress and strain data, under a modified variational framework, aiming to eschew the empirical models that inevitably involve incomplete experimental information [41,42] and the process of material parameter identification [44–46] that remains numerically intractable. In DMDD, the data-driven problem is solved by minimizing the distance between the computed physical solutions (i.e. the set of equilibrium admissible stress and kinematically admissible strain jointly) and a given set of experimental data under a proper energy norm. A similar idea was proposed by Ibañez et al. [47] where manifold learning techniques are applied to material database to construct the tangent stiffness approximation of constitutive relation, with which the convergent solution could be attained by using directional search solvers [48,49]. In these methods, the data selection or automated machine learning techniques on material data are carried out during the computation of the associated initial-boundary-value problem, thus bypassing the traditional construction of constitutive models. Again, these methods fall into the third class of the data-driven approaches discussed above, and they are usually defined as data-driven computational mechanics (DDCM)

[41,47]. This data-driven paradigm has been recently extended to dynamics [43], nonlinear elasticity [35,50,51,82], material identification [52], and data completion [53]. Overall, the key idea of the above mentioned methods is to seek the intersection of the hidden constitutive (material) manifold represented by experimental data and the physical manifold by using iterative processes with appropriate search directions, as shown in **Fig. 1**.
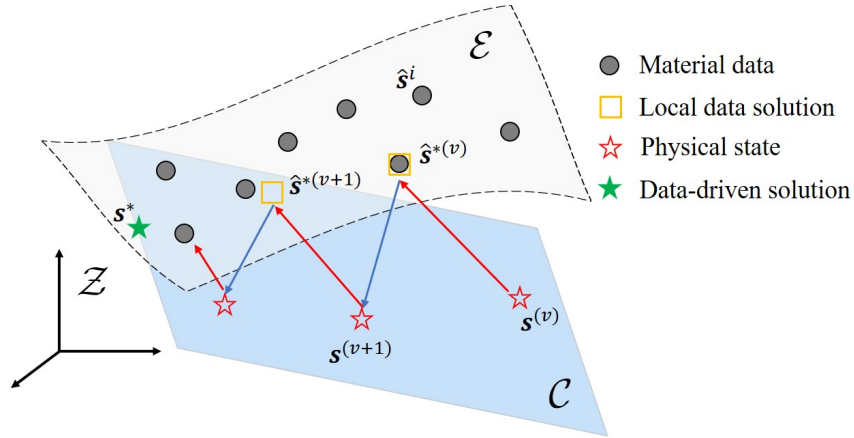

Fig. 1. Schematics of data-driven computing for predictive modeling, where a set of data points $\{\hat{s}^i, i = 1, 2, ...\}$ are given to represent the material behavior, E is an imaginary manifold of the material database representing the underlying constitutive relation, and C is the physical manifold of admissible stress and strain states $s = (\varepsilon, \sigma)$ satisfying equilibrium and compatibility. The data-driven solution s^* is solved by a fixed point iteration that searches the physical states $s^{(v)} \in C$ and the local data solutions $\hat{s}^{*(v)} \in E$ via iterative projections, where the subscript 'v' is the step indicator. In this study, only stresses and strains are considered in the material database.

Despite the major advancement made in the field, it remains challenging in dealing with noisy and sparse data [1]. The standard DMDD paradigm [41] is shown sensitive to noisy data and outliers [42,54], while the approaches based on manifold learning [47] or local regression [55] may fail to converge due to the over-relaxed manifold constrsuction and lack of convexity. To enhance robustness, the DMDD approach was extended to the max-ent data driven computing [42], which utilizes entropy estimation to analyze the statistics information of data. However, as

simulated annealing algorithms are used to solve the resulting data-driven minimization problem, the solution procedures become computationally intensive. Alternatively, Ayensa-Jimenez et al. [54] proposed to consider data uncertainty by explicitly incorporating statistical quantities into the standard DMDD approach and defined a stochastic analogous problem. In this approach, the expected value and variance-covariance matrix need to be estimated and the local data structure is still not taken into consideration and hence it becomes ineffective in dealing with high-dimensional dataset.

In this paper, we propose a novel data-driven approach which utilizes the intrinsic local data structure to enhance accuracy and robustness against noisy data while computationally feasible. By assuming that each data point and its neighbors lie on or close to a locally linear patch of the manifold, the proposed approach approximates the underlying constitutive manifold near the physically admissible state by locally constructing a convex envelop based on the associated experimental neighbor data points. As a result, the proposed approach can utilize the local data structure without explicitly constructing the local manifold or regression models needed in the approaches [47,54,55] utilizing the associated tangent spaces for data-driven iterations. With this locally convex construction, the solution space for searching optimum local data is regularized onto a bounded, continuous, and convex subset (polytope) for enhanced robustness and convergence stability in data-driven computing. The proposed approach is, thus, referred to as local convexity data-driven (LCDD) computing. In this approach, a cluster of experimental data associated with the physical solution (e.g., the pair of strains and stresses) is first identified by the *k-nearest neighbor* (*k*-NN) algorithm, and the optimum data solution is searched within the associated locally convex hull instead of the discrete material set. To solve this local search problem efficiently, we recast the approach into a *non-negative least squares* (NNLS) problem [56] by introducing the invariance constraint into the objectivity function. Because of the inherited manifold learning capacity in the NNLS solvers, the proposed LCDD permits the locally linear approximation for the underlying material manifold, which means that LCDD could reproduce the solutions given by the classical model-based simulation if the constitutive relation represents a locally linear pattern.

On the other hand, LCDD can be viewed as a generalization of DMDD by equipping a suitable manifold learning technique that naturally takes the local data information into account

5

and retains a simple computing framework. In the solution phase on the physical manifold, a constrained minimization problem is solved by introducing a reproducing kernel approximation (RK) [57,58] in conjunction with a stabilized conforming nodal integration (SCNI) [59] such that the displacements, stresses and strains are computed at the nodal points. This approach significantly reduces the needed search for the optimal stress-strain data from the data set. The employment of the RK approximation also introduces higher order smoothness to the solution space of the physical manifold, making it consistent to the continuous and convex solution space of the regularized LCDD learning solver. It is noted that there is no assumption of isotropy and homogeneity in the proposed LCDD data-driven computational framework. The learning algorithm can identify the intrinsic properties of the given dataset. In this study, we only consider the modeling of homogeneous material, that is, the same dataset is used to characterize the material behavior at every evaluation point over the domain.

The objective of the present work is to study the main issues of data-driven approaches when dealing with noisy data in high-dimensional space. The paper is organized as follows. In Section 2, a generalized data-driven computational formalism is reviewed. In Section 3, locally convex reconstruction is introduced and the local manifold learning for data-driven solver formulated under the NNLS framework is presented. Section 4 provides numerical tests of truss structures to demonstrate the effectiveness of LCDD against noisy data. In Section 5, continuum mechanics with elastic solid is considered to assess the accuracy and convergence properties of LCDD when the noisy data is of high-dimensional phase space. Finally, concluding remarks and discussions are given in Section 6.

6

2. Physics-constrained data-driven modeling

In this section, we revisit the methodology of data-driven computational mechanics (DDCM) in [41,47,60] and formulate the associated data-driven problem under the variational framework for distance-minimizing between a physical set and an experimental data set.

The deformation of an elastic solid occupying the domain Ω bounded by Neumann boundary Γ_t and Dirichlet boundary Γ_u can be described by two basic laws:

(i) Equilibrium

$$\nabla \cdot \boldsymbol{\sigma} + \boldsymbol{b} = \boldsymbol{0}, \text{ in } \Omega, \tag{1a}$$

$$\boldsymbol{\sigma} \cdot \boldsymbol{n} = \bar{\boldsymbol{t}}, \text{ on } \Gamma_t, \tag{1b}$$

(ii) Compatibility

$$\varepsilon = \frac{1}{2}\left(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^T\right), \text{ in } \Omega, \tag{2a}$$

$$\boldsymbol{u} = \bar{\boldsymbol{u}}, \text{ on } \Gamma_u, \tag{2b}$$

where $\boldsymbol{\sigma}$, $\boldsymbol{u}$, ε, and $\boldsymbol{b}$ are the Cauchy stress, displacement, strain, and body force, respectively, and $\bar{\boldsymbol{u}}$ and $\bar{\boldsymbol{t}}$ are the prescribed displacement and the applied traction on Γ_u and Γ_t, respectively.

Traditionally, to solve a boundary value problem (BVP) in (1) and (2), the constitutive law, e.g. $\boldsymbol{\sigma} = \boldsymbol{\sigma}(\varepsilon)$, is required to relate $\boldsymbol{\sigma}$ and ε. In data-driven computing, the equilibrium and compatibility equations in (1) and (2) are solved numerically, while the constitutive relation is given by a set of experimental data, $E = \{(\hat{\varepsilon}^i, \hat{\sigma}^i), i = 1, ..., p\}$, where p is the number of measurement data collected from experiments.

Remark 2.1 *For modeling homogeneous material, the same dataset* E *is used to characterize the material behavior at every evaluation point over the domain. Note that this data-driven computing can be applied to heterogeneous materials if space-dependent databases are available. In this approach, there is no predefined material models or material parameters that need to be identified, which makes data-driven computing different from the classical simulation methods and material identification problems.*

It is convenient to introduce the notion of phase space Z as the space of the strain-stress pairs (ε, σ), and denote C as the admissible set for elements $(\varepsilon, \sigma) \in Z$ that satisfy the physical constraints in (1) and (2), which is also called the physical manifold. Ideally, the data-driven solution is the intersection of the global data set E_{em} and the physical manifold set C, i.e. $E_{em} \cap C$, where $E_{em} = E \times ... \times E \subset Z$ denotes the ensemble of the experimental set E over Ω. Since E consists of a finite set of discrete data points which could lead to non-existence of the intersection $E_{em} \cap C$, a distance-minimizing relaxation is usually employed.

2.1. Distance-minimizing data-driven problem

Data-driven computing [41,47] and data-enabled applications, such as dynamics data-driven application systems (DDDAS) [1] and parameter identification for pre-defined material models [44–46], introduce distance-minimization between the simulation data and the measurement data. The main difference between these approaches is that data-driven computing is a forward problem while parameter identification is an inverse problem for material calibration. We refer interested readers to the literatures for more details of parameter identification [44–46]. Data-driven computing can be stated as one of the following double-minimization problems:

$$\min_{(\hat{\varepsilon}, \hat{\sigma}) \in E_{em}} \min_{(u, \sigma) \in C_u \times C_\sigma} H(u, \sigma, \hat{\varepsilon}, \hat{\sigma}) \text{ or } \min_{(u, \sigma) \in C_u \times C_\sigma} \min_{(\hat{\varepsilon}, \hat{\sigma}) \in E_{em}} H(u, \sigma, \hat{\varepsilon}, \hat{\sigma}) \qquad (3)$$

where H is a given functional to define a distance measure, which is to be elaborated in the next section, and C_σ and C_u denote the sets of *equilibrium admissible* stress fields and *kinematically admissible* displacement fields, respectively, i.e.,

$$C_\sigma = \{\tau \in V_\sigma \mid \nabla \cdot \tau + b = 0 \text{ in } \Omega, \text{ and } \tau \cdot n = \bar{t} \text{ on } \Gamma_t\}, \qquad (4a)$$

$$C_u = \{v \in V_u \mid v = \bar{u} \text{ on } \Gamma_u\}, \qquad (4b)$$

in which $V_\sigma = [L^2(\Omega)]^6$ is the symmetric stress space, and $V_u = [H^1(\Omega)]^3$ is the displacement space. Then the physical manifold set is defined as

$$C = \{(\varepsilon[u], \sigma) \mid u \in V_u, \sigma \in V_\sigma\}. \qquad (4c)$$

Note that strain ε is obtained from the displacement $\boldsymbol{u} \in C_u$ using (2a), which is denoted by $\varepsilon = \varepsilon[\boldsymbol{u}]$.

Corresponding to the strain-stress state $(\varepsilon, \sigma) \in C$ obtained from the physical manifold C, $(\hat{\varepsilon}, \hat{\sigma}) \in E_{em}$ is used to denote the data from the experimental data set E_{em}. As illustrated in **Fig. 1**, the data-driven computing in (3) is to find the state (ε, σ) constrained to the physical set C while closest to the dataset E_{em} under a certain "distance" measure defined by the functional H, such that the system response is determined directly from the experimental data without specifying any constitutive models.

The data-driven problem in (3) can be decomposed into a two-step problem:

Global step: $\quad J(\hat{\varepsilon}, \hat{\sigma}) = \min_{(\boldsymbol{u}, \sigma) \in C_u \times C_\sigma} H(\boldsymbol{u}, \sigma, \hat{\varepsilon}, \hat{\sigma}),$ (5a)

Local step: $\quad (\hat{\varepsilon}^*, \hat{\sigma}^*) = \arg \min_{(\hat{\varepsilon}, \hat{\sigma}) \in E_{em}} J(\hat{\varepsilon}, \hat{\sigma})$ (5b)

where $(\hat{\varepsilon}^*, \hat{\sigma}^*)$ is the optimum experimental point closest to the computed state (ε, σ) given in (5 a). From an optimization perspective, the solution procedures of this data-driven problem involve an alternate-direction search where a minimization with respect to $(\boldsymbol{u}, \sigma)$ is followed by a minimization with respect to $(\hat{\varepsilon}, \hat{\sigma})$, denoted as a *global step* and a *local step*, respectively.

Compared to the problem setting in material parameter identification [45], the data-driven computing in (5) does not rely on any pre-assumed elasticity tensor to relate ε and σ. Instead, it iteratively searches a representative stress-strain pair from the experimental dataset for performing simulation.

2.2. Data-driven solver

The norm $\|\boldsymbol{g}\|_Z$ associated to the phase space Z has been defined as a combination of the energy-like and complementary energy-like functional [41] as follows:

$$\| (\varepsilon, \sigma) \|_Z^2 = \frac{1}{2} \int_\Omega \varepsilon : \boldsymbol{M}^\varepsilon : \varepsilon + \sigma : \boldsymbol{M}^\sigma : \sigma \, \mathrm{d}\Omega, \quad (6)$$

where $\boldsymbol{M}^\varepsilon$ and $\boldsymbol{M}^\sigma$ are two tensors to balance the contribution of the strain and stress data measured in different physical units.

For numerical implementation, the state variables $(\boldsymbol{\varepsilon}, \boldsymbol{\sigma})$ are computed at integration points $(\boldsymbol{\varepsilon}_\alpha, \boldsymbol{\sigma}_\alpha) \equiv (\boldsymbol{\varepsilon}(\boldsymbol{x}_\alpha), \boldsymbol{\sigma}(\boldsymbol{x}_\alpha)) \in {}^\circ{}^q \times {}^\circ{}^q$, where $\{\boldsymbol{x}_\alpha\}_{\alpha=1}^m$ are the coordinates of the m integration points (i.e., stress-strain evaluation points) and q is the dimension of stress and strain. As such, we denote $\{(\boldsymbol{\varepsilon}_\alpha, \boldsymbol{\sigma}_\alpha)\}_{\alpha=1}^m \in Z^h$, where $Z^h \subset Z$ is the discrete counterpart of the phase space. Correspondingly, the distance minimization in the local step searches for the local data solution $(\hat{\boldsymbol{\varepsilon}}_\alpha, \hat{\boldsymbol{\sigma}}_\alpha) \in E$ at every integration point $\boldsymbol{x}_\alpha$, $\alpha = 1, ..., m$. In the subsequent discussion, we define $\boldsymbol{s}_\alpha = [\boldsymbol{\varepsilon}_\alpha^T \; \boldsymbol{\sigma}_\alpha^T]^T \in {}^\circ{}^{2q}$ and $\hat{\boldsymbol{s}}_\alpha = [\hat{\boldsymbol{\varepsilon}}_\alpha^T \; \hat{\boldsymbol{\sigma}}_\alpha^T]^T \in {}^\circ{}^{2q}$ as the computational and experimental strain-stress pairs, respectively, in the local phase space.

A functional H defined as the discrete form of (6) to measure the distance between $\{(\boldsymbol{\varepsilon}_\alpha, \boldsymbol{\sigma}_\alpha)\}_{\alpha=1}^m$ and $\{(\hat{\boldsymbol{\varepsilon}}_\alpha, \hat{\boldsymbol{\sigma}}_\alpha)\}_{\alpha=1}^m$ is given as

$$H(\boldsymbol{u}, \boldsymbol{\sigma}, \hat{\boldsymbol{\varepsilon}}, \hat{\boldsymbol{\sigma}}) = \| (\boldsymbol{\varepsilon} - \hat{\boldsymbol{\varepsilon}}, \boldsymbol{\sigma} - \hat{\boldsymbol{\sigma}}) \|_Z^2 \approx \sum_{\alpha=1}^m d^2(\boldsymbol{s}_\alpha, \hat{\boldsymbol{s}}_\alpha) V_\alpha, \tag{7}$$

where $\{V_\alpha\}_{\alpha=1}^m$ are the quadrature weights associated with the m integration points, and

$$d(\boldsymbol{s}_\alpha, \hat{\boldsymbol{s}}_\alpha) = (d^{\varepsilon 2}(\boldsymbol{\varepsilon}_\alpha, \hat{\boldsymbol{\varepsilon}}_\alpha) + d^{\sigma 2}(\boldsymbol{\sigma}_\alpha, \hat{\boldsymbol{\sigma}}_\alpha))^{1/2}, \tag{8}$$

where

$$d^\varepsilon(\boldsymbol{\varepsilon}_\alpha, \hat{\boldsymbol{\varepsilon}}_\alpha) = \left(1/2(\boldsymbol{\varepsilon}_\alpha - \hat{\boldsymbol{\varepsilon}}_\alpha)^T \boldsymbol{M}_\alpha^\varepsilon (\boldsymbol{\varepsilon}_\alpha - \hat{\boldsymbol{\varepsilon}}_\alpha)\right)^{1/2},$$
$$d^\sigma(\boldsymbol{\sigma}_\alpha, \hat{\boldsymbol{\sigma}}_\alpha) = \left(1/2(\boldsymbol{\sigma}_\alpha - \hat{\boldsymbol{\sigma}}_\alpha)^T \boldsymbol{M}_\alpha^\sigma (\boldsymbol{\sigma}_\alpha - \hat{\boldsymbol{\sigma}}_\alpha)\right)^{1/2}. \tag{9}$$

Here $\boldsymbol{M}_\alpha^\varepsilon \in {}^\circ{}^{q \times q}$ and $\boldsymbol{M}_\alpha^\sigma \in {}^\circ{}^{q \times q}$ are symmetric and positive-definite coefficient matrices for multivariate distance measures, and usually $\boldsymbol{M}_\alpha^\sigma = \left(\boldsymbol{M}_\alpha^\varepsilon\right)^{-1}$. One approach for selecting the coefficient matrices is by computing the covariance of the material data set and using the so-called Mahalanobis distance for multivariate data, as proposed in [54]. Investigating the effect of coefficient matrices is out of the scope of this study. Numerical examples show that by using the proposed locally convex construction scheme with a coefficient matrix representing linear

elasticity, which can be extracted from stress-strain dataset at small strain range, satisfactory data-driven results are achieved.

2.3.1. Global step of data-driven solver

The global step of the data-driven problem (5a) is reformulated as:

$$\min_{u \in C_u, \sigma \in V_\sigma} H(u, \sigma, \hat{\varepsilon}, \hat{\sigma}),$$

$$\text{subject to:} \quad \text{div}\sigma + b = 0 \quad \text{in } \Omega, \tag{10}$$

$$\sigma \cdot n = \bar{t} \quad \text{on } \Gamma_t.$$

This global step searches for the physically admissible state $s = (\varepsilon[u], \sigma) \in C$ closest to a given experimental data $\hat{s} = (\hat{\varepsilon}, \hat{\sigma})$ by means of Lagrange multipliers

$$L_{DD}(u, \sigma, \lambda, \eta) = H(u, \sigma, \hat{\varepsilon}, \hat{\sigma}) + \int_\Omega \lambda \cdot (\text{div}\sigma + b) d\Omega + \int_{\Gamma_t} \eta \cdot (\sigma \cdot n - \bar{t}) d\Gamma, \tag{11}$$

where λ and η are the Lagrange multipliers in proper function spaces. The Euler-Lagrange equations of (11) reveals $\eta = -\lambda$ on Γ_t [61]. Considering the equations (6) - (9), and $\varepsilon = \varepsilon[u]$, the variational form is

$$\delta L_{DD}(u, \sigma, \lambda) = \int_\Omega \left(\delta\varepsilon[u] : M^\varepsilon : (\varepsilon[u] - \hat{\varepsilon}) + \delta\sigma : (M^\sigma : (\sigma - \hat{\sigma})) \right) d\Omega$$

$$- \int_\Omega \delta\sigma : \varepsilon[\lambda] d\Omega + \int_{\Gamma_u} (\delta\sigma \cdot n) \cdot \lambda d\Gamma - \int_\Omega \delta\varepsilon[\lambda] : \sigma d\Omega \tag{12}$$

$$+ \int_\Omega \delta\lambda \cdot b d\Omega + \int_{\Gamma_t} \delta\lambda \cdot \bar{t} d\Gamma,$$

where $\varepsilon[\lambda] = 1/2(\nabla\lambda + \nabla\lambda^T)$. Consequently, we have

$$\int_\Omega \delta\varepsilon[u] : M^\varepsilon : \varepsilon[u] d\Omega = \int_\Omega \delta\varepsilon[u] : M^\varepsilon : \hat{\varepsilon} d\Omega, \tag{13a}$$

$$\int_\Omega \delta\varepsilon[\lambda] : \sigma d\Omega = \int_\Omega \delta\lambda \cdot b d\Omega + \int_{\Gamma_t} \delta\lambda \cdot \bar{t} d\Gamma, \tag{13b}$$

$$\int_\Omega \delta\sigma : (M^\sigma : \sigma - \varepsilon[\lambda]) d\Omega = \int_\Omega \delta\sigma : M^\sigma : \hat{\sigma} d\Omega. \tag{13c}$$

Note that $\lambda = 0$ on Γ_u has been introduced. In this study displacement u, the Lagrange multipliers λ and stress σ are approximated by

11

$$u(x) \approx u^h(x) = \sum_{I=1}^{N} \Psi_I(x) d_I, \tag{14a}$$

$$\lambda(x) \approx \lambda^h(x) = \sum_{I=1}^{N} \Psi_I(x) \Lambda_I, \tag{14b}$$

$$\sigma(x) \approx \sigma^h(x) = \sum_{\alpha=1}^{m} \chi_\alpha(x) \sigma_\alpha, \tag{14c}$$

where N is the number of discretization nodes, m is the number of stress-strain evaluation points at x_α, $\{d_I\}_{I=1}^{N}$ are the nodal displacement vectors, $\{\Lambda_I\}_{I=1}^{N}$ are the nodal Lagrange multiplier vectors, $\chi_\alpha(x)$ is an indicator function such that $\chi_\alpha(x) = 1$ if $x \in \Omega_\alpha$ and $\chi_\alpha(x) = 0$ if $x \notin \Omega_\alpha$, where Ω_α is the subdomain associated to the integration point x_α. Here, we employ $\{\Psi_I\}_{I=1}^{N}$ the reproducing kernel (RK) shape functions [57,58] constructed using the cubic-B splines kernel function and linear basis functions. The introduction of RK approximation is summarized in Appendix B. Stress in (13c) is discretized by a collocation approach in (14c). Thus, the discrete form of Equation (13c) yields

$$\sum_{\alpha=1}^{m} V_\alpha \delta \sigma_\alpha^T \left(M_\alpha^\sigma \sigma_\alpha - \sum_{I=1}^{N} B_{\alpha I} \Lambda_I \right) = \sum_{\alpha=1}^{m} V_\alpha \delta \sigma_\alpha^T M_\alpha^\sigma \hat{\sigma}_\alpha, \tag{15}$$

where $B_{\alpha I} = B_I(x_\alpha)$ is the strain-displacement matrix (the smoothed strain-displacement matrix $\tilde{B}_{\alpha I}$ is used in this study, refer to Appendix B.2), and $M_\alpha^\sigma = M^\sigma(x_\alpha)$. As a result, the matrix equations of (13) result in:

$$\sum_{J=1}^{N} \left(\sum_{\alpha=1}^{m} V_\alpha B_{\alpha I}^T M_\alpha^\varepsilon B_{\alpha J} \right) d_J = \sum_{\alpha=1}^{m} V_\alpha B_{\alpha I}^T M_\alpha^\varepsilon \hat{\varepsilon}_\alpha, \quad I = 1,...,N, \tag{16a}$$

$$\sum_{\alpha=1}^{m} V_\alpha B_{\alpha I}^T \sigma_\alpha = f_I, \quad I = 1,...,N, \tag{16b}$$

$$M_\alpha^\sigma \sigma_\alpha - \sum_{I=1}^{N} B_{\alpha I} \Lambda_I = M_\alpha^\sigma \hat{\sigma}_\alpha, \quad \alpha = 1,...,m. \tag{16c}$$

where $\{V_\alpha\}_{\alpha=1}^{m}$ are the quadrature weights as defined in (7), and $\{f_I\}_{I=1}^{N}$ are the nodal force vectors associated with the employed RK approximation of body force b and surface traction $\bar{t}$.

It can be seen that $\{\boldsymbol{d}_I\}_{I=1}^{N}$ are solved from (16a) directly, and $\{\boldsymbol{\Lambda}_I\}_{I=1}^{N}$ represent the displacement adjustment related to the difference between the computational stress and the given stress data $\{\hat{\boldsymbol{\sigma}}_\alpha\}_{\alpha=1}^{m}$, as shown in Equation (16c). Plugging (16c) into (16b) it yields

$$\sum_{J=1}^{N}\left(\sum_{\alpha=1}^{m}V_\alpha\boldsymbol{B}_{\alpha I}^{T}\boldsymbol{M}_\alpha^{\sigma-1}\boldsymbol{B}_{\alpha J}\right)\boldsymbol{\Lambda}_J = \boldsymbol{f}_I - \sum_{\alpha=1}^{m}V_\alpha\boldsymbol{B}_{\alpha I}^{T}\hat{\boldsymbol{\sigma}}_\alpha,\ I=1,...,N, \tag{17a}$$

where $\{\boldsymbol{\Lambda}_I\}_{I=1}^{N}$ can be solved readily. Then, the computational stress $\{\boldsymbol{\sigma}_\alpha\}_{\alpha=1}^{m}$ are obtained by the following equation

$$\boldsymbol{\sigma}_\alpha = \hat{\boldsymbol{\sigma}}_\alpha + \boldsymbol{M}_\alpha^{\sigma-1}\sum_{I=1}^{N}\boldsymbol{B}_{\alpha I}\boldsymbol{\Lambda}_I,\ \alpha=1,...,m. \tag{17b}$$

In summary, Equations (16a), (17a) and (17b) constitute the global step of the data-driven solver. In each global step, the displacement vector $\{\boldsymbol{d}_I\}_{I=1}^{N}$ is obtained from strain data $\{\hat{\boldsymbol{\varepsilon}}_\alpha\}_{\alpha=1}^{m}$ by complying with compatibility, while the displacement adjustment $\{\boldsymbol{\Lambda}_I\}_{I=1}^{N}$ is driven by the force residuals between the external force and the internal force computed by the experimental stress data $\{\hat{\boldsymbol{\sigma}}_\alpha\}_{\alpha=1}^{m}$, as shown in (17a).

In this study, we propose to use a stabilized conforming nodal integration (SCNI) [59] for the integration of the weak form (13) due to its nodal representation nature of both state and field variables at nodal points. The brief summary about SCNI can be found in Appendix B. In this approach, the continuum domain is partitioned by a Voronoi diagram (see **Fig. 15**), and both the state variables $\{(\boldsymbol{\varepsilon}_\alpha,\boldsymbol{\sigma}_\alpha)\}_{\alpha=1}^{N} = \{(\boldsymbol{\varepsilon}(\boldsymbol{x}_\alpha),\boldsymbol{\sigma}(\boldsymbol{x}_\alpha))\}_{\alpha=1}^{N}$ and the nodal displacement vectors $\{\boldsymbol{u}_I\}_{I=1}^{N} = \{\boldsymbol{u}(\boldsymbol{x}_I)\}_{I=1}^{N}$ are computed at the set of nodes located at $\{\boldsymbol{x}_\alpha\}_{\alpha=1}^{N}$, i.e. $m=N$. This approach minimizes the number of integration points where the stress and strain experimental data are searched in the local step (5b), allowing an enhanced effectiveness in the learning solver. The introduction of a smooth reproducing kernel (RK) shape function in the displacement approximation in (14a), such as the employment of a cubic B-spline in the RK approximation in Equations (B.1) - (B.7) in Appendix B, yields a C^1 continuous strain-displacement matrix $\boldsymbol{B}_I(\boldsymbol{x}_\alpha)$, and consequently a smooth tangent matrices in the displacement adjustment and stress update equations in (17a) and (17b), respectively. This smooth solution space of the physical

13

manifold is made consistent with the continuous and convex solution space of the regularized LCDD learning solver to be introduced in Section 3.

2.3.2. Local step of data-driven solver

In this approach, the experimental data $\hat{s} = (\hat{\varepsilon}, \hat{\sigma})$ as used for the global step solution (see (10) or (13)) is searched by the local step in (5b) during each local-global iteration. Considering the functional H as shown in (7), the local step of (5b) can be decomposed into m local minimization problems: find *the optimal local data* $\hat{s}_\alpha^* = (\hat{\varepsilon}_\alpha^*, \hat{\sigma}_\alpha^*)$, such that the distance to a given local state $s_\alpha = (\varepsilon_\alpha, \sigma_\alpha)$ is minimized, i.e.

$$\hat{s}_\alpha^* = \arg\min_{\hat{s}_\alpha \in E} d^2(s_\alpha, \hat{s}_\alpha) = \arg\min_{\hat{s}_\alpha \in E} d^{\varepsilon 2}(\varepsilon_\alpha, \hat{\varepsilon}_\alpha) + d^{\sigma 2}(\sigma_\alpha, \hat{\sigma}_\alpha), \tag{18}$$

for $\alpha = 1, ..., m$.

2.3.3. Standard data-driven solver

The procedures for solving the data-driven problem are summarized below. Given a set of local data solutions $\{\hat{s}_\alpha^{(v)}\}_{\alpha=1}^m \in E_{em}$ (or $\hat{s}_\alpha^{(v)} \in E, \alpha = 1, ..., m$ for homogeneous material) of the v-th iteration, the following global step and local step are iterated until convergence:

1) *Global Step*. Input: $\{\hat{s}_\alpha^{(v)}\}_{\alpha=1}^m \rightarrow$ Output: $\{s_\alpha^{(v)}\}_{\alpha=1}^m$

    1.1 Solve Equations (16a) for $\{d_I^{(v)}\}_{I=1}^N$ and (17a) for $\{\Lambda_I^{(v)}\}_{I=1}^N$.

    1.2 Update computational states for $\{s_\alpha^{(v)}\}_{\alpha=1}^m = \{(\varepsilon_\alpha^{(v)}, \sigma_\alpha^{(v)})\}_{\alpha=1}^m$ via

    $$\varepsilon_\alpha^{(v)} = \sum_{I=1}^N B_{\alpha I} d_I^{(v)} \text{ and } \sigma_\alpha^{(v)} = \hat{\sigma}_\alpha^{(v)} + M_\alpha^{\sigma-1} \sum_{I=1}^N B_{\alpha I} \Lambda_I^{(v)} \text{ in (17b)}$$

2) *Local Step*. Input: $\{s_\alpha^{(v)}\}_{\alpha=1}^m \rightarrow$ Output: $\{\hat{s}_\alpha^{(v+1)}\}_{\alpha=1}^m$

    for $\alpha = 1, ..., m$, solve Equation (18) for $\hat{s}_\alpha^{(v+1)}$.

Remark 2.2. It has been observed that distance minimizing data-driven (DMDD) computing solver [41] with distance measure in (7) - (9) is sensitive to data noise and outliers [42,54]

because the local minimization stage (18) only searches for the nearest data point from the given experimental data set regardless of any latent data structure. The data-driven solution could be strongly influenced by the outliers locating near to the physical manifold C but do not conform to the hidden material data pattern (or the latent statistical model) of E. Without the knowledge of the underlying data manifold, it requires a large amount of data to achieve sufficiently accurate predictions which is costly [20,31].

3. Local convexity-preserving data-driven approach

In this section, we introduce the LCDD approach, which introduces the locally convex reconstruction technique inspired by manifold learning strategies [21,62,63]. Our aim in this work is to develop a computationally feasible data-driven predictive modeling framework to enhance accuracy and robustness against noise and outliers in the experimental data set by constructing the local manifold with the desired smoothness and convexity.

For ease of exposition, we define a weighted vector norm "$\|\cdot\|_M$" based on (8) as follows

$$\| s_\alpha \|^2_M = s_\alpha^T \bar{M}_\alpha s_\alpha \equiv \| \bar{M}_\alpha^{1/2} s_\alpha \|^2, \tag{19}$$

where $s_\alpha = [\varepsilon_\alpha^T \, \sigma_\alpha^T]^T \in \degree^{2q}$, $\bar{M}_\alpha = diag([M_\alpha^\varepsilon, M_\alpha^\sigma])$, and $\bar{M}_\alpha^{1/2}$ can be determined by the singular value decomposition of $\bar{M}_\alpha$. For a given s_α, the local step (18) is rewritten as

$$\hat{s}_\alpha^* = \arg\min_{\hat{s}_\alpha \in E} \| s_\alpha - \hat{s}_\alpha \|^2_M, \tag{20}$$

for $\alpha = 1,...,m$.

3.1. Locally convex construction

It has been shown [21,62,64–66] that naturally occurring data usually reside on a lower dimensional submanifold which is embedded in the high-dimensional ambient space, as shown in **Fig. 2**. In this study, inspired by locally linear embedding (LLE) approach [62], we assume there exists an underlying manifold of low dimensionality corresponding to the raw experimental data set, i.e. $E = \{\hat{s}^i, i = 1,..., p\}$ where p is the number of data points, that is locally linear and smooth varying. Therefore, a data point $\hat{s}^i \in E$ can be linearly reconstructed from its neighbors in the data set, i.e.

$$\hat{s}^i \approx \hat{s}^i_{recon} = \sum_{j \in N_k(\hat{s}^i)} w_{ij} \hat{s}^j, \tag{21}$$

where $\hat{s}^i_{recon}$ is the reconstruction of $\hat{s}^i$, $N_k(\hat{s}^i)$ is the set of the k nearest neighbor (k-NN) data points to $\hat{s}^i$ in E, and w_{ij} are the unknown coefficients. In LLE, the optimal reconstruction weights w_{ij}^* can be obtained by solving the following problem:

16

$$\{w_{ij}^*\}_{i,j=1,\ldots,p} = \arg\min \sum_{i=1}^{p} \| \hat{\boldsymbol{s}}^i - \sum_{j=1,j\neq i}^{p} w_{ij}\hat{\boldsymbol{s}}^j \|^2$$

$$\text{subject to: } \sum_{j=1}^{p} w_{ij} = 1, \ i = 1,\ldots,p \tag{22}$$

$$w_{ij} = 0 \ \text{if} \ j \notin \mathrm{N}_k(\hat{\boldsymbol{s}}^i)$$

Note that $w_{ij} = 0$ when $i = j$. The data reconstruction procedures in (21) and (22) provide the projection of $\hat{\boldsymbol{s}}^i$, i.e. $\hat{\boldsymbol{s}}^i_{recon}$, onto the subspace spanned by $\{\hat{\boldsymbol{s}}^j\}_{j\in \mathrm{N}_k(\hat{\boldsymbol{s}}^i)}$ with respect to the norm.
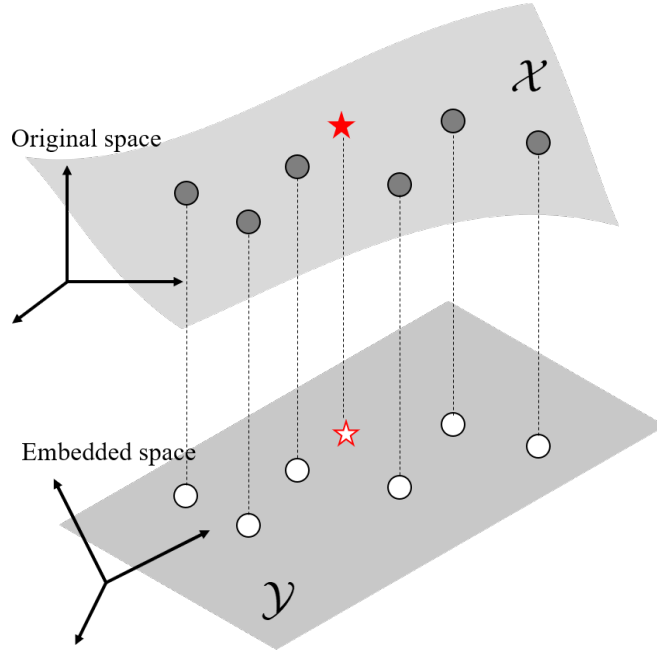

Fig. 2. Schematic of a manifold embedded in the original space and the associated low-dimensional embedding, where the training samples and the new sample are denoted by gray circles and red star, respectively.

Different from the standard LLE, the search for solution data point in data-driven computing is constrained by the physical manifold associated with (1) and (2). But considering that the data-driven algorithm in (5) performs a fixed point iteration on the experimental data points that are closest to the physical manifold, the locally linear reconstruction remains suitable for this scenario. In this sense, it is similar to the out-of-sample extension problem [63,67,68] where it is desirable to add new projected data points (new samples) to a previously learnt low-

dimensional embedding, as shown in **Fig. 2**. In addition, from physical perspective, the data-driven solution constrained by the physics laws need to be close enough to the graph of the experimental data with underlying constitutive data structure. Thus, we need to prevent the reconstructed data point by (21) from projecting to a point that is far away from the underlying material data structure on the embedded subspace. To this end, we propose a local manifold learning algorithm to reconstruct the given local state on the locally convex manifold of the experimental data set.

Given a local state s_α, the most representative k nearest neighbor (k-NN) points in E are first identified using the same metric induced by the given norm "$\|\cdot\|_M$", and collected as $\{\hat{s}^i\}_{i \in N_k(s_\alpha)} \subset E$, in which the indices for the nearest neighbors of s_α are stored in a set $N_k(s_\alpha)$. Then we project the local state onto the convex hull of $\{\hat{s}^i\}_{i \in N_k(s_\alpha)}$ associated to s_α, which is defined as:

$$E(s_\alpha) = \mathrm{Conv}(\{\hat{s}^i\}_{i \in N_k(s_\alpha)}) = \left\{ \sum_{i \in N_k(s_\alpha)} w_i \hat{s}^i \,\middle|\, \sum_{i \in N_k(s_\alpha)} w_i = 1, \text{ and } w_i \geq 0, \forall i \in N_k(s_\alpha) \right\}, \quad (23)$$

or concisely denoted as E_α. Accordingly, the optimal reconstruction coefficients are given by solving the following minimization problem:

$$w_\alpha^* = \arg\min_{w \in \circ^k} \| s_\alpha - \sum_{i \in N_k(s_\alpha)} w_i \hat{s}^i \|_M^2$$

$$\text{subject to: } \sum_{i \in N_k(s_\alpha)} w_i = 1, \qquad\qquad (24a)$$

$$w_i \geq 0, \ \forall i \in N_k(s_\alpha),$$

where $w \in \circ^k$ denotes the vector consisting of the weights $\{w_i\}_{i \in N_k(s_\alpha)}$ corresponding to the k selected neighbor points, and w_α^* with the subscript α denotes the optimal weights corresponding to the given local state s_α. The reconstruction $\hat{s}_\alpha^*$ can be retrieved by using the linear combination of $\{\hat{s}^i\}_{i \in N_k(s_\alpha)}$ with the computed weight vector w_α^* as follows

$$\hat{s}_\alpha^* = \sum_{j \in N_k(s_\alpha)} w_j^* \hat{s}^j = \hat{S}_\alpha w_\alpha^*, \qquad\qquad (24b)$$

where $\hat{\boldsymbol{S}}_\alpha \in \circ^{2q \times k}$ is the matrix composed of the k-NN data points $\{\hat{\boldsymbol{s}}^i\}_{i \in \mathrm{N}_k(\boldsymbol{s}_\alpha)}$. This approach in (24) is called *locally convex construction*. Compared to the equation (22), the main differences in (24a) are: 1. a new data point $\boldsymbol{s}_\alpha$ obtained from the physical solver, instead of other points in the experimental data set, is used for local construction; 2. a weighted vector norm "$\| \cdot \|_M$" representing energy is adopted for distance measure.

Based on the idea of locally convex construction, the local step of data-driven computing in (20) is modified as: Given the data-set neighbors $\{\hat{\boldsymbol{s}}^i\}_{i \in \mathrm{N}_k(\boldsymbol{s}_\alpha)} \subset \mathrm{E}$ for $\boldsymbol{s}_\alpha$, solve $\hat{\boldsymbol{s}}_\alpha^*$ such that

$$\hat{\boldsymbol{s}}_\alpha^* = \arg \min_{\hat{s}_\alpha \in \mathrm{E}_\alpha} \| \boldsymbol{s}_\alpha - \hat{\boldsymbol{s}}_\alpha \|_M^2, \tag{25}$$

for $\alpha = 1, ..., m$. By comparing (20) and (25), we can observe that the space E used in the standard data-driven scheme [41,54] is now replaced by the associated convex hull E_α that is locally reconstructed around the input $\boldsymbol{s}_\alpha$ by learning techniques, allowing to capture the local material manifold. Consequently, the reconstruction data (i.e., the optimal local data) $\hat{\boldsymbol{s}}_\alpha^*$ is sought from the set E_α with convexity and smoothness. With the definition in (23), the solution of the minimization problem (25) is obtained by solving (24).

Remark 3.1. Equation (24a) is a *constrained regression* or *constrained least-squares* problem under a *invariance* constraint and a *non-negative* constraint. The invariance constraint imposes the partition of unity on the weight array $\boldsymbol{w}$, i.e. $\boldsymbol{I}^T \boldsymbol{w} = 1$, where $\boldsymbol{I} = [1, 1, ..., 1]^T \in \circ^k$. It ensures the invariance of the reconstruction weights $\boldsymbol{w}_\alpha^*$ to rotations, rescaling, and translations of the same k-NN data points, and thus, the weights characterize geometric properties independent of a particular frame of reference [62,63]. It also guarantees the linear approximation property such that $\hat{\boldsymbol{s}}_\alpha^*$ is in the subspace $\mathrm{span}(\{\hat{\boldsymbol{s}}^i\}_{i \in N_k(\boldsymbol{s}_\alpha)})$. When we further consider the non-negative constraint, the approximation $\hat{\boldsymbol{s}}_\alpha^*$ is restricted to the convex hull $\mathrm{E}(\boldsymbol{s}_\alpha)$ (see **Fig. 3**). The imposed convexity and locality yields enhanced robustness of linear regressions to outliers [63,69], and reduces numerical instability across different clusters of neighbor points during data-driven iterations. Moreover, it is well known that the non-negative constraint naturally imposes sparseness on the coefficient solution $\boldsymbol{w}_\alpha^*$. Lastly, by specifying the number of

k-NN points, it provides an opportunity to incorporate a priori knowledge about the experimental data structure and therefore, enhance the robustness of data learning [63].
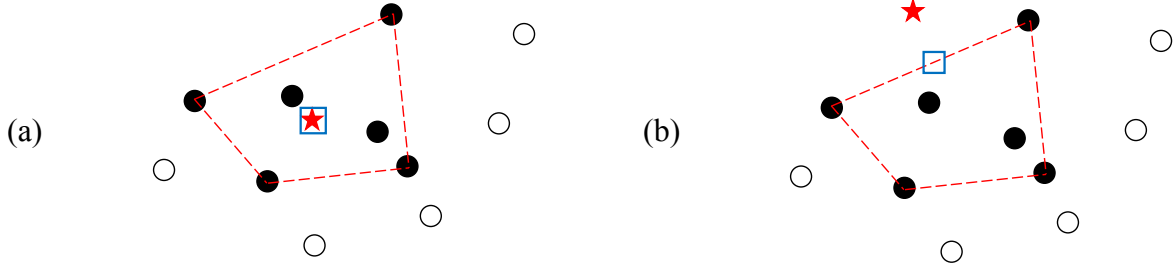

Fig. 3. Sketch of the projection $\hat{s}_\alpha^*$ (the blue square) on a convex hull E_α (the region is depicted by red dashed lines) of k-NN points (the solid circles in black) when a local state s_α (the red star) locates (a) inside and (b) outside E_α. Neighbor points of $k = 6$ are used for demonstration.

Essentially, the modified local step of data-driven problem in (25) can be interpreted as a process of seeking the data approximation based on the previously learnt low-dimensional manifold E_α associated with the given local state. From a geometrical point of view, it searches the projection (i.e. the optimal local material data $\hat{s}_\alpha^* \in E_\alpha$) in the associated convex set E_α. If s_α locates inside E_α, the projection is represented by s_α itself (**Fig. 3**a). Otherwise, the local state is optimally projected on the convex hull E_α and the projection point is considered as the best representative on the constitutive manifold (**Fig. 3**b).

3.2. Solving non-negative least squares

In this section, a computationally feasible algorithm is developed to solve the local step minimization problem in (24a) by relating it to the *non-negative least squares* (NNLS) problem that has been well established. The NNLS problem is reviewed in Append A.

However, to solve the minimization problem (24a) under the NNLS framework, see (A.1), the partition of unity constraint in (24a) needs to be properly handled. In this end, we propose to employ the quadratic penalty method [70] to penalize the residuals of the partition of unity constraint in the auxiliary objective, and the modified minimization problem becomes

20

$$w_\alpha^* = \arg\min_{w \in \circ^k} \| s_\alpha - \hat{S}_\alpha w \|_M^2 + \xi(I^T w - 1)^2,$$
$$\text{subject to: } w_i \geq 0, \ i = 1,...,k, \tag{26}$$

where $\xi > 0$ is a regularized coefficient to impose the associated constraint.

Note that to conform with the Euclidean metric used in the standard NNLS solver (A.1), the local state s_α can be easily rescaled to $z_\alpha = \bar{M}_\alpha^{1/2} s_\alpha$ (similarly, $\hat{Z}_\alpha = \bar{M}_\alpha^{1/2} \hat{S}_\alpha$) by using the relation given in (19). As a result, the minimization problem (26) can be recast into a NNLS form as shown by augmenting the vector s_α and the matrix $\hat{Z}_\alpha$ with additional components as follows

$$w_\alpha^* = \arg\min_{w \in \circ^k} \| z_\alpha^{\text{aug}} - \hat{Z}_\alpha^{\text{aug}} w \|^2,$$
$$\text{subject to: } w_i \geq 0, \ i = 1,...,k, \tag{27}$$

where

$$\hat{Z}_\alpha^{\text{aug}} := \begin{bmatrix} \hat{Z}_\alpha \\ \sqrt{\xi} I^T \end{bmatrix} \in \circ^{(2q+1) \times k}, \quad z_\alpha^{\text{aug}} := \begin{bmatrix} z_\alpha \\ \sqrt{\xi} \end{bmatrix} \in \circ^{2q+1}. \tag{28}$$

To properly impose the penalty term, we set $\xi = \bar{\bar{\xi}} \, \text{tr}(\hat{Z}_\alpha^T \hat{Z}_\alpha)/k$, where $\bar{\bar{\xi}}$ is a large parameter (usually set as $10^4 - 10^6$). With the weight solution w_α^* solved by NNLS algorithm, the reconstruction $\hat{s}_\alpha^*$ can be obtained via (24b).

It is possible that constrained least squares in (27) could suffer from numerical instability due to rank deficiency when the number of neighbors is larger than the rank of the neighborhood, i.e. $k > \text{rank}(\hat{Z}_\alpha^{\text{aug}})$. As has been well studied in machine learning field [5], a further regularization can be introduced to the NNLS problem. In this study, a commonly used ridge regression [71], or called Tikhonov regularization, is applied to address the ill-posed issues, and the NNLS problem (27) is modified as

$$w_\alpha^* = \arg\min_{w \in \circ^k} \| \hat{Z}_\alpha^{\text{aug}} w - z_\alpha^{\text{aug}} \|^2 + \mu \| w \|^2,$$
$$\text{subject to: } w_i \geq 0, \ i = 1,...,k, \tag{29}$$

where the regularized coefficient is

$$\mu = \bar{\mu}\,\mathrm{tr}(\hat{\pmb{Z}}_\alpha^T \hat{\pmb{Z}}_\alpha)\big/k, \tag{30}$$

Here $\bar{\mu}$ is a small constant (set as 10^{-4} by default) such that the regularization has minor effect on the solution $\pmb{w}_\alpha^*$ and the reconstruction $\hat{\pmb{s}}_\alpha^*$. It is also shown that [5] this regularization imposes certain smoothness on the solution and guarantees a unique solution.

Remark 3.2. As discussed in [63], the size range of k-NN points depends on various features of the data, such as the manifold geometry and the sampling density. In principle, k should be greater than the underlying manifold dimensionality of the material data set E in order to explore the data structural and prevent overwhelming influence from outliers/noise. Meanwhile, the resultant neighborhoods should be localized enough to ensure the validity of locally linear approximation.

3.3. Local convexity-preserving data-driven solver

A simple algorithm for the proposed LCDD solver is shown as follows: Given a convergence tolerance TOL and the material database E , then

1. Initialize $\hat{\pmb{s}}_\alpha^{*(0)} = [\hat{\pmb{\varepsilon}}_\alpha^{*(0)T}\ \hat{\pmb{\sigma}}_\alpha^{*(0)T}]^T$, $\alpha = 1,...,m$ randomly, and $v = 0$.

2. While $\displaystyle\max_{\alpha=1,...,m} \| \hat{\pmb{s}}_\alpha^{*(v)} - \hat{\pmb{s}}_\alpha^{*(v-1)} \|_M > TOL$

   a. Solve equations (16), and output $\{\pmb{s}_\alpha^{(v)}\}_{\alpha=1}^m$

   b. Construct k-NN neighborhood $\mathrm{N}_k(\pmb{s}_\alpha^{(v)})$ and $\hat{\pmb{S}}_\alpha$ for each local state $\pmb{s}_\alpha^{(v)}$.

   c. Solve NNLS (27) (or (29)) by **Algorithm 1**, and use $\pmb{w}_\alpha^*$ to output $\hat{\pmb{s}}_\alpha^{*(v+1)}$ via (24b).

   d. Update: $v \leftarrow v+1$

3. Solution is $\pmb{s}_\alpha = [\pmb{\varepsilon}_\alpha^T\ \pmb{\sigma}_\alpha^T]^T \leftarrow \pmb{s}_\alpha^{(v)} = [\pmb{\varepsilon}_\alpha^{(v)T}\ \pmb{\sigma}_\alpha^{(v)T}]^T$, $\alpha = 1,...,m$.

It has been shown that the Lawson-Hanson method [72,73] (**Algorithm 1** in Appendix A) used for solving NNLS converges in a finite number of iterations less than the size of the output coefficient vector, which is the size of k-NN in LCDD. In addition, considering the small size of the local matrix $\hat{\pmb{S}}_\alpha \in \circ^{\ 2q\times k}$, $k, q = \min(N,m)$, where N and m are the numbers of discretization nodes and integration points, respectively, the additional computational cost in

solving the NNLS problem in (27) or (29) is negligible compared to solving the linear system (16
).

4. Numerical Examples: Truss Structures

In this section, the accuracy, convergence and robustness properties of the proposed LCDD approach are examined. Synthetic data sets are employed to verify the performance of the proposed method. The standard DMDD approach without considering any local data structure is also provided for comparison. For simplicity, we consider homogeneous material in the following numerical examples in Section 4 and Section 5. That is, the same material database $E = \{\hat{s}^i = (\hat{\varepsilon}^i, \hat{\sigma}^i), i = 1, ..., p\}$, where p is the number of data points, is introduced for all material points.

In this section, the material behavior of the α-th bar member is characterized by a simple uniaxial strain ε_α and uniaxial stress σ_α relationship. As such, the local state vector is defined as $s_\alpha = [\varepsilon_\alpha \ \sigma_\alpha]^T \in \degree^2$, and the associated norm to measure the distances of local states is given as

$$\| s_\alpha \|_M = \left(\frac{1}{2} M \varepsilon_\alpha^2 + \frac{1}{2} M^{-1} \sigma_\alpha^2 \right)^{1/2}, \tag{31}$$

where M is a positive constant analogous to the Young's modulus of the reference material.

In the following examples, k is the number of the k-NN used in the local step of the LCDD solver and $\bar{\mu}$ is the regularization coefficient in (30) with a default value 10^{-4}.
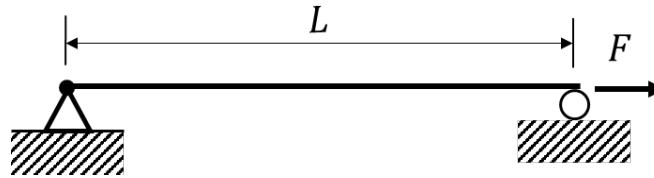
4.1. Example I: One-dimensional truss

Fig. 4. One-bar truss structure with the cross-section area $A = 200 \text{ cm}^2$ subjected to a uniaxial load $F = 10 \text{ kN}$.

This example examines data-driven computing in a single truss member ($m = 1$) when dealing with irregular material data that exhibits noise and outliers. A truss member with the cross-section area $A = 200 \text{ cm}^2$ is subjected to an axial load of $F = 10 \text{ kN}$ as shown in **Fig. 4**.

24

4.1.1. Data set with random noise

To examine the performance of LCDD when dealing with noisy material data, two material databases (denoted by the circle points) with different levels of Gaussian random noise (i.e. $\chi = 0.05$ and $\chi = 0.15$ in (32)) are considered, as shown in **Fig. 5** and **Fig. 6**. The data sets are given by a set of noiseless strain and stress points, i.e. $\{\bar{\varepsilon}^i, \bar{\sigma}^i\}_{i=1}^p$, superimposed with a Gaussian noise as follows:

$$
\begin{aligned}
\hat{\varepsilon}^i &= \bar{\varepsilon}^i + \mathrm{N}\,(0, \chi\bar{\varepsilon}_{\max}),\ i = 1,...,p \\
\hat{\sigma}^i &= \bar{\sigma}^i + \mathrm{N}\,(0, \chi M\bar{\varepsilon}_{\max}),\ i = 1,...,p
\end{aligned}
\tag{32}
$$

where the coefficient constant is $M = 100\ \mathrm{MPa}$, and χ is the given random level. In (32), the noiseless strain data $\{\bar{\varepsilon}^i\}_{i=1}^p$ is generated by a uniform distribution within a given strain range $[-\bar{\varepsilon}_{\max}, \bar{\varepsilon}_{\max}]$ with $\bar{\varepsilon}_{\max} = 0.01$, the noiseless stress data $\{\bar{\sigma}^i\}_{i=1}^p$ is defined as $\bar{\sigma}^i = M\bar{\varepsilon}^i$, and the Gaussian noise $\mathrm{N}\,(0, \chi\bar{\varepsilon}_{\max})$ and $\mathrm{N}\,(0, \chi M\bar{\varepsilon}_{\max})$ of strain and stress represent the normal distribution with zero mean and standard deviations of $\chi\bar{\varepsilon}_{\max}$ and $\chi M\bar{\varepsilon}_{\max}$, respectively.

The external force is incrementally loaded via 5 equal steps at the one-bar truss structure in **Fig. 4**, and thus 5 corresponding incremental data-driven solutions (depicted by the asterisk points) are shown in **Fig. 5** and **Fig. 6**. The associated optimal local data to the minimization problem in (25), i.e. $\hat{s}_\alpha^* = [\hat{\varepsilon}_\alpha^*\ \hat{\sigma}_\alpha^*]^T$, at these 5 loading steps are denoted by the triangle points. As shown in **Fig. 5**, both the data-driven solutions of DMDD and LCDD can finally converge to a state point at the physical (equilibrium) manifold of $\sigma = 0.5\ \mathrm{Mpa}$ and stay close to the given data set. Although the mild randomness is considered, DMDD yields a less desirable result compared to the result obtained by LCDD with the number of neighbor points $k = 6$. The data-driven solution from LCDD converges at the intersection of the physical manifold and the conjectural material graph, which is an ideal solution as discussed in Section 2. Note that by using the proposed locally convex reconstruction (see (23) or (24)), the local solver in (25) allows to attain the optimal local data from the reconstructed local convex hull instead of a direct search in the raw experimental points.

When the given material database presents stronger randomness, as shown in **Fig. 6**, DMDD performs poorly as it prematurely converges to a suboptimal solution far outside the data set, and its incremental solutions attain some local minimum points (**Fig. 6**a), implying susceptibility to data noise. In contrast, LCDD yields more desirable results (see **Fig. 6**b), where the incremental data-driven solutions move consistently to the intersection of the equilibrium manifold and the underlying material submanifold.

(a)   (b)
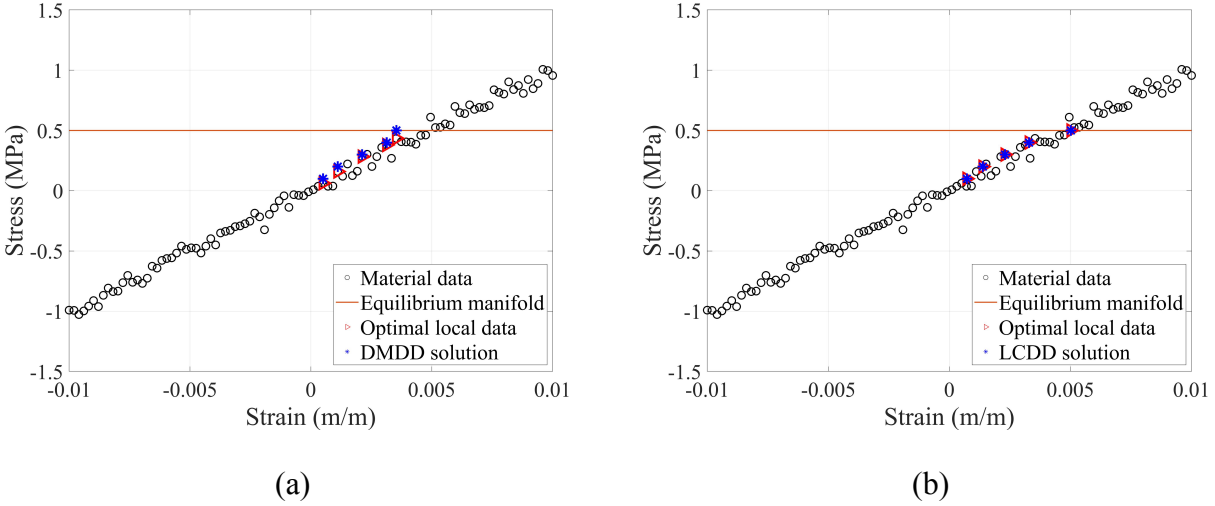
Fig. 5. Comparison of the (a) DMDD and (b) LCDD solvers for the one-bar truss structure using a material database with mild Gaussian random noise $\chi = 0.05$. The database contains $p = 100$ stress-strain data points. The number of neighbor points used in LCDD is $k = 6$.

Fig. 6. Comparison of the (a) DMDD and (b) LCDD solvers for the one-bar truss structure using a material database with strong Gaussian random noise $\chi = 0.15$. The database contains $p = 100$ stress-strain data points. The number of neighbor points used in LCDD is $k = 6$.
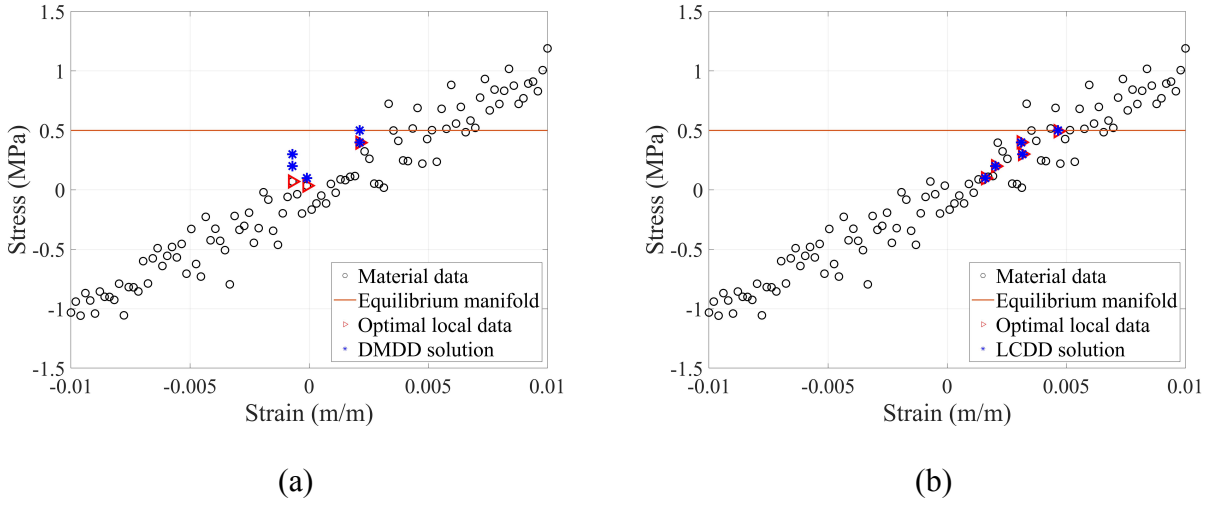
The comparison of the results in **Fig. 5** and **Fig. 6** suggests that LCDD is a robust scheme against large noise and local minimum wells, yielding similar pattern of convergence with material databases of different random levels, whereas the iterative solution path of DMDD is sensitive to the level and distribution of noise. It is also observed that the optimal data solutions $(\hat{\varepsilon}_\alpha^*, \hat{\sigma}_\alpha^*)$ obtained from LCDD usually coincide with the data-driven solutions $(\varepsilon_\alpha, \sigma_\alpha)$ even though no experiment data in E is exactly at those locations. This study indicates the advantage of LCDD in forming an implicit local material graph (via the convex hull) for searching the optimal data points. This unique feature allows LCDD to capture the local data structure, providing not only robustness against noise due to clustering analysis, but also the reproducibility to a locally linear manifold if the data is well sampled, which will be further discussed in the following examples.

4.1.2. Data set with outliers

The robustness of the proposed LCDD solver is further demonstrated by consider a dataset with outliers. Let a material data set represent an underlying linear graph, i.e.

27

$F(\hat{\varepsilon}^i, \hat{\sigma}^i) = \hat{\sigma}^i - M\hat{\varepsilon}^i = 0$ with an outlier positioned near to the linear manifold, as shown in **Fig. 7**. The data-driven solutions and the associated local projected data at each incremental step are plotted in **Fig. 7**. The results show that the DMDD solutions are misled by the presence of the outlier (**Fig. 7**a), whereas the LCDD solutions successfully converge to a reasonable location at the material manifold (**Fig. 7**b). In addition, to verify the effect of the locally convex construction parameters on the performance of the LCDD solver, two exemplary results using a different k-NN numbers ($k = 12$) and a different regularization coefficient ($\bar{\mu} = 10^{-2}$) are provided in **Fig. 7**c and **Fig. 7**d, respectively. As shown in **Fig. 7**c, while using more neighbor points in LCDD tends to increase the probability to involve the outlier for local manifold construction and causes the intermediate results during simulation more influenced by the outlier, it yields a final data-driven solution consistent to the one using $k = 6$. This robustness is achieved from the clustering analysis based on the reconstruction of multiple data points that prevents the dominance of outliers. The effect of the regularization coefficient $\bar{\mu}$ on LCDD is also studied, see **Fig. 7**d. It shows that with larger $\bar{\mu}$ the optimal local data solution sought from the associated convex hull favors the region with higher data density because the reconstructed weights of the outliers have been penalized (refer to (29)). In this study, $\bar{\mu} = 10^{-4}$ is adopted as the default setting. There are other regularization methods and robust penalty functions that allow further suppressing the influence of noise or outliers, e.g. the Huber penalty function. The interested readers are encouraged to consult the reference [5].
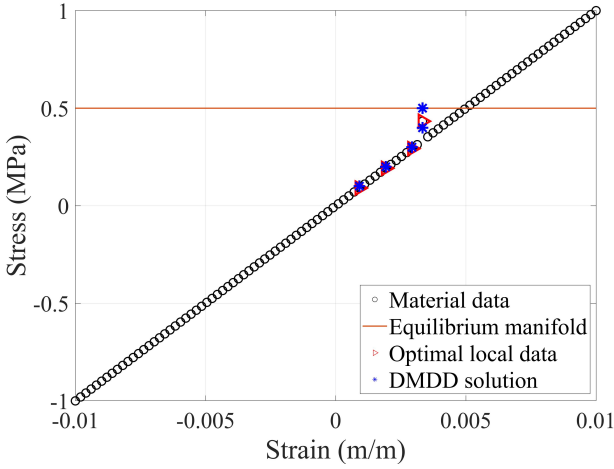
(a) DMDD
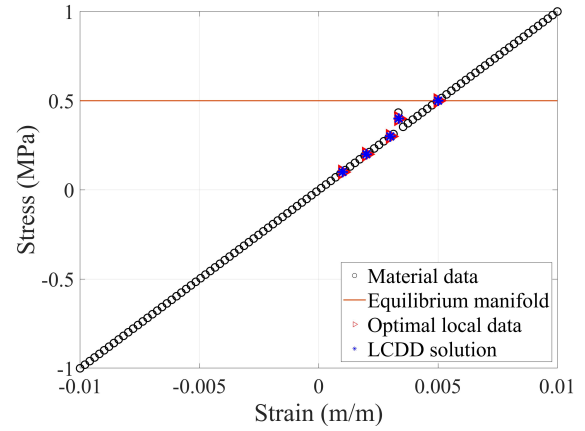
(b) LCDD: $k = 6$, $\bar{\mu} = 10^{-4}$
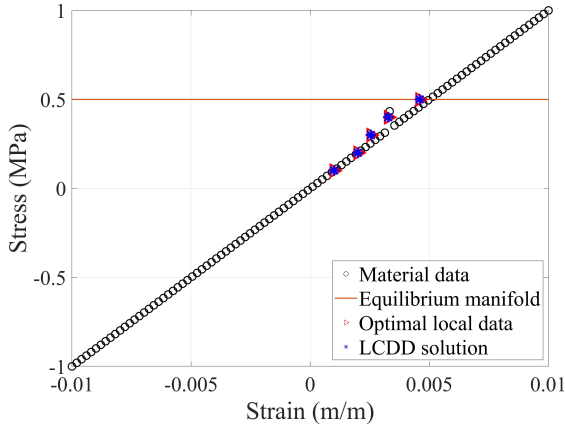
(c) LCDD: $k = 12$, $\bar{\mu} = 10^{-4}$
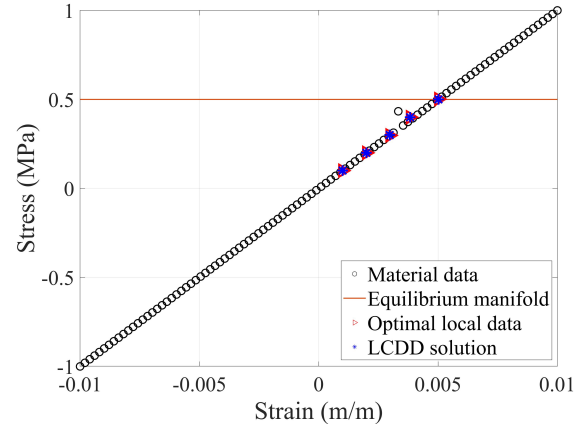
(d) LCDD: $k = 6$, $\bar{\mu} = 10^{-2}$

Fig. 7. Comparison of the data-driven solvers, DMDD and LCDD, for the one-bar truss structure using a material database with an outlier. The database contains $p = 100$ stress-strain data points. Different numbers of neighbor points k and regularization coefficient values $\bar{\mu}$ are used in LCDD.

4.1.3. Data set characterizing a nonlinearly elastic material

To reveal another pathology resulting from the discrete nature of data when using distance-minimizing approaches, consider a "nonlinear" database generated by a sigmoid

function, as shown in **Fig. 8**. In the phase space, the plot of data points transits from a nearly linear stage of slope with $\sigma^i/\varepsilon^i = 100\text{Mpa}$ to a plateau at $\sigma = 0.51\,\text{Mpa}$, which can be viewed as approximating a uniaxial perfect plasticity behavior.

Fig. 8a shows that although no noise is presented in this database, DMDD converges to a suboptimal solution before approaching the flat plateau, indicating the limitation of using discrete data space in the local solver (see (20)). This is because once the direction of data projection (linearly scaled with M in this example) during data-driven iterations is nearly normal to the underlying material graph, the resulting displacement increment driven by the force residual (see (17a)) is too small to move the computational stresses and strains toward other data points in E that are closer to the physical manifold. As a result, the data-driven scheme converges at an undesirable solution. This issue is attributed to the non-continuous nature of discrete data, resulting in the susceptibility of DMDD to the selection of measure coefficient M, the associated metric norm used to measure distance in the phase space, and the density and the underlying structure of data [54,60].

On the other hand, LCDD converges to a better solution (see **Fig. 8**b) at which the physical and material manifolds intersect. This is because when using the LCDD solver the inherited locally convex approximation represents a smooth constitutive (material) submanifold (i.e. the convex envelop) associated with the nonlinear material behavior. Since the locally convex reconstruction resembles the manifold learning technique introduced in [62] or local regression [74], LCDD is expected to reproduce a locally linear constitutive model corresponding to the sampled data points.

It should be emphasized that this linear reproducibility is very attractive in dealing with higher-dimensional phase space when data is relatively scarce, e.g., the elasticity problems in Section 5. As the reconstruction of local convexity confines the solution space for searching optimum local data in a bounded smooth domain, the proposed LCDD approach also avoids the non-convergence issue during the data-driven iterations, which usually appears in the regression based data-driven methods [47,55].

(a)

(b)

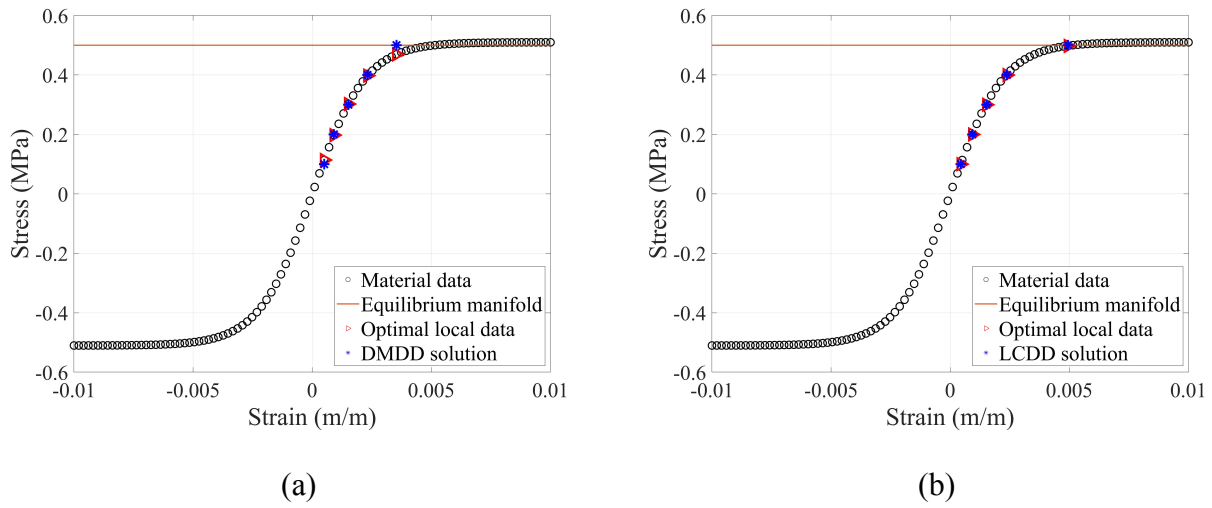**Fig. 8.** Comparison of the (a) DMDD and (b) LCDD solvers for the one-bar truss structure using a sigmoid material database. The database contains $p = 100$ stress-strain data points. The number of neighbor points used in LCDD is $k = 6$.
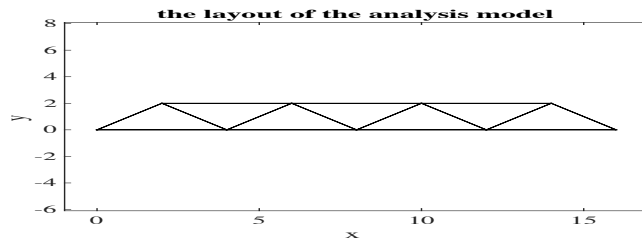
4.2. Example II: Truss system

Fig. 9. A 15-bar truss structure with prescribed displacements and applied loads: $a = 4$ m, $h = 2$ m, $\bar{u}_x = 0.01$ m, and $F = 100$ kN.
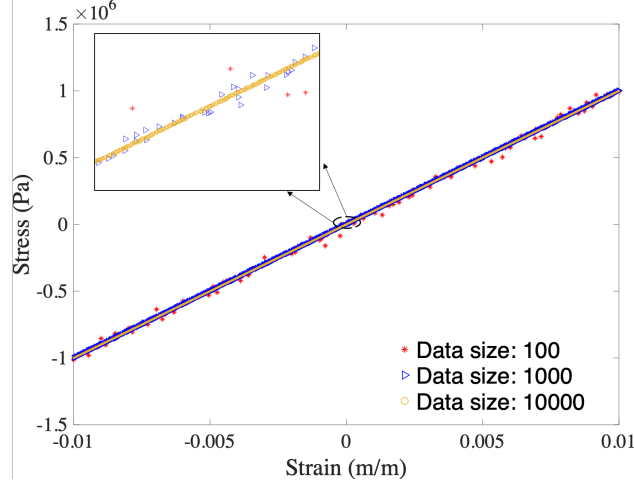
Fig. 10. Three noisy databases with different sizes, $p = 10^2, 10^3, 10^4$, where the random level of noise χ corresponding to each database is inverse to the data size, defined as $\chi = 2p^{-1}$.

To examine the convergence behavior with respect to the material data set size, consider a 15-bar truss structure (i.e., $m = 15$ for the local state vectors $\{s_\alpha = [\varepsilon_\alpha \ \sigma_\alpha]^T\}_{\alpha=1}^m$) with unity cross-sectional area, as illustrated in **Fig. 9**. The solution obtained from different data-driven solvers are compared against the reference solution using the following normalized root-mean-square (%RMS) state errors

$$\varepsilon_{(\%\mathrm{RMS})} = \frac{1}{\varepsilon_{\max}^{\mathrm{ref}}} \left(\frac{1}{m} \sum_{\alpha=1}^m l_\alpha (\varepsilon_\alpha - \varepsilon_\alpha^{\mathrm{ref}})^2 \right)^{1/2}, \tag{33a}$$

$$\sigma_{(\%\mathrm{RMS})} = \frac{1}{\sigma_{\max}^{\mathrm{ref}}} \left(\frac{1}{m} \sum_{\alpha=1}^m l_\alpha (\sigma_\alpha - \sigma_\alpha^{\mathrm{ref}})^2 \right)^{1/2}, \tag{33b}$$

where $\{l_\alpha\}_{\alpha=1}^m$ are the length of the bars, $\{(\varepsilon_\alpha, \sigma_\alpha)\}_{\alpha=1}^m$ are the data-driven solutions for all bar members, $\{(\varepsilon_\alpha^{\mathrm{ref}}, \sigma_\alpha^{\mathrm{ref}})\}_{\alpha=1}^m$ are the strain and stress reference solutions corresponding synthetic material model, and $(\varepsilon_{\max}^{\mathrm{ref}}, \sigma_{\max}^{\mathrm{ref}})$ are the largest absolute values of strain and stress among all bar members.

In this numerical study, we consider three material data sets (see **Fig. 10**) with different sizes (i.e. $p = 10^2, 10^3, 10^4$), where the database with more data points is said to have higher

density or larger data size, for the data-driven simulations. These noisy data sets are superimposed with the strain and stress perturbations given by the random Gaussian noise $\chi = 2p^{-1}$ (refer to (32)). In this case, the underlying structure of the data set uniformly converge to a linear curve with a slope of $M = 100\ \text{MPa}$ as the number of data points increases.

The convergence results of different data-driven solvers (DMDD and LCDD) measured by the normalized RMS state errors in (33) are shown in **Fig. 11** and **Fig. 12**. As suggested by the estimate given in [41,60], the data-driven solutions obtained by both methods converge towards the classical model-based solution with a rate close to 1 as the number of data points increases. However, a less satisfactory result is obtained by DMDD compared to LCDD. In **Fig. 11**, where LCDD yields nearly 1 order of accuracy higher than DMDD. This is due to the locally convex reconstruction that recovers the locally linear manifold. In LCDD, the inherent manifold learning ability contributes to the improved accuracy in addition to the enhancement of robustness against noisy data.
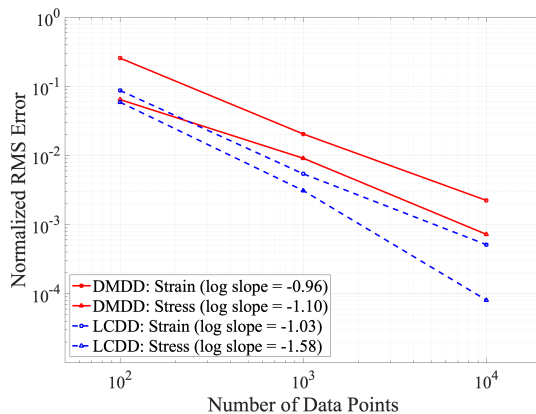

Fig. 11. Truss structure case. The convergence plot of the normalized RMS errors, $\varepsilon_{(\%\text{RMS})}$ and $\sigma_{(\%\text{RMS})}$, against increasing the size of database. The DMDD and LCDD solvers are compared using the three noisy data sets with different sizes, $p = 10^2, 10^3, 10^4$. The number of neighbor points used in LCDD is $k = 12$.
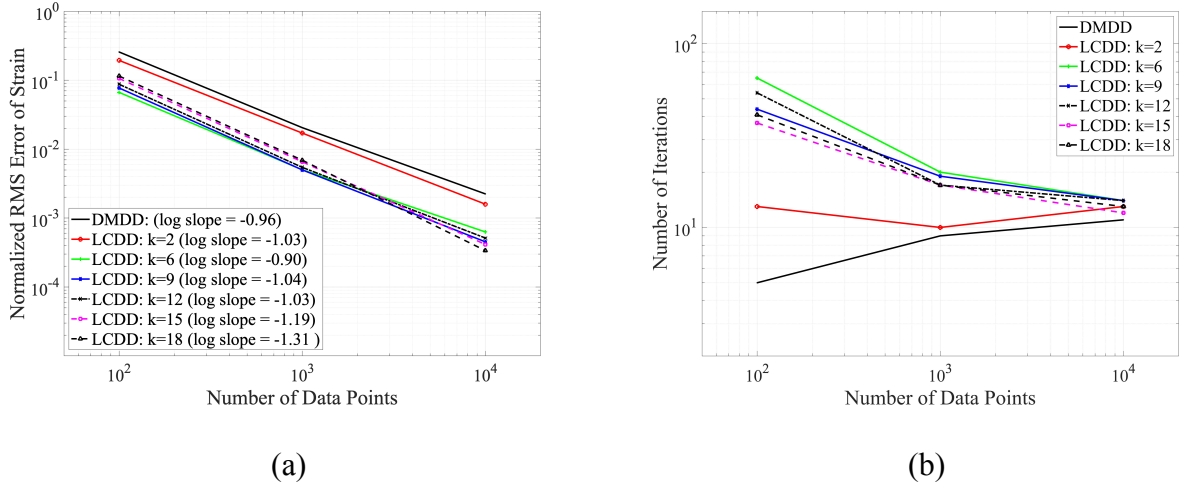
Fig. 12. Truss structure case. (a) The convergence plot of the normalized RMS strain error $\varepsilon_{(\%RMS)}$ and (b) the number of iterations for convergence against increasing the size of database. The DMDD and LCDD solvers are compared using the noisy data sets with three different sizes, $p = 10^2, 10^3, 10^4$. Different number of neighbor points k are used in LCDD.

It is also observed that the performance of the proposed LCDD solver appears to be insensitive to the numbers of convex hull neighbors (from $k = 6$ to $k = 18$) as demonstrated by its solution errors (**Fig. 12**a) and the number of iterations to attain convergence (**Fig. 12**b). Surprisingly, the results in **Fig. 12**b suggests that the LCDD solution converges faster as the data set size increases. This phenomenon is significantly distinct from other data-driven solvers, such as DMDD [41] and the max-ent data-driven solver [42], which require more iterations to achieve convergence when using larger data sets. We believe this is because the local manifold learning of LCDD better represents the underlying manifold. It is worth noting that when using $k = 2$, LCDD appears to lose accuracy and yields solutions approaching DMDD (shown in **Fig. 12**a), implying that LCDD would recover DMDD in the limit of using one neighbor.

A close comparison between the data-driven solutions of DMDD and LCDD using the data set with 100 material points (see **Fig. 10**) is given in **Fig. 13**. The reference solution (denoted by the diamond points) is obtained by utilizing the synthetic linear model (i.e. $\sigma = M\varepsilon$). As can be seen from **Fig. 13**a, the variations of the noisy data set substantially influence the

DMDD performance such that several data-driven solutions (the asterisk points in a dashed box) converge poorly at some local minima that deviate from the linear graph, resulting in an overall unsatisfying performance of DMDD. In contrast, the LCDD solver overcomes such issues with noisy data as shown in **Fig. 13**b.

(a)

(b)
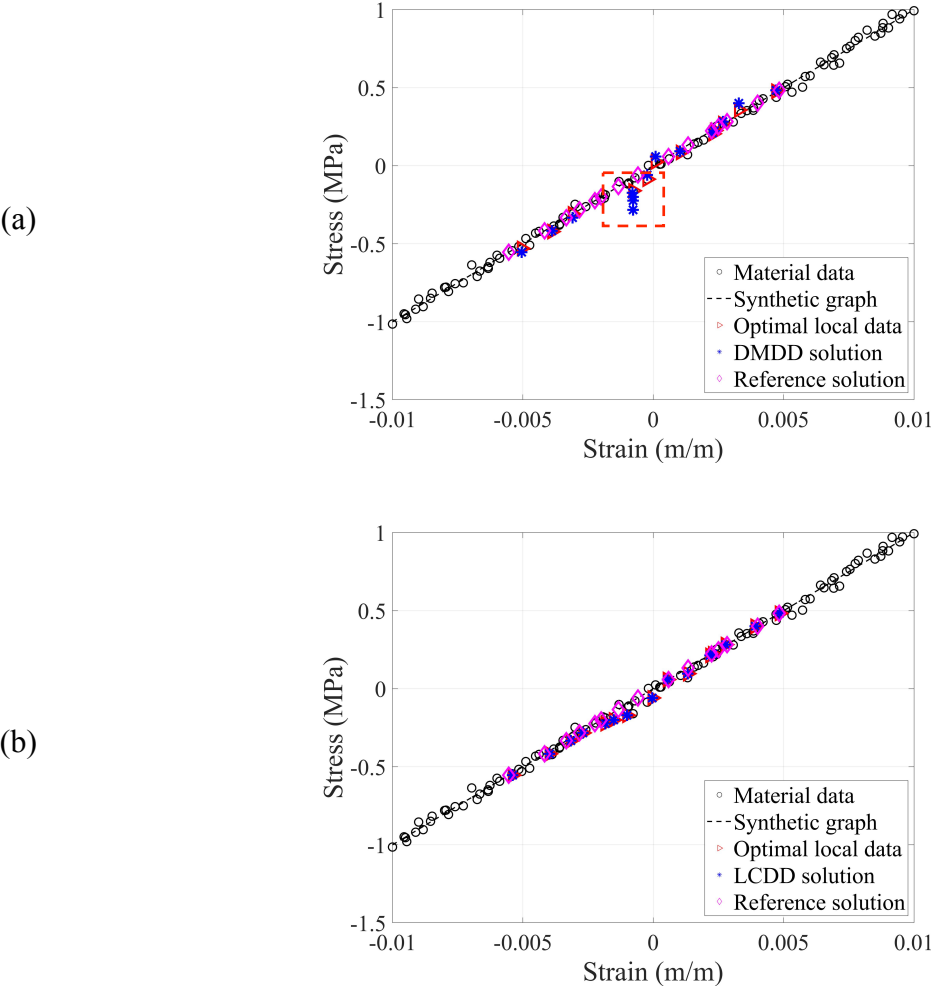
Fig. 13. Comparison of the data-driven solutions of (a) DMDD and (b) LCDD for the truss structure using the noisy material data set of $p = 100$ stress-strain data points. The number of neighbor points used in LCDD is $k = 12$.

5. Numerical Examples: Elasticity

In this section, we apply the proposed method to a two-dimensional elasticity problem. The accuracy and robustness of data-driven solvers are verified with high dimensionality of the associated phase space in the example problems.
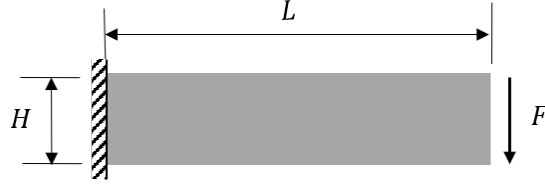
Fig. 14. A beam model subjected to a shear load: $L = 48\text{m}$, $H = 12\text{m}$ and $F = 1000\,\text{N}$.

| (a) | (b) |
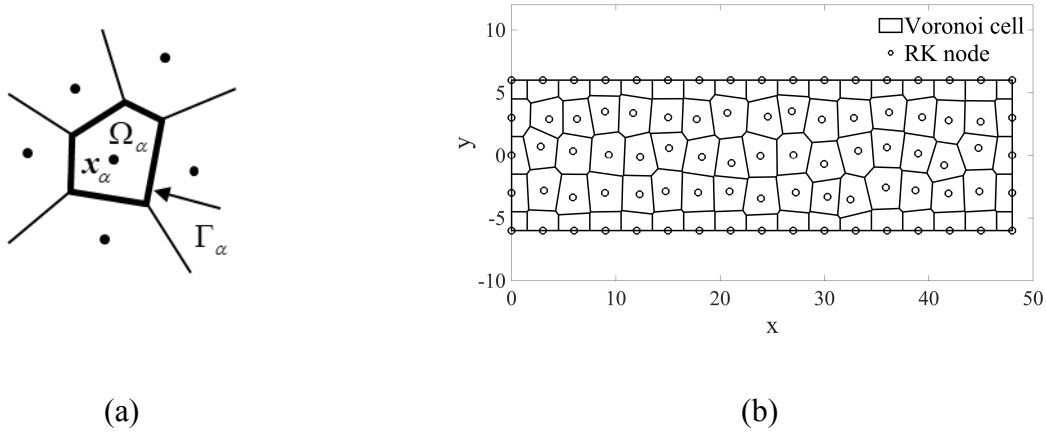
Fig. 15. Schematics of (a) Voronoi diagram and (b) the discretization of the beam model. The RK nodes are also the integration points under the SCNI framework.

A plane stress elastic beam subjected to a shear load as shown in **Fig. 14** is to be modeled. The reproducing kernel particle method (RKPM) with a stabilized conforming nodal integration (SCNI) approach are used for discretization, as shown in **Fig. 15**. More details about RKPM and SCNI can be referred to Appendix B. Under this RKPM-SCNI framework, the volumes of the Voronoi cells Ω_α, $\alpha = 1,...,N$, are used as the weights V_α in (16).

The synthetic elastic material model is given by the classical elasticity law with Young's modulus $E = 30$ MPa and Poisson's ratio $v = 0.3$. In this problem, the coefficient matrix used for norm (19) at each integration point is defined as

$$\boldsymbol{M} = \boldsymbol{M}_\alpha^\varepsilon = \boldsymbol{M}_\alpha^{\sigma^{-1}} = \frac{E}{1-v^2}\begin{bmatrix} 1 & v & 0 \\ v & 1 & 0 \\ 0 & 0 & (1-v)/2 \end{bmatrix}. \tag{34}$$

To evaluate the performance of the data-driven solvers, the following normalized root-mean-square (%RMS) state error is defined for high-dimensional state,

$$\omega_{(\%\mathrm{RMS})} = \left(\frac{\displaystyle\sum_{\alpha=1}^{N} V_\alpha \| \boldsymbol{s}_\alpha - \boldsymbol{s}_\alpha^{\mathrm{ref}} \|_M^2}{\displaystyle\sum_{\alpha=1}^{N} V_\alpha \| \boldsymbol{s}_\alpha^{\mathrm{ref}} \|_M^2} \right)^{\frac{1}{2}}, \tag{35}$$

Where $\boldsymbol{s}_\alpha^{\mathrm{ref}} = [\boldsymbol{\varepsilon}_\alpha^{\mathrm{ref}\,T}\ \boldsymbol{\sigma}_\alpha^{\mathrm{ref}\,T}]^T$ denotes the nodal strain and stress reference solutions solved by using the synthetic material model, while $\boldsymbol{s}_\alpha = [\boldsymbol{\varepsilon}_\alpha^T\ \boldsymbol{\sigma}_\alpha^T]^T$ denotes the solutions solved by data-driven solvers using a given material data set.

Following the procedures of data generation in (32), the noiseless stress-strain data points using synthetic elastic material model in (34) are first generated, where the strains ε_x, ε_y and γ_{xy} are defined within the range $[-5 \times 10^{-4}, 5 \times 10^{-4}]$. The noisy data set is generated by superimposing the three components of noiseless strain and stress data, respectively, with the associated Gaussian noise term defined in (32), where $\chi = 0.4 / \left(\sqrt[3]{p} l_{(j)} \right)$ and $l_{(j)}$ is the maximum value associated to j-th component of the noiseless data. Four material data sets in various size (i.e., $p = 10^3, 20^3, 40^3, 80^3$) are considered for the beam model.

The performance of the data-driven solvers using the noiseless data sets and the noisy data sets are given in **Fig. 16** and **Fig. 17**, respectively. Consistent to the convergence estimate in [41], the DMDD solutions converge linearly to the reference solution against the cubic root of the number of data points, regardless of using noiseless or noisy databases. LCDD using noiseless data sets (**Fig. 16**a) generates data-driven solution with the error as small as the

convergence tolerance in the iterative data-driven computing (see Section 3.3). This implies that LCDD perfectly captures the underlying linear material graph even in such a high dimensional phase space. The convergence study with noisy data sets (**Fig. 17**a) shows that the LCDD solution using a sparse data set ($p = 10^3$) is able to achieve higher accuracy than the DMDD solution obtained using a very dense data set ($p = 80^3$), suggesting the superiority of LCDD over DMDD. Considering that it is difficult, in practice, to obtain a database with sufficiently dense data for high-dimensional spaces, the proposed LCDD approach is attractive.

As the intrinsic dimensionality of the employed linear elastic database is $d = 2$, it is interesting to observe from **Fig. 16**a and **Fig. 17**a that the LCDD solutions obtained by using $k = 3$ ($d < k < 2q$, where $2q$ is the dimension of the material dataset E) present an intermediate solution between the DMDD solution (i.e. $k = 1$) and the other LCDD solutions with using more neighbor points $k \geq 2q = 6$. The results indicate the importance of including enough neighbors in convex hull to fully preserve the manifold learning capacity in LCDD. In the case of noiseless data we observe that LCDD with $k = 6$ is sufficient as its results are almost identical to the case with $k = 9$ results(see **Fig. 16**).

(a)   (b)

Fig. 16. Shear beam model with noiseless data sets. (a) The convergence plot of the normalized RMS state error $\omega_{(\%RMS)}$ and (b) the number of iterations for convergence against increasing the size of database. The DMDD and LCDD solvers are compared using the noiseless data sets with

four different sizes, $p = 10^3, 20^3, 40^3, 80^3$. Different numbers of neighbor points k are used in LCDD.

The associated number of convergence steps for the data-driven solvers are also presented in **Fig. 16**b and **Fig. 17**b. In contrast to the DMDD solver where the number of iterations increases with using a larger database, there is no evident increase for the LCDD solver. Moreover, the comparison of **Fig. 16**b and **Fig. 17**b shows that LCDD does not require more iterations between the local and global steps to converge when dealing with the noisy database rather than the noiseless database. It suggests that the convergence of LCDD is not sensitive to the size of database as well as the data sampling quality.

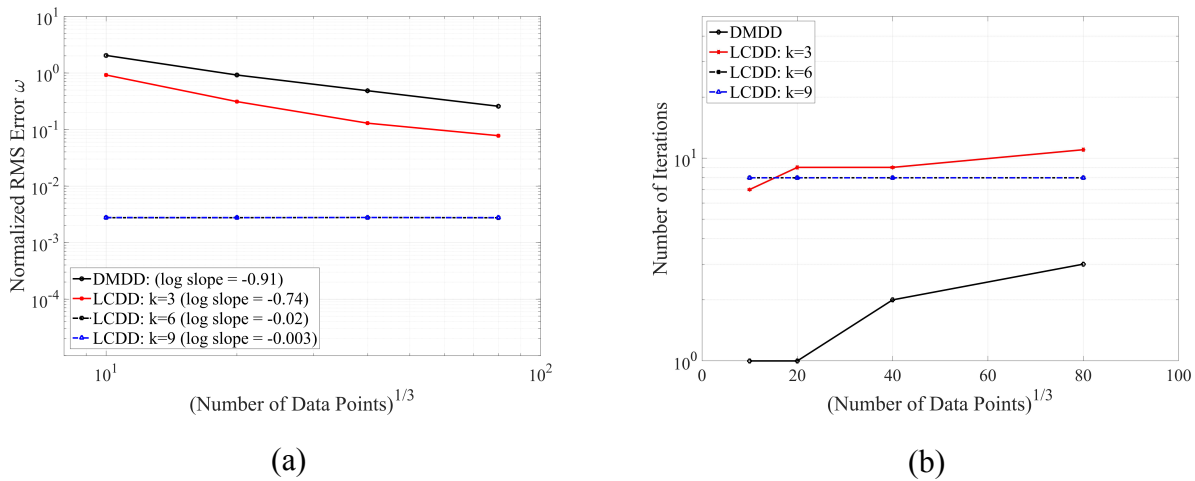(a)   (b)
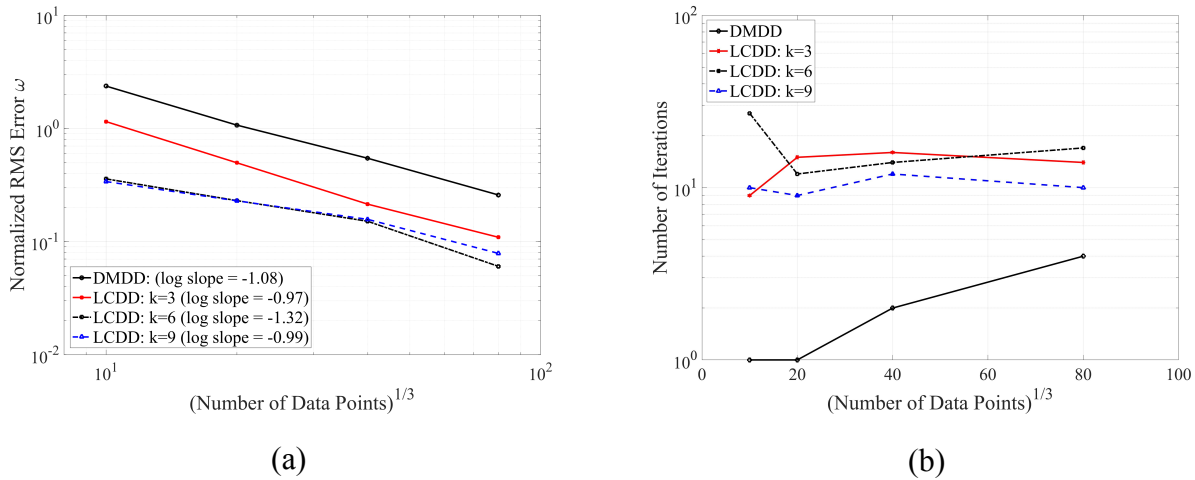
Fig. 17. Shear beam model with noisy data sets. (a) The convergence plot of the normalized RMS state error $\omega_{(\%RMS)}$ and (b) the number of iterations for convergence against increasing the size of database. The DMDD and LCDD solvers are compared using the noisy data sets with four different sizes, $p = 10^3, 20^3, 40^3, 80^3$. Different numbers of neighbor points k are used in LCDD.

The beam deformations simulated by the data-driven solvers are also compared in **Fig. 18** and **Fig. 19**. It is observed that DMDD performs poorly (**Fig. 18**a and **Fig. 19**a) due to the susceptibility to noisy data and local minimum issues that are more pronounced in elasticity

problems with a high-dimensional phase space. On the other hand, LCDD exactly reproduces the reference solutions when using noiseless database (**Fig. 18**b), and results in marginal deviations from the reference when using the noisy database (**Fig. 19**b). Moreover, **Fig. 20** shows the stress solutions, σ_{xx} and σ_{xy}, obtained by LCDD compared to the constitutive model-based reference solutions. It shows that LCDD can yield accurate stress solutions across the problem domain with using a noisy database. This demonstrates that the LCDD approach remains robust with noisy data in solving elasticity problems.

(a)   (b)
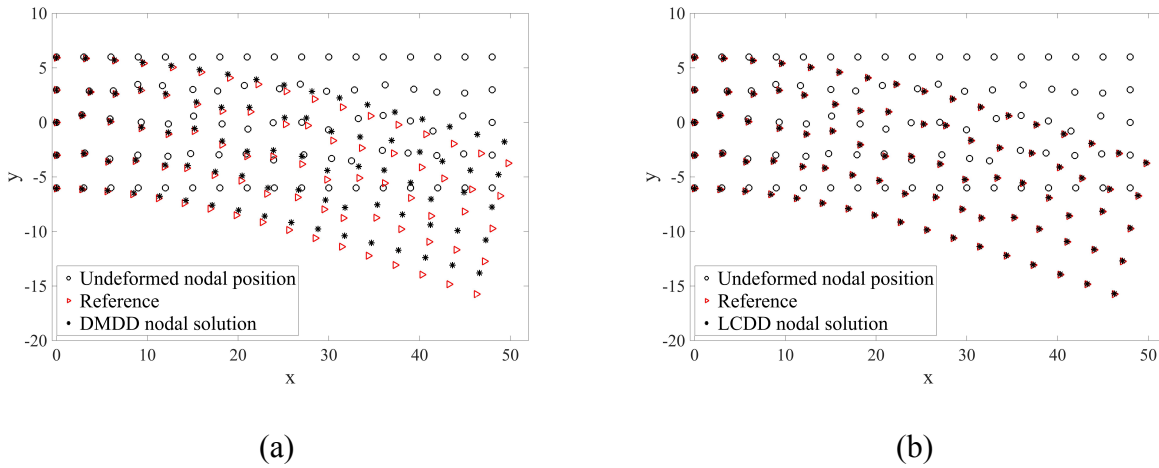
Fig. 18. Comparison of the data-driven displacement solutions for the shear beam model by using (a) DMDD with a noiseless data set of $p = 80^3$ stress-strain data points and (b) LCDD with a noiseless data set of $p = 10^3$ stress-strain data points. The number of neighbor points used in LCDD is $k = 6$.
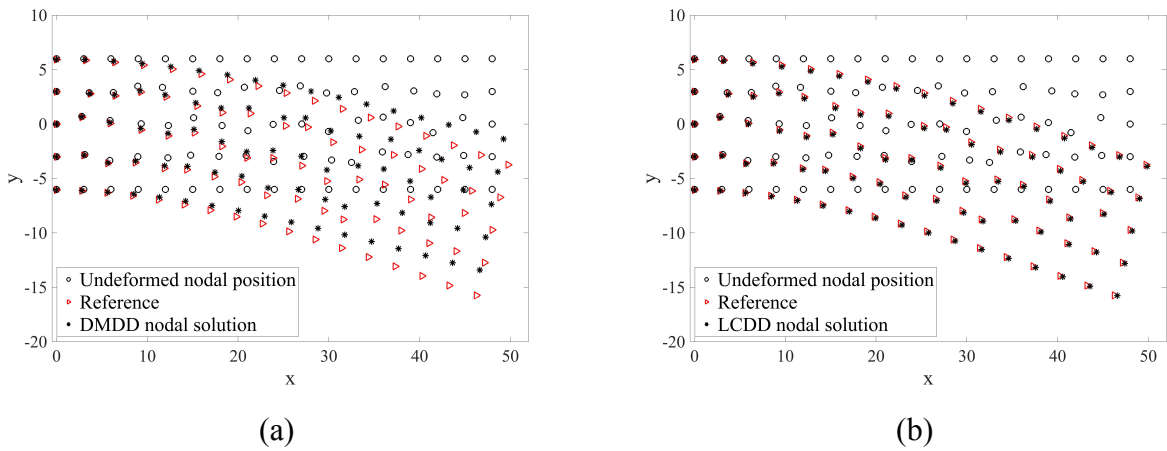
Fig. 19. Comparison of the data-driven displacement solutions for the shear beam model by using (a) DMDD with a noisy data set of $p = 80^3$ stress-strain data points and (b) LCDD with a noisy data set of $p = 10^3$ stress-strain data points. The number of neighbor points used in LCDD is $k = 6$
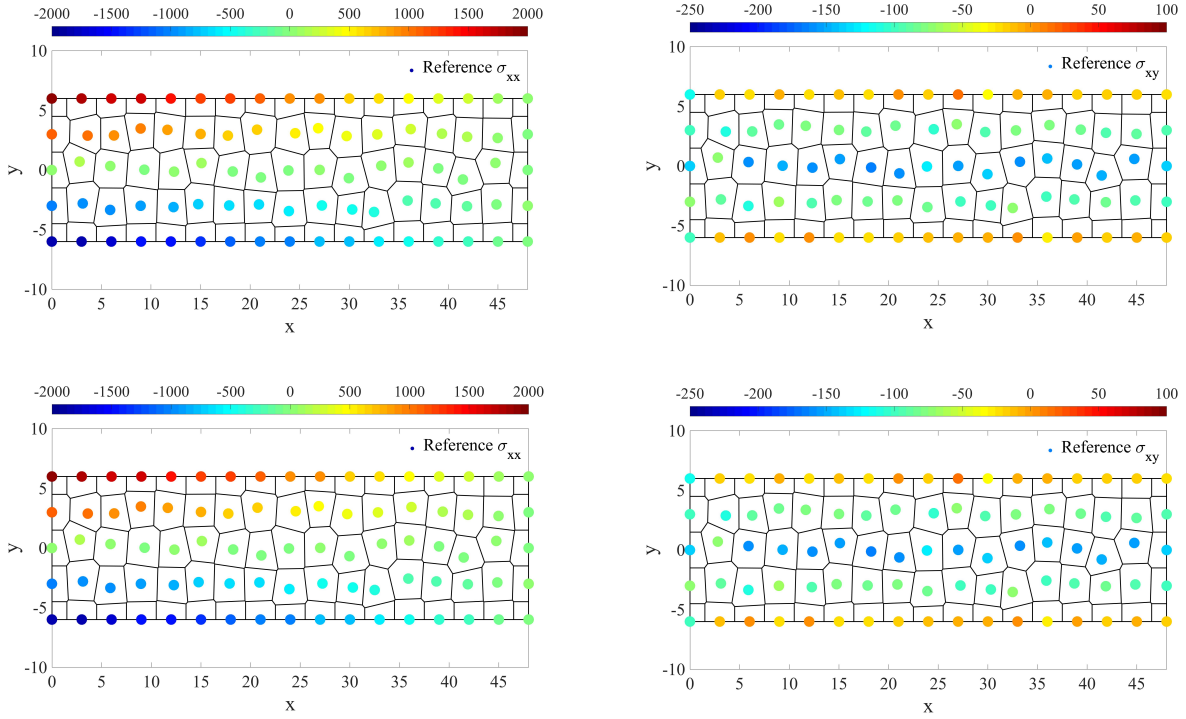
Fig. 20. Comparison of the stress solutions, σ_{xx} and σ_{xy}, of the reference constitutive model-based computing (upper) and the proposed LCDD computing (bottom) for the shear beam model. A noisy data set of $p = 10^3$ stress-strain data points and $k = 6$ neighbor points are used in LCDD.

6. Discussion & Conclusion

We have formulated a local convexity data-driven (LCDD) solver as a new data-driven computing paradigm integrated with manifold learning techniques that generalizes the distance-minimizing data-driven (DMDD) computing [41] with enhanced accuracy and robustness of data-driven computing against noise/outliers. The proposed method adaptively selects k nearest neighbor (k-NN) material data points for physical local states to be updated and searches for optimal data solutions from a bounded solution space defined by the convex hull of the selected k-NN points. This local data searching procedure has been formulated under a *non-negative least squares* (NNLS) solver and can be solved efficiently. By means of the clustering analysis based on k-NN and preserving the convexity of the constructed local material manifold, LCDD yields enhanced robustness and convergence stability.

From the *pure data-driven approach* point of view (refer to [54]), LCDD is inspired by measuring the distance to a local convex set E, instead of from a single discrete data, aiming to enhance the robustness against noise and prevent undesirable local minima. LCDD can be reduced to the standard DMDD approach when using only one neighbor data point, i.e. $k = 1$. Thus, LCDD retains the simplicity and is computationally efficiency compared to other robustness enhanced data-driven methods with introducing statistics information [42,54]. From the *fitted data-driven (or linearization) approach* point of view, on the other hand, LCDD relies on the approximation of locally linear material graph by the manifold learning methodologies [21,63] to capture the global structure via local data information. However, the proposed LCDD scheme distinguishes itself from the other manifold learning based data-driven approaches [47, 55] with the following two aspects: first, the iteration process of LCDD does not need to explicitly construct constitutive manifold and use tangent information; second, LCDD introduces the convexity condition on the reconstructed material graph, thus avoiding convergence issues that occur when using standard regression approaches. In addition, we believe preserving convexity is also of physical importance. For example, it is expected to better preserve the positivity of strain energy via LCDD rather than other manifold learning techniques because the proposed approach learns the underlying constitutive manifold based on a convex combination of a cluster of neighboring data points, although a rigorous analysis of this question is beyond the scope of this paper.

It has also been shown that the embedded NNLS solver seeks the projection point of a given computational state onto a nearby material graph implicitly constructed based on the k-NN points, which ensures local linear reproducibility in the approximation. Hence, in addition to the improved robustness and accuracy in dealing with noisy database, LCDD exactly represent the underlying linear stress-strain relationship. In the proposed global-local data-driven algorithm, smooth solution spaces are employed for the global physical solution and the local project on the convex set. This is achieved by introducing a RK shape functions in the approximation of the global physics laws and the regularized LCDD learning solver in the optimal data set search. With the SCNI domain integration employed in the global Galerkin equations, it significantly reduces the needed stress-strain data search in the local LCDD learning solver, leading to an effective data-driven computing. The proposed LCDD dada-driven method has been applied to truss problems with linear and nonlinear stress-strain relationship, and continuum elasticity problems and demonstrated its effectiveness in robustness, convergence, accuracy in high-dimensional phase spaces.

This paper is intended to introduce manifold learning techniques, or dimensionality reduction [21], to data-driven computing. Our numerical studies show that it is effective in applying manifold learning for problems with high-dimensional data, because in high-dimensional spaces the data can be extremely sparse and the acquisition of sufficient data is not practical. This demands effective dimensionality reduction to identify and extract the essential information from the database, and elasticity example in Section 5 demonstrates the suitability of LCDD with inherent manifold learning for such problems.

Although manifold learning has been shown to enhance the convergence performance during data-driven iteration, the computational cost of each local step remains linearly scaled with the size of material database. Thus, when the datasets become large with high-dimensional information, such as time-dependent states [43], inelastic quantities [35,50,51,82], and so on, more advanced machine learning models for manifold learning like autoencoder can be employed. This is the direction of our further research.

Acknowledgements

Appendix A. Non-negative least squares solver

Let us recall a standard non-negative least squares (NNLS) problem: *given a matrix* $A \in \circ^{n \times p}$ *(usually* $p > n$*) and a observed vector* $z \in \circ^{n}$*, find a nonnegative vector* $y^* \in \circ^{p}_{+}$ *to minimize the following function,*

$$y^* = \arg\min_{x \in \circ^{p}} \| Ay - z \|,$$

$$\text{subject to: } y_i \geq 0, \ i = 1, ..., p, \tag{A.1}$$

where $\| \cdot \|$ *stands for the standard Euclidean norm.*

A variety of methods have been applied to tackle the NNLS problem since 1980s. Those algorithms in general can be roughly categorized into active-set methods and iterative approaches [75]. Lawson and Hanson [56] seems to propose the first standard algorithm to solve NNLS problem (A.1). Their method is essentially an active set method [76], which is based on the observation that only a small subset of the non-negative constraints are usually active at the solution. It shows in [56] that the iteration in the active set method converges and terminates without any cutoff in iterations. The standard algorithm for active-set method is reviewed in Algorithm 1.

Algorithm 1 Non-negative least squares solver: $\boldsymbol{y}^* \leftarrow NNLS(\boldsymbol{A}, \boldsymbol{z}, TOL)$
Input: $\boldsymbol{A} \in \degree^{n \times p}$, $\boldsymbol{z} \in \degree^{n}$, TOL
Output: $\boldsymbol{y}^*$ f $= \boldsymbol{0}$ such that $\boldsymbol{y}^* = \arg\min \|\boldsymbol{Ay} - \boldsymbol{z}\|^2$
Initialization: $Z \leftarrow \varnothing$, $Y \leftarrow \{1, 2, K, p\}$, $\boldsymbol{y} = \boldsymbol{0}$, $\boldsymbol{r} \leftarrow \boldsymbol{z}$
WHILE $\|\boldsymbol{r}\| / \|\boldsymbol{z}\| > TOL$ and $Y \neq \varnothing$, DO
$\quad \boldsymbol{q} \leftarrow \boldsymbol{A}^T(\boldsymbol{z} - \boldsymbol{Ay})$, $j = \arg\max\limits_{i=1,\dots,p}(q_i)$
$\quad$ Include the index j in Z and remove it from Y
$\quad \boldsymbol{s}_Z \leftarrow (\boldsymbol{A}_Z^T \boldsymbol{A}_Z)^{-1} \boldsymbol{A}_Z^T \boldsymbol{z}$, $\boldsymbol{s}_Y \leftarrow \boldsymbol{0}$, where $\boldsymbol{A}_Z \in \degree^{n \times |Z|}$
$\quad$ WHILE $\min\limits_{i \in Z} \boldsymbol{s}_Z \leq 0$, DO
$\quad\quad \alpha = -\min\limits_{i \in Z} y_i / (y_i - s_i)$
$\quad\quad \boldsymbol{y} \leftarrow \boldsymbol{y} + \alpha(\boldsymbol{s} - \boldsymbol{y})$
$\quad\quad$ Update Y with zero value indices of $\boldsymbol{y}$ and Z with the positive indices of $\boldsymbol{y}$
$\quad\quad \boldsymbol{s}_Z \leftarrow (\boldsymbol{A}_Z^T \boldsymbol{A}_Z)^{-1} \boldsymbol{A}_Z^T \boldsymbol{z}$, $\boldsymbol{s}_Y \leftarrow \boldsymbol{0}$
$\quad \boldsymbol{y} \leftarrow \boldsymbol{s} = \boldsymbol{s}_Z \cup \boldsymbol{s}_Y$
$\quad \boldsymbol{r} \leftarrow \boldsymbol{z} - \boldsymbol{Ay}$
Return $\boldsymbol{y}^* \leftarrow \boldsymbol{y}$

Appendix B. Meshfree approximation

In the section we review the reproducing kernel paritile method (RKPM) [57,58] and the stabilized conforming nodal integration (SCNI) approach [59] that are used for solving the weak form in (13).

B.1. Reproducing kernel approximation

The reproducing kernel (RK) shape functions $\{\Psi_I\}_{I=1}^{N}$ used for approximating displacement and Lagrange multipliers in (14) are expressed as

$$\Psi_I(\boldsymbol{x}) = C(\boldsymbol{x}; \boldsymbol{x} - \boldsymbol{x}_I)\Phi_a(\boldsymbol{x} - \boldsymbol{x}_I) \tag{B.1}$$

The kernel function Φ_a defines a local support for the shape functions by a support size "a" as well as the smoothness of the approximation. A widely used kernel function is the cubic-B splines that provides C^2 continuity, expressed as

$$\Phi_a(\boldsymbol{x} - \boldsymbol{x}_I) = \Phi_a(z) = \begin{cases} 2/3 - 4z^2 + 4z^3 & \text{for } 0 \leq z < 1/2 \\ 4/3 - 4z + 4z^2 - 4/3\,z^3 & \text{for } 1/2 \leq z < 1 \\ 0 & \text{for } z \geq 1 \end{cases} \tag{B.2}$$

where $z = \| \boldsymbol{x} - \boldsymbol{x}_I \|/a$. The term $C(\boldsymbol{x}; \boldsymbol{x} - \boldsymbol{x}_I)$ is a correction function constructed using a set of basis functions,

$$C(\boldsymbol{x}; \boldsymbol{x} - \boldsymbol{x}_I) = \sum_{i+j+k=0}^{n} (x_1 - x_{1I})^i (x_2 - x_{2I})^j (x_3 - x_{3I})^k b_{ijk}(\boldsymbol{x}) = \boldsymbol{H}^T(\boldsymbol{x} - \boldsymbol{x}_I)\boldsymbol{b}(\boldsymbol{x}) \tag{B.3}$$

in which $\boldsymbol{H}(\boldsymbol{x} - \boldsymbol{x}_I)$ is a vector consisting of all the monomial basis functions upto *n*-th order, and $\boldsymbol{b}$ is an unknown parameter vector determined by enforcing the *n*-th order reproducing conditions as follows,

$$\sum_{I=1}^{N} \Psi_I(\boldsymbol{x}) x_{1I}^i x_{2I}^j x_{3I}^k = x_1^i x_2^j x_3^k, \quad |i + j + k| = 0, 1, ..., n \tag{B.4}$$

Introducing Eqs. (B.1) and (B.3) into (B.4), the coefficient vector can be obtained by

$$\boldsymbol{b}(\boldsymbol{x}) = \boldsymbol{M}^{-1}(\boldsymbol{x})\boldsymbol{H}(\boldsymbol{0}) \tag{B.5}$$

where the moment matrix is

$$M(x) = \sum_{I=1}^{N} H(x - x_I) H^T(x - x_I) \Phi_a(x - x_I) \qquad \text{(B.6)}$$

Finially, the RK shape functions is expressed as,

$$\Psi_I(x) = H^T(0) M^{-1}(x) H(x - x_I) \Phi_a(x - x_I) \qquad \text{(B.7)}$$

It should be noted that the above RK shape functions do not possess the Kronecker delta property. Thus, certain techniques are needed to impose kinematically admissible approximation on the essential boundary, such as Lagrange multipliers method [77], Nitsche's method [78,79], boudary singular kernel method [80], and transformation methods [58,80]. More discussions on the mathematical properties of the reproducing kernel approaximation can be found in the review paper [81].

B.2. Nodal integration scheme

The SCNI approach is employed for the domain integration of the weak form (13) to achieve computational efficiency and accuracy when using meshfree shape functions with nodal integration quadrature schemes.

The key idea behind SCNI is to satisfy the linear patch test (thus, ensure the linear consistency) by leveraging a condition, i.e. the divergence constraint on the test function space and numerical integration [59], expressed as:

$$\int_{\Omega}^{\wedge} \nabla \Psi_I \, d\Omega = \int_{\partial\Omega}^{\wedge} \Psi_I \boldsymbol{n} \, d\Gamma, \qquad \text{(B.8)}$$

where '^' over the integral symbol denotes numerical integration. In SCNI, an effective way to achieve Eq. (B.8) is based on nodal integration with gradients smoothed over conforming representative nodal domains, as shown in **Fig. 21**, converted to boundary integration using the divergence theorem

$$\tilde{\nabla}\Psi_I(x_L) = \frac{1}{V_L} \int_{\Omega_L} \nabla \Psi_I \, d\Omega = \frac{1}{V_L} \int_{\partial\Omega_L} \Psi_I \boldsymbol{n} \, d\Gamma, \qquad \text{(B.9)}$$

where $V_L = \int_{\Omega_L} d\Omega$ is the volume of a conforming smoothing domain associated to the node x_L, and $\tilde{\nabla}$ denotes the smoothed gradient operator. In this method, smoothed gradients are employed for both test and trial functions, as the approximation in (B.9) enjoys first order completeness

49

and leads to a quadratic rate of convergence for solving linear solid problems by meshfree Galerkin methods. As shown in **Fig. 21**, the continuum domain Ω is partitioned into N conforming cells by Voronoi diagram, and both the nodal displacement vectors and the state variables (e.g., stress, strain) are defined at the set of nodes at $\{x_L\}_{L=1}^{N}$.

Therefore, if we consider two-dimensional elasticity problem under the SCNI framework, the smoothed strain-displacement matrix $\tilde{B}_I^0(x_L)$ used in (16) is expressed as:

$$\tilde{B}_I^0(x_L) = \begin{bmatrix} \tilde{B}_{I1}^0(x_L) & 0 \\ 0 & \tilde{B}_{I2}^0(x_L) \\ \tilde{B}_{I2}^0(x_L) & \tilde{B}_{I1}^0(x_L) \end{bmatrix}, \tag{B.10}$$

$$\tilde{B}_{Ii}^0(x_L) = \frac{1}{V_L} \int_{\partial\Omega_L} \Psi_I(x) n_i(x) d\Gamma. \tag{B.11}$$

Since the employment of the smoothed gradient operator in (B.9) and (B.11) satisfies the divergence constraint regardless of the numerical boundary integration, a two trapezoidal rule for each segment of $\partial\Omega_L$ is used in this study.
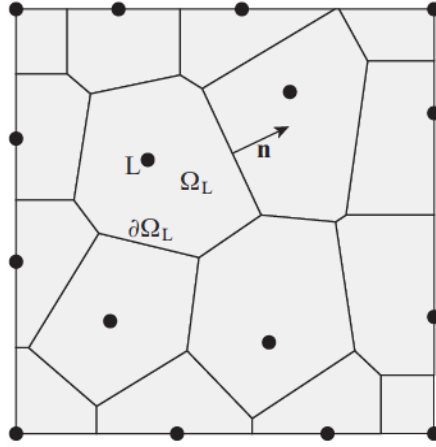

Fig. 21. Illustration of Voronoi diagram for SCNI.

Reference

[1] F. Darema, Dynamic data driven applications systems: A new paradigm for application simulations and measurements, in: Int. Conf. Comput. Sci., 2004: pp. 662–669.

[2] S. Baker, J. Berger, P. Brady, K. Borne, S. Glotzer, R.H.D. Johnson, A. Karr, D. Keyes, B. Pate, H. Prosper, Data-enabled science in the mathematical and physical sciences, 2010.

[3] J.N. Kutz, Data-driven modeling & scientific computation: methods for complex systems & big data, Oxford University Press, 2013.

[4] D.T. Larose, Discovering knowledge in data: an introduction to data mining, John Wiley & Sons, 2014. doi:10.1016/j.cll.2007.10.008.

[5] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning, Springer series in statistics, New York, 2009. doi:10.1007/b94608.

[6] K.P. Murphy, Machine Learning: A Probabilistic Perspective, Cambridge: MIT press, Cambridge, 2012. doi:10.1007/978-3-642-21004-4_10.

[7] C. Angermueller, T. Pärnamaa, L. Parts, O. Stegle, Deep learning for computational biology, Mol. Syst. Biol. (2016). doi:10.15252/msb.20156651.

[8] I. Kononenko, Machine learning for medical diagnosis: History, state of the art and perspective, Artif. Intell. Med. (2001). doi:10.1016/S0933-3657(01)00077-X.

[9] K. Rajan, Materials informatics, Mater. Today. 8 (2005) 38–45.

[10] R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, C. Kim, Machine learning in materials informatics: Recent applications and prospects, Npj Comput. Mater. (2017). doi:10.1038/s41524-017-0056-5.

[11] McKinsey & Company, Big data: The next frontier for innovation, competition, and productivity, McKinsey Glob. Inst. (2011). doi:10.1080/01443610903114527.

[12] J. Ghaboussi, J.H. Garrett, X. Wu, Knowledge Based Modeling of Material Behavior with Neural Networks, J. Eng. Mech. 117 (1991) 132–153. doi:10.1061/(ASCE)0733-9399(1991)117:1(132).

[13] J. Ghaboussi, D.E. Sidarta, New Nested Adaptive Neural Networks (NANN) for Constitutive Modeling, Comput. Geotech. 22 (1998) 29–52. doi:10.1016/S0266-352X(97)00034-7.

[14] M. Lefik, B.A. Schrefler, Artificial neural network as an incremental non-linear constitutive model for a finite element code, Comput. Methods Appl. Mech. Eng. 192 (2003) 3265–3283. doi:10.1016/S0045-7825(03)00350-5.

[15] M. Milano, P. Koumoutsakos, Neural network modeling for near wall turbulent flow, J. Comput. Phys. 182 (2002) 1–26. doi:10.1006/jcph.2002.7146.

[16] B.D. Tracey, K. Duraisamy, J.J. Alonso, A Machine Learning Strategy to Assist Turbulence Model Development, in: 53rd AIAA Aerosp. Sci. Meet., 2015. doi:10.2514/6.2015-1287.

[17]   Z.J. Zhang, K. Duraisamy, Machine Learning Methods for Data-Driven Turbulence Modeling, in: 22nd AIAA Comput. Fluid Dyn. Conf., 2015. doi:10.2514/6.2015-2460.

[18]   M. Schmidt, H. Lipson, Distilling free-form natural laws from experimental data, Science (80-.). 324 (2009) 81–85. doi:10.1126/science.1165893.

[19]   S.L. Brunton, J.L. Proctor, J.N. Kutz, Discovering governing equations from data: Sparse identification of nonlinear dynamical systems, Proc. Natl. Acad. Sci. 113 (2016) 3932–3937. doi:10.1073/pnas.1517384113.

[20]   M. Raissi, G.E. Karniadakis, Hidden physics models: Machine learning of nonlinear partial differential equations, J. Comput. Phys. 357 (2018) 125–141. doi:10.1016/j.jcp.2017.11.039.

[21]   J.A. Lee, M. Verleysen, Nonlinear Dimensionality Reduction, Springer Science & Business Media, 2007. doi:10.1109/TNN.2008.2005582.

[22]   S. Haykin, Neural Networks and Learning Machines, Pearson, NJ, USA, 2009. doi:978-0131471399.

[23]   B.A. Le, J. Yvonnet, Q.-C. He, Computational homogenization of nonlinear elastic materials using neural networks, Int. J. Numer. Methods Eng. 104 (2015) 1061–1084. doi:10.1002/nme.4953.

[24]   S. Bhattacharjee, K. Matouš, A nonlinear manifold-based reduced order model for multiscale analysis of heterogeneous hyperelastic materials, J. Comput. Phys. 313 (2016) 635–653. doi:10.1016/j.jcp.2016.01.040.

[25]   M.A. Bessa, R. Bostanabad, Z. Liu, A. Hu, D.W. Apley, C. Brinson, W. Chen, W.K. Liu, A framework for data-driven analysis of materials under uncertainty: Countering the curse of dimensionality, Comput. Methods Appl. Mech. Eng. (2017). doi:10.1016/j.cma.2017.03.037.

[26]   G.B. Olson, Designing a New Material World, Science (80-.). 288 (2000) 993–998. doi:10.1126/science.288.5468.993.

[27]   K. Wang, W.C. Sun, A multiscale multi-permeability poroplasticity model linked by recursive homogenizations and deep learning, Comput. Methods Appl. Mech. Eng. (2018). doi:10.1016/j.cma.2018.01.036.

[28]   A. Oishi, G. Yagawa, Computational mechanics enhanced by deep learning, Comput. Methods Appl. Mech. Eng. (2017). doi:10.1016/j.cma.2017.08.040.

[29]   D. Stoecklein, K.G. Lore, M. Davies, S. Sarkar, B. Ganapathysubramanian, Deep Learning for Flow Sculpting: Insights into Efficient Learning using Scientific Simulation Data, Sci. Rep. (2017). doi:10.1038/srep46368.

[30]   J.T. Oden, T. Belytschko, J. Fish, T.J.R. Hughes, C. Johnson, D. Keyes, A. Laub, L. Petzold, D. Srolovitz, S. Yip, Simulation-based engineering science: Revolutionizing engineering science through simulation, Natl. Sci. Found. (2006) 1–88.

[31]   R. Ibañez, D. Borzacchiello, J.V. Aguado, E. Abisset-Chavanne, E. Cueto, P. Ladeveze, F. Chinesta, Data-driven non-linear elasticity: constitutive manifold construction and

problem discretization, Comput. Mech. (2017) 1–14. doi:10.1007/s00466-017-1440-1.

[32] J. Ling, R. Jones, J. Templeton, Machine learning strategies for systems with invariance properties, J. Comput. Phys. 318 (2016) 22–35. doi:10.1016/j.jcp.2016.05.003.

[33] M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics Informed Deep Learning (Part I): Data-driven Discovery of Nonlinear Partial Differential Equations, (2017) 1–22. doi:10.1103/PhysRevE.87.030803.

[34] A. Koscianski, S. De Cursi, Physically Constrained Neural Networks and Regularization of Inverse Problems, in: 6 Th World Congr. Struct. Multidiscip. Optim., 2005.

[35] R. Ibáñez, E. Abisset-Chavanne, D. González, J.L. Duval, E. Cueto, F. Chinesta, Hybrid constitutive modeling: data-driven learning of corrections to plasticity models, Int. J. Mater. Form. (2018). doi:10.1007/s12289-018-1448-x.

[36] G. Evensen, Data assimilation: The ensemble kalman filter, Springer Science & Business Media, 2010. doi:10.1007/978-3-540-38301-7.

[37] Z. Liu, M.A. Bessa, W.K. Liu, Self-consistent clustering analysis: An efficient multi-scale scheme for inelastic heterogeneous materials, Comput. Methods Appl. Mech. Eng. (2016). doi:10.1016/j.cma.2016.04.004.

[38] K. Matouš, M.G.D. Geers, V.G. Kouznetsova, A. Gillman, A review of predictive nonlinear theories for multiscale modeling of heterogeneous materials, J. Comput. Phys. 330 (2017) 192–220.

[39] B. Peherstorfer, K. Willcox, Dynamic data-driven reduced-order models, Comput. Methods Appl. Mech. Eng. 291 (2015) 21–41. doi:10.1016/j.cma.2015.03.018.

[40] Q. He, J.S. Chen, C. Marodon, A decomposed subspace reduction for fracture mechanics based on the meshfree integrated singular basis function method, Comput. Mech. 63 (2019) 593–614. doi:10.1007/s00466-018-1611-8.

[41] T. Kirchdoerfer, M. Ortiz, Data-driven computational mechanics, Comput. Methods Appl. Mech. Eng. 304 (2016) 81–101. doi:10.1016/j.cma.2016.02.001.

[42] T. Kirchdoerfer, M. Ortiz, Data Driven Computing with noisy material data sets, Comput. Methods Appl. Mech. Eng. 326 (2017) 622–641. doi:10.1016/j.cma.2017.07.039.

[43] T. Kirchdoerfer, M. Ortiz, Data-driven computing in dynamics, Int. J. Numer. Methods Eng. 113 (2018) 1697–1710. doi:10.1002/nme.5716.

[44] M. Bonne, A. Constantinescu, Inverse problems in elasticity, Inverse Probl. 21 (2005) R1–R50. doi:10.1088/0266-5611/21/2/R01.

[45] S. Avril, M. Bonnet, A.S. Bretelle, M. Grédiac, F. Hild, P. Ienny, F. Latourte, D. Lemosse, S. Pagano, E. Pagnacco, F. Pierron, Overview of identification methods of mechanical parameters based on full-field measurements, Exp. Mech. 48 (2008) 381–402. doi:10.1007/s11340-008-9148-y.

[46] M. Ben Azzouna, P. Feissel, P. Villon, Robust identification of elastic properties using the Modified Constitutive Relation Error, Comput. Methods Appl. Mech. Eng. 295 (2015) 196–218. doi:10.1016/j.cma.2015.04.004.

[47] R. Ibañez, E. Abisset-Chavanne, J.V. Aguado, D. Gonzalez, E. Cueto, F. Chinesta, A Manifold Learning Approach to Data-Driven Computational Elasticity and Inelasticity, Arch. Comput. Methods Eng. (2016) 1–11. doi:10.1007/s11831-016-9197-9.

[48] P. Ladevèze, The large time increment method for the analysis of structures with non-linear behavior described by internal variables, COMPTES RENDUS L Acad. DES Sci. Ser. II. 309 (1989) 1095–1099.

[49] J.C. Simo, N. Tarnow, M. Doblare, Non linear dynamics of three dimensional rods: Exact energy and momentum conserving algorithms, Int. J. Numer. Methods Eng. (1995). doi:10.1002/nme.1620380903.

[50] L.T.K. Nguyen, M.-A. Keip, A data-driven approach to nonlinear elasticity, Comput. Struct. 194 (2018) 97–115.

[51] D. González, F. Chinesta, E. Cueto, Thermodynamically consistent data-driven computational mechanics, Contin. Mech. Thermodyn. (2018) 1–15. doi:10.1007/s00161-018-0677-z.

[52] A. Leygue, M. Coret, J. Réthoré, L. Stainier, E. Verron, Data-based derivation of material response, Comput. Methods Appl. Mech. Eng. 331 (2018) 184–196.

[53] J. Ayensa-Jiménez, M.H. Doweidar, J.A. Sanz-Herrera, M. Doblaré, An unsupervised data completion method for physically-based data-driven models, Comput. Methods Appl. Mech. Eng. 344 (2019) 120–143.

[54] J. Ayensa-Jiménez, M.H. Doweidar, J.A. Sanz-Herrera, M. Doblaré, A new reliability-based data-driven approach for noisy experimental data with physical constraints, Comput. Methods Appl. Mech. Eng. 328 (2018) 752–774. doi:10.1016/j.cma.2017.08.027.

[55] Y. Kanno, Simple heuristic for data-driven computational elasticity with material data involving noise and outliers: a local robust regression approach, Jpn. J. Ind. Appl. Math. (2018). doi:10.1007/s13160-018-0323-y.

[56] C.L. Lawson, R.J. Hanson, Solving Least Squares Problems, Society for Industrial and Applied Mathematics, 1987. doi:10.1137/1.9781611971217.

[57] W.K. Liu, S. Jun, Y.F. Zhang, Reproducing kernel particle methods, Int. J. Numer. Methods Fluids. 20 (1995) 1081–1106. doi:10.1002/fld.1650200824.

[58] J.-S. Chen, C. Pan, C.T. Wu, W.K. Liu, Reproducing Kernel Particle Methods for large deformation analysis of non-linear structures, Comput. Methods Appl. Mech. Eng. 139 (1996) 195–227. doi:10.1016/S0045-7825(96)01083-3.

[59] J.-S. Chen, C.-T. Wu, S. Yoon, Y. You, A stabilized conforming nodal integration for Galerkin mesh-free methods, Int. J. Numer. Methods Eng. 50 (2001) 435–466. doi:10.1002/1097-0207(20010120)50:2<435::aid-nme32>3.0.co;2-a.

[60] S. Conti, S. Müller, M. Ortiz, Data-Driven Problems in Elasticity, Arch. Ration. Mech. Anal. 229 (2018) 79–123. doi:10.1007/s00205-017-1214-0.

[61] C.A. Felippa, A survey of parametrized variational principles and applications to computational mechanics, Comput. Methods Appl. Mech. Eng. 113 (1994) 109–139.

doi:10.1016/0045-7825(94)90214-3.

[62] S.T. Roweis, L.K. Saul, Nonlinear dimensionality reduction by locally linear embedding., Science (80-.). 290 (2000) 2323–6. doi:10.1126/science.290.5500.2323.

[63] L.K.L. Saul, S.S.T. Roweis, Think globally, fit locally: unsupervised learning of low dimensional manifolds, J. Mach. Learn. Res. 4 (2003) 119–155. doi:10.1162/153244304322972667.

[64] J.B. Tenenbaum, V. de Silva, J.C. Langford, A global geometric framework for nonlinear dimensionality reduction., Science. 290 (2000) 2319–23. doi:10.1126/science.290.5500.2319.

[65] M. Belkin, P. Niyogi, Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering, Nips. 14 (2001) 585–591.

[66] Z. Zhang, H. Zha, Principal manifolds and nonlinear dimensionality reduction via tangent space alignment, SIAM J. Sci. Comput. 26 (2004) 313–338.

[67] Y. Bengio, J.-F. Paiement, P. Vincent, O. Delalleau, N. Le Roux, M. Ouimet, Out-of-sample extensions for lle, isomap, mds, eigenmaps, and spectral clustering, Adv. Neural Inf. Process. Syst. (2003) 177–184.

[68] E. Vural, C. Guillemot, Out-of-sample generalizations for supervised manifold learning for classification, IEEE Trans. Image Process. 25 (2016) 1410–1424.

[69] H. Cevikalp, W. Triggs, Face recognition based on image sets, in: Comput. Vis. Pattern Recognit., IEEE Conference, 2010: pp. 2567–2573. doi:10.1109/CVPR.2010.5539965.

[70] S. Boyd, L. Vandenberghe, Convex optimization, Cambridge university press, 2004.

[71] A.E. Hoerl, R.W. Kennard, Ridge Regression: Biased Estimation for Nonorthogonal Problems, Technometrics. 12 (1970) 55–67. doi:10.1080/00401706.1970.10488634.

[72] J.A. Tropp, A.C. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, IEEE Trans. Inf. Theory. 53 (2007) 4655–4666. doi:10.1109/TIT.2007.909108.

[73] M. Yaghoobi, M.E. Davies, Fast non-negative orthogonal least squares, in: 2015 23rd Eur. Signal Process. Conf. EUSIPCO 2015, 2015: pp. 479–483. doi:10.1109/EUSIPCO.2015.7362429.

[74] W.S. Cleveland, S.J. Devlin, Locally weighted regression: An approach to regression analysis by local fitting, J. Am. Stat. Assoc. (1988). doi:10.1080/01621459.1988.10478639.

[75] D. Chen, R.J. Plemmons, Nonnegativity constraints in numerical analysis, in: Birth Numer. Anal., World Scientific, 2010: pp. 109–139. doi:10.1142/9789812836267_0008.

[76] P.E. Gill, W. Murray, M.H. Wright, Practical optimization, Emerald Group Publishing Limited, 1981.

[77] T. Belytschko, Y.Y. Lu, L. Gu, Element free Galerkin methods, Int. J. Numer. Methods Eng. 37 (1994) 229–256.

[78] J. Nitsche, Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind, Abhandlungen Aus Dem Math. Semin. Der Univ. Hambg. 36 (1971) 9–15. doi:10.1007/BF02995904.

[79] S. Fernández-Méndez, A. Huerta, Imposing essential boundary conditions in mesh-free methods, Comput. Methods Appl. Mech. Eng. 193 (2004) 1257–1275. doi:10.1016/j.cma.2003.12.019.

[80] J.-S. Chen, H.P. Wang, New boundary condition treatments in meshfree computation of contact problems, Comput. Methods Appl. Mech. Eng. 187 (2000) 441–468.

[81] J.-S. Chen, M. Hillman, S.-W. Chi, Meshfree Methods: Progress Made after 20 Years, J. Eng. Mech. 143 (2017) 04017001. doi:10.1061/(ASCE)EM.1943-7889.0001176.

[82] R. Eggersmann, T. Kirchdoerfer, S. Reese, L. Stainier, M. Ortiz, Model-free data-driven inelasticity, Comput. Methods Appl. Mech. Eng. 350 (2019) 81-99.