SCIENCE CHINA
Physics, Mechanics & Astronomy

• Article •

June 2022 Vol. 65 No. 6: 264611
https://doi.org/10.1007/s11433-021-1840-9

Extracting governing system for the plastic deformation of metallic glasses using machine learning

Liping Yu[1], Xiaoxiang Guo[2], Gang Wang[3], Baoan Sun[4], Dongxue Han[4], Cun Chen[2], Jingli Ren[2*], and Weihua Wang[4*]

[1] *College of Science, Henan University of Technology, Zhengzhou 450001, China;*
[2] *Henan Academy of Big Data, Zhengzhou University, Zhengzhou 450052, China;*
[3] *Institute of Materials, Shanghai University, Shanghai 200444, China;*
[4] *Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China*

Received August 29, 2021; accepted December 22, 2021; published online April 18, 2022

This paper shows hidden information from the plastic deformation of metallic glasses using machine learning. $Ni_{62}Nb_{38}$ (at.%) metallic glass (MG) film and $Zr_{64.13}Cu_{15.75}Al_{10}Ni_{10.12}$ (at.%) BMG, as two model materials, are considered for nano-scratching and compression experiment, respectively. The interconnectedness among variables is probed using correlation analysis. The evolvement mechanism and governing system of plastic deformation are explored by combining dynamical neural networks and sparse identification. The governing system has the same basis function for different experiments, and the coefficient error is $\leq 0.14\%$ under repeated experiments, revealing the intrinsic quality in metallic glasses. Furthermore, the governing system is conducted based on the preceding result to predict the deformation behavior. This shows that the prediction agrees well with the real value for the deformation process.

metallic glasses, sparse identification, dynamical neural networks, correlation analysis

PACS number(s): 81.05.Kf, 81.40.Lm, 07.05.Mh, 02.30.Hq

Citation: L. Yu, X. Guo, G. Wang, B. Sun, D. Han, C. Chen, J. Ren, and W. Wang, Extracting governing system for the plastic deformation of metallic glasses using machine learning, Sci. China-Phys. Mech. Astron. **65**, 264611 (2022), https://doi.org/10.1007/s11433-021-1840-9

1 Introduction

The plastic flow of metallic glasses is reported to initiate from certain loosely packed atomic-scale regions or soft spots [1-3], which are often correlated by long-range elastic interactions. After plastic yielding, the plastic strain is highly localized into shear bands [4, 5]. The shear banding process could profoundly affect their macroscopic mechanical behavior. One direct consequence is that a shear band tends to become runaway with the work softening during deformation [6], resulting in catastrophic failure [7-9]. Therefore, the understanding of the physical mechanism of plasticity serves as a prominent topic for predicting the dynamics of deformation, failure of metallic glasses, and developing applications using these materials.

Previous theoretical analyses of plastic deformation in metallic glasses fall into two categories. First, researchers aim to quantitatively analyze the dynamics using time-series

*Corresponding authors (Jingli Ren, email: renjl@zzu.edu.cn; Weihua Wang, email: whw@iphy.ac.cn)

or statistical analysis such as Lyapunov exponents, multifractals or mean-field theory. The representative dynamical behaviors including self-organized criticality [10-12], fractals [13-17], and chaotic behavior [18-20] are presented. Second, constitutive models are developed to describe material deformation and failure, depending on shear transformation zone (STZ) theory [21-25]. In the STZ theory, plastic deformation occurs when localized regions rearrange because of applied stress; furthermore, the density of these regions is determined by a dynamically evolving effective disorder temperature [3, 26, 27]. Certain factors at STZ can cause local configuration changing. Once the configuration change, the STZ is no longer available for additional transformations in the original shear direction [28].

These pioneering studies increase our understanding of plastic deformation in metallic glasses; however, the quantitative analysis for certain experimental data, generally maybe not be thorough and lack systematization. Furthermore, the constitutive model is macroscopic and may not describe certain specific phenomena in experiments. How to bridge the gap between systematization and particularity? In this study, we intend to balance them by extracting governing systems from experimental data using machine learning (ML) to present systematic and objective description. A mathematical equation that can describe a physical phenomenon or law is known as the governing system of the physical phenomenon or law. Governing system in plastic deformation has different representations as per different conditions.

ML is an effective tool for data analysis [29-31]. Various ML algorithms have been proposed and discussed for the past decades: linear regression [32], logistic regression [33], elastic net [32], ridge regression [34], Bayesian regression [35], decision tree [36], k-nearest neighbors [32], support vector machines [37], sparse representation [32, 38], Gaussian process regression [39], neural networks [40], random forest [32], bagging [41], stacking [42]. Among these methods, sparse identification can accurately show the data indentifying the fewest terms in the dynamic governing systems [43, 44]. Dynamical neural networks is a common approach to learn the black-box models, which reproduce the observable input-output behavior of the target system and reveal internal physical mechanisms [45]. Accordingly, we employ sparse identification and dynamical neural networks to excavate governing systems and evolvement mechanism.

In this study, we consider $Ni_{62}Nb_{38}$ (at.%) metallic glass (MG) film and $Zr_{64.13}Cu_{15.75}Al_{10}Ni_{10.12}$ (at.%) BMG as model materials. First, we determine relevant variables using correlation analysis, for selecting input data in the dynamical neural network. Then, we explore the deformation mechanism by dynamical neural networks and obtain the governing systems corresponding to the black-box using sparse identification. To explore the dynamic variation during deformation, the largest Lyapunov exponent is presented for the preceding experimental data. Our results show that the identified parsimonious models can accurately describe the deformation behavior in nano-scratching and compression experiments. The extracted governing systems have the same basis function in different experiments. The coefficient error is $\leq 0.14\%$ under repeated experiments. Moreover, using governing system to predict the deformation behavior is applicable.

2 Method and results

The $Ni_{62}Nb_{38}$ MG films under unrelaxed glassy states were deposited on a cleaned monocrystalline silicon wafer with a size of 10 mm × 10 mm × 0.5 mm in a high-vacuum chamber (a base pressure of 5×10^{-5} mbar (1 mbar = 10^2 Pa)) with a load-lock system, and a probe handler using a direct current (DC) planar magnetron sputtering device (Shenyang ZKY Technology Development Co., JGP-450). The thickness of the $Ni_{62}Nb_{38}$ MG film was ~ 2 μm, obtained from a cross-section view observed by high-resolution scanning electron microscopy (SEM) (Phenom world, Phenom Proxy). Nano-scratching experiments were conducted in a Triboindentor TI-900 machine (Hysitron, Inc.). The nanoscratch indentor was a conic diamond with a tip radius of 1 mm and a half-angle of $\pi/6$. The scratch length was 20 mm, and the moving speed of the nanoindenter was 2 mm/s. For $Ni_{62}Nb_{38}$ (at.%) MG, nano-scratching with a scratch time of 5 s, loading forces of 2000 and 500 μN were performed. At each loading force, at least three scratches were repeatedly performed. Under repeated experiments, the scratches were conducted on a single sample in different locations.

Figure 1(a)-(d) show signals with a scratch time of 5 s, loading force of 2000, and 500 μN in horizontal and vertical directions, including the lateral force (LF), lateral displacement (LD), normal force (NF), and normal displacement (ND)-time, respectively. The nanoscratching has not really started between 0 to 12 s, which is preparatory work before the experiment. Then, the nanoscratching process, which is the subject of investigation, occurs from the time duration from 12 to 17 s. Therefore, the time axis starts from 12 to 17 s in Figure 1(a)-(d). Note that the curve of LF versus time comprises several serration events. Each serration event includes force-increasing and force-decreasing processes. In the force-increasing process, the structural strengthening in the atomic scale occurs, thereby facilitating the accommodation of the plastic shear strain. However, the force-decreasing process corresponds to a force relaxation and usually causes a shear separation of the MG surface.

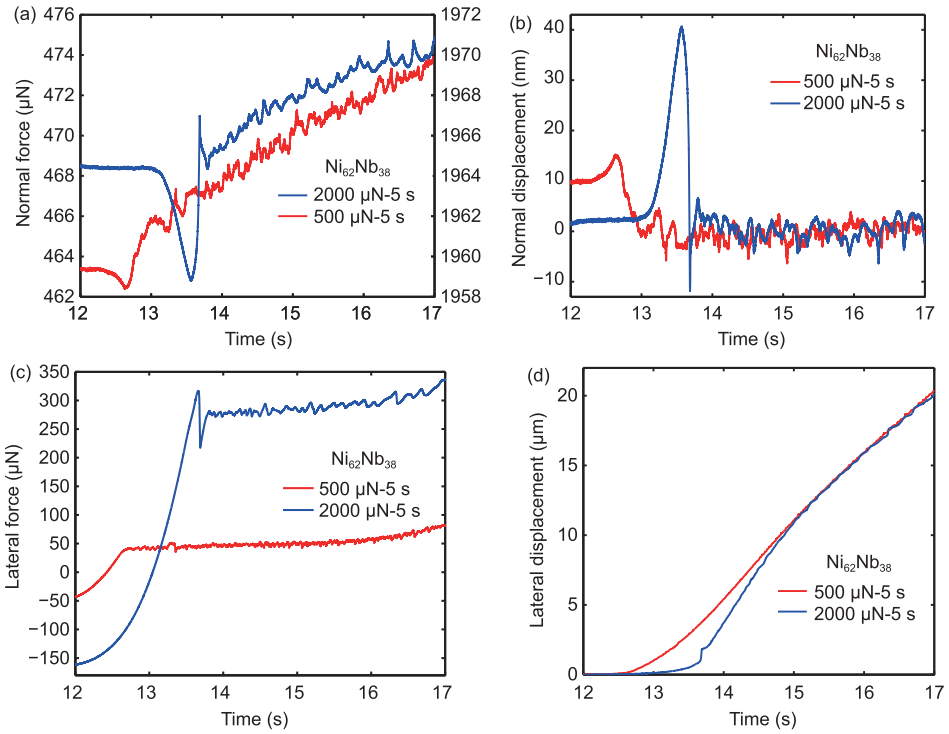

Figure 1 (Color online) (a), (b) Normal force and displacement as functions of scratch time under loading forces of 2000 and 500 μN with a scratch time of 5 s; (c), (d) lateral force and displacement as functions of scratch time under loading forces of 2000 and 500 μN with a scratch time of 5 s.

2.1 Variable correlation analysis

A correlation analysis is conducted to select input data in the dynamical neural networks and sparse identification. As discussed in ref. [16], stick-slip behavior reflects a surface damage process during friction associated with an intermittent deformation process underlying LF. The investigation of LF during the nanoscratching process may provide valuable information to understand the damage process of metals; therefore, the primary objective of this study is LF. Assume that noncontinuous deformation can be expressed by a dynamic system

$$Z(k+1) = f(Z(k); P), \tag{1}$$

where $Z(k) = (z_1(k), \ldots, z_n(k))$ is observed data, i.e., time-series data of deformation, which are variables describing the dynamical state of the system; P contains parameters representing variable factors, such as temperature, deformation time, or strain rate, which drive the deformation from one state to another; and $f = (f_1, \ldots, f_n)$ comprises generally of nonlinear functions of $Z(k)$ with fixed point $Z = \overline{Z}$. The Pearson correlation coefficient among variables is given by

$$\mathrm{PCC}(z_i, z_j) = \frac{\mathrm{Cov}(z_i, z_j)}{\sqrt{\mathrm{Var}(z_i)\mathrm{Var}(z_j)}},$$

where $\mathrm{Var}(z_i) = \mathrm{Cov}(z_i, z_i)$, $1 \le i, j \le n$.

Figure 2(a) shows the correlation for each pair among LF, LD, NF, and ND as a function of time during the nanoscratching process under 500 μN-5 s. Then, we present a correlation matrix among four variables to obtain specific relation. From the correlation matrix shown in Figure 2(b), we obtain that LD or NF has a high correlation with LF. This indicates that, LD or NF can be selected as input data if we consider LF as training data using dynamical neural networks.

Figure 2(c) and (d) show the correlation analysis for each pair among four indicators under 2000 μN-5 s. As per the mutual correlation in Figure 2(d), we obtain that LF has high correlation with LD. This indicates that, LD can be selected if we take LF as training data. Moreover, we obtain that the absolute value of correlation for each pair of variables (|PCC(LF,LD)|, |PCC(LF,NF)|, |PCC(LF,ND)|, |PCC(LD,NF)|, |PCC(LF,LD)|, and |PCC(NF,ND)|) drastically increases to 1 before plastic deformation in Figure 2(a) and (c). We will explain this phenomenon as follows.

Assuming that there is a value P_c in the system (1), such that at least one of the eigenvalues of the Jacobian matrix $\frac{\partial f(Z,P_c)}{\partial Z}|_{Z=\overline{Z}}$ at $Z = \overline{Z}$, which characterizes the change rate of the orbit of a system around fixed points, equals 1 in modulus or absolute value. This indicates that the system undergoes Neimark-Sacker bifurcation, flip bifurcation, or fold bifurcation at fixed points when P reaches the critical value P_c. The transition from elastic deformation to plastic deformation can

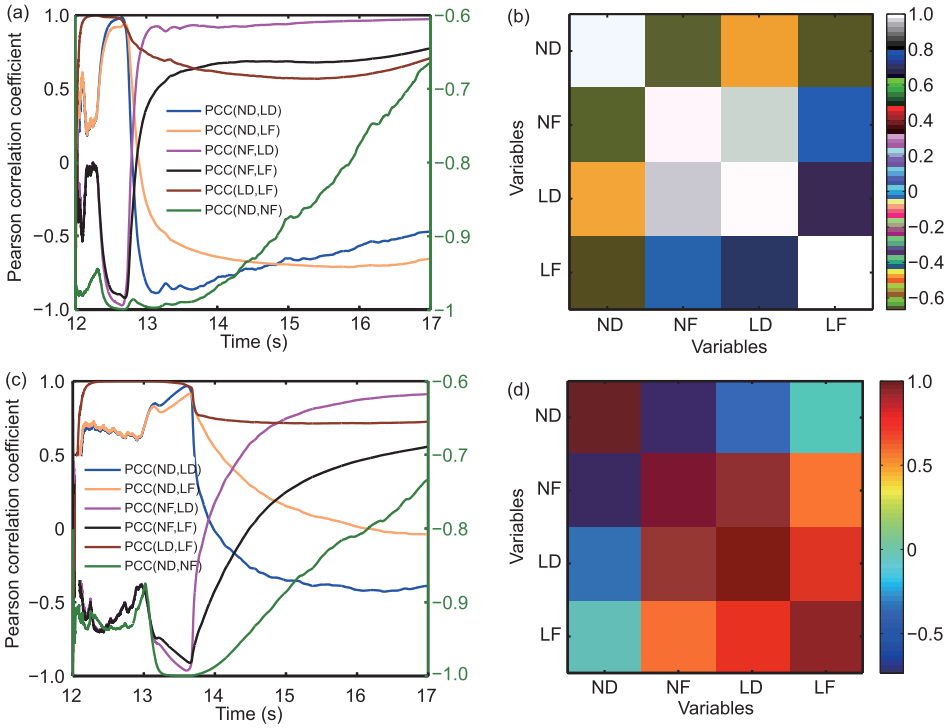

Figure 2 (Color online) Curves of Pearson correlation coefficient for the variables against time and correlation matrix with a scratch time of 5 s, loading force of 500 μN ((a), (b)), and loading force of 2000 μN ((c), (d)). The green right y-axis corresponds to the correlation between ND and NF. (b), (d) The matrix confirms that LF has a positive correlation with LD, NF, and a negative correlation with ND.

be viewed as a bifurcation in dynamic system. Following the analyses [46,47], the dynamic system (1) can be linearized at $Z = \overline{Z}$ as follows:

$$Z(k+1) = \overline{Z} + \frac{\partial f(Z, P_c)}{\partial Z}|_{Z=\overline{Z}}(Z(k) - \overline{Z}) + o(Z(k) - \overline{Z}). \quad (2)$$

If we introduce $Y(k) = S^{-1}(Z(k) - \overline{Z})$, of which the transformation S comprises the corresponding eigenvectors, the system can be transformed to

$$Y(k+1) = \Lambda(P)Y(k) + \xi(k), \quad (3)$$

where $\Lambda(P)$ is the diagonalizable matrix of $\frac{\partial f(Z,P_c)}{\partial Z}|_{Z=\overline{Z}}$, and $\xi(k) = (\xi_1(k), \ldots, \xi_n(k))$ are Gaussian noise with zero means and covariances $k_{ij} = \mathrm{Cov}(\xi_i, \xi_j)$. Here, $\Lambda(P) = \mathrm{diag}(\lambda_1(P), \ldots, \lambda_n(P))$, with each $|\lambda_i|$ between 0 and 1 corresponding to the stable point. With the parameter P tending to be P_c, if the largest eigenvalue in modulus, say λ_1, which we call the dominant eigenvalue, approaches 1, and then bifurcation occurs.

Without loss in generality, suppose that $Y(k)$ has zero mean, and the covariance and Pearson correlation coefficient are given by [46,47]

$$\mathrm{Cov}(y_i, y_j) = \frac{k_{ij}}{1 - \lambda_i\lambda_j},$$

$$\mathrm{PCC}(y_i, y_j) = \frac{k_{ij}}{\sqrt{k_{ii}k_{jj}}} \frac{\sqrt{(1-\lambda_i^2)(1-\lambda_j^2)}}{1 - \lambda_i\lambda_j}.$$

Note that $\mathrm{Var}(y_1) = \mathrm{Cov}(y_1, y_1) \to +\infty$ for the eigenvalue $|\lambda_1| \to 1$, while $\mathrm{Cov}(y_i, y_j)$ for $i \neq j$ and $\mathrm{Var}(y_i)$ for $i \neq 1$ show no drastic change (bounded). $\mathrm{PCC}(y_1, y_j) \to 0$ for any $j \neq 1$ when $|\lambda_1| \to 1$, i.e., the variable y_1 is related to the dominant eigenvalue λ_1.

For variables Z whose values are directly measured by $z_i = \sum_{j=1}^{n} s_{ij}y_j + \overline{z_i}$, where s_{ij} is the (i, j) element of linear transformation S. The covariance trend can be obtained when the parameter P tends to P_c using the following expressions

$$\mathrm{Cov}(z_i, z_j) = s_{i1}s_{j1}\mathrm{Var}(y_1) + \cdots + s_{in}s_{jn}\mathrm{Var}(y_n) + \sum_{k,m=1,k\neq m}^{n} \mathrm{Cov}(y_k, y_m),$$

$$\mathrm{PCC}(z_i, z_j) = \frac{\mathrm{Cov}(z_i, z_j)}{\sqrt{\mathrm{Var}(z_i)\mathrm{Var}(z_j)}},$$

where $\mathrm{Var}(z_i) = \mathrm{Cov}(z_i, z_i)$, $1 \leq i, j \leq n$. With the results from the space Y, we can obtain that if z_i and z_j are related to y_1, so s_{i1} and s_{j1} are nonzero, $\mathrm{Cov}(z_i, z_j)$ will tend to ∞ and $|\mathrm{PCC}(z_i, z_j)| \to 1$ drastically if $|\lambda_1| \to 1$ on the observable space Z. This is the reason that the absolute value of correlation drastically increases to 1 before plastic deformation in

Figure 2(a) and (c), which may be considered as a predicting index for discussion in future.

2.2 Dynamical neural networks

Before learning the black-box models, a reconstructed phase space is necessary. As described by Strogatz [48], an unknown dynamic system has an equivalent topological property with the model in the reconstructed phase space. The basic idea is to unfold dynamics through a phase space reconstruction by embedding the time series in a higher-dimensional space using a suitable time delay [49, 50]. Assume the original one-dimensional signal, denoted as $x(i), i = 1, 2, \ldots, N$, the point of the reconstructed phase space, $Y(i, j)$, is

$$\begin{pmatrix} x(1) & x(2) & \cdots & x(N-(m-1)\tau) \\ x(1+\tau) & x(2+\tau) & \cdots & x(N-(m-2)\tau) \\ \vdots & \vdots & \ddots & \vdots \\ x(1+(m-1)\tau) & x(2+(m-1)\tau) & \cdots & x(N) \end{pmatrix},$$

where N is the length of time series; τ is the time delay, calculated using the mutual information method [51]; m is embedding dimension using the Cao method [52]. The number of points in the reconstructed phase space is $N-(m-1)\tau$.

Considering the increasing trend in the correlation curves, we present the largest Lyapunov exponent [53],

$$\lambda_1 = \frac{1}{t_M - t_0} \sum_{k=1}^{M} \ln \frac{L'(t_k)}{L(t_{k-1})},$$

where $L(t_{k-1})$ or $L'(t_k)$ is the Euclidian module between $Y(i, j)$ and its nearest neighbor point in the reconstructed phase space, see details in ref. [25]. The largest Lyapunov exponent, λ_1, is a parameter to characterize the rate of the divergence of the trajectories in the phase space. A positive value of λ shows chaotic dynamic behavior, whereas, a negative value of λ shows a stable dynamic behavior. Figure 3 shows the largest Lyapunov exponent as a function of deformation time. As shown in Figure 3(a) and (b), the Lyapunov exponent has a large variation (drastically increases to 0) before plastic deformation under 500 μN-5 s and 2000 μN-5 s, indicating the undulation of dynamic behavior in a pre-transition stage. The fluctuation of the largest Lyapunov exponent agrees with the trend of Pearson correlation coefficient before plastic deformation in Figure 2(a) and (c), which can be combined as a predicting index.

Combining the preceding correlation analysis and phase space reconstruction, we use dynamical neural networks to train data, and identify a black-box model. A two-layer feed-forward network with sigmoid hidden neurons and linear output neurons can arbitrarily fit multidimensional mapping problems, given the consistent data and a sufficient number of neurons in its hidden layer. The input matrix $\mathbf{P}$ is constituted by an m-dimensional observational variable

$$\mathbf{P} = \begin{pmatrix} x_1(1) & x_1(2) & \cdots & x_1(N) \\ x_2(1) & x_2(2) & \cdots & x_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ x_m(1) & x_m(2) & \cdots & x_m(N) \end{pmatrix}.$$

Moreover, the target matrix $\mathbf{Q}$ is in the reconstructed phase space

$$\mathbf{Q} = \begin{pmatrix} x_k(1) & \cdots & x_k(M) \\ x_k(1+\tau) & \cdots & x_k(M+\tau) \\ \vdots & \ddots & \vdots \\ x_k(1+(m-1)\tau) & \cdots & x_k(M+(m-1)\tau) \end{pmatrix},$$

where $M = N-(m-1)\tau$, $k = 1, 2, \ldots, m$. Generally, 70% of

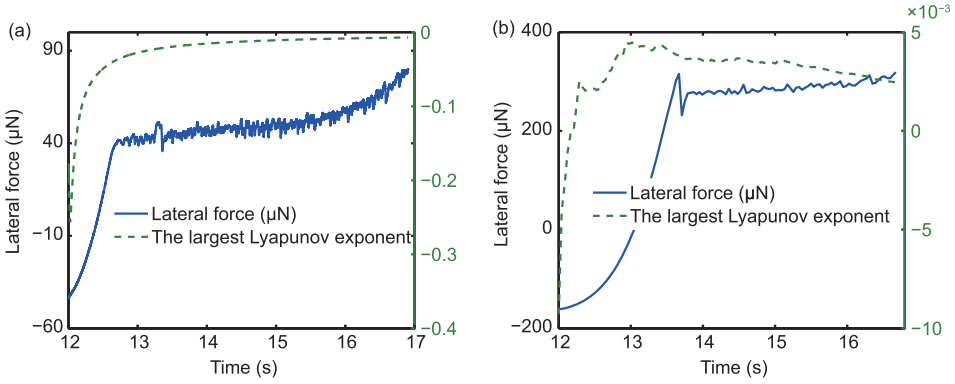

Figure 3 (Color online) Largest Lyapunov exponent as a function of deformation time for $Ni_{62}Nb_{38}$ MG film during a nanoscratch process under (a) 500 μN-5 s and (b) 2000 μN-5 s. The green right y-axis corresponds to the largest Lyapunov exponent. The largest Lyapunov exponent goes from negative to positive in (a) and (b) before the transition of plastic deformation.

the samples are used to train the network; 15% of the samples are used to validate the trained network; and the remaining 15% of samples are used to test the network.

As per the correlation analysis, we set LF and LD as input vectors, and reconstructed LF as target vector to confirm the correlation. The number of samples points is 15000, and the corresponding algorithm is presented in Algorithm 1.

Figure 4(a) and (b), compare testing data and real data to estimate the efficiency of dynamical neural networks under 500 μN-5 s and 2000 μN-5 s (the blue line is real data; the red line is testing data. The subfigure is a magnification of the region in the square box). To obtain additional evidence about testing, Figure 4 (c) and (d) show group of new experimental data (the second experiment) under the same condition with 50% of samples used to train the network; and 25% of samples for testing. Moreover, the mean squared error (MSE) of the variable, x_k is considered as the criteria standards to evaluate the validity of the algorithm,

$$\mathrm{MSE}(x_k) = \frac{1}{K}\sum_{j=1}^{K}(x_k(M+j) - \widehat{x_k}(M+j))^2,$$

where $x_k(M+j)$ is the original value, and $\widehat{x_k}(M+j)$ is the

Algorithm 1 Algorithm for dynamical neural networks

Input: The combined variable of lateral displacement and lateral force, $X(i), i = 1, 2, \ldots, N$;

Output: The training, validation and testing of $Y(i, j)$;

1. Reconstruct the variable of lateral force $y(i), i = 1, 2, \ldots, N$, denote as $Y(i, j)$;

2. Data normalization, [Y_1,Y_s]=mapminmax(Y), [X_1,X_s]=mapminmax(X);

3. Dividing data, [trainsample.$Y(X)$, valsample.$Y(X)$, testsample.$Y(X)$]
=dividerand($Y(X)$,0.7,0.15,0.15);

4. Create network, net=newff(minmax(Y),[10,1],tansig purelin,"traingdm");

5. Set network parameters, net.trainParam.epochs=1000;
net.trainParam.goal=1×10^{-7}; net.trainParam.lr=0.01; net.trainParam.mc=0.9;

6. Train network, net.trainFcn="trainlm", [net,tr]=train(net,trainsample.Y,trainsample.X);

7. Training result, validation and testing, [normoutput,Perf]=sim(net,sample.Y,[],[],sample.X);

8. Reversely normalized, output=mapminmax("reverse", normoutput,Y_s).

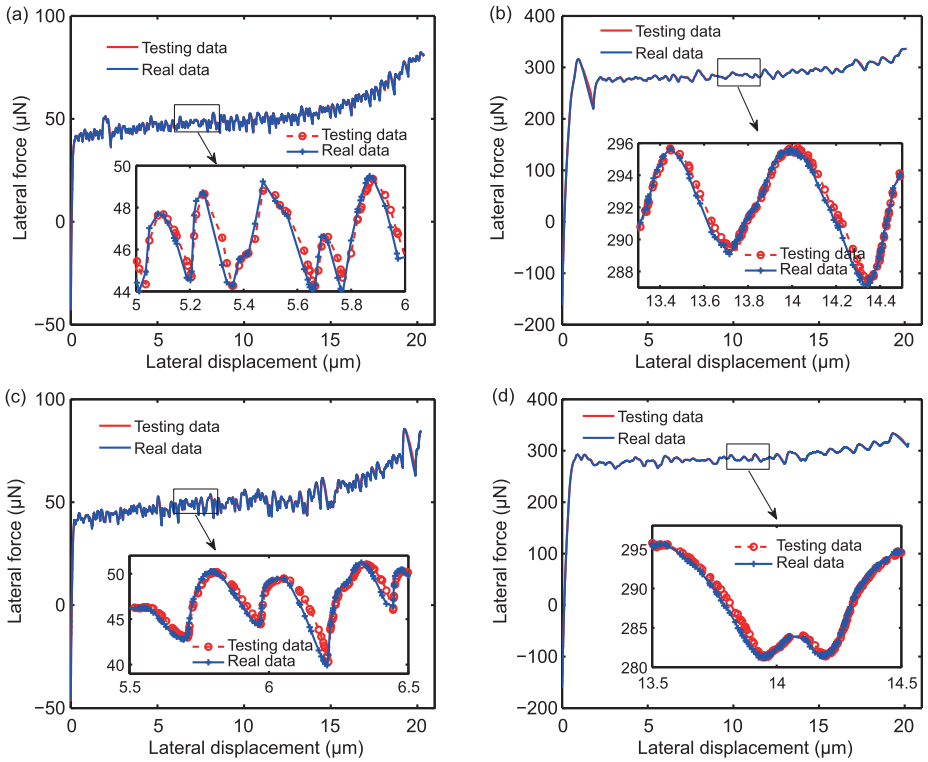

Figure 4 (Color online) Explore the physical mechanism of deformation process by dynamical neural networks. Performance of LF for $Ni_{62}Nb_{38}$ MG film under (a) and (c) 500 μN-5 s with 15% and 25% of samples used for testing, (b) and (d) 2000 μN-5 s with 15% and 25% of samples used for testing. The subfigure is a magnification of the region in the square box.

testing data. Table 1 lists the errors under the condition of 15% and 25% of samples for testing. The smaller the error value, the better the algorithm effect. The results show that neural networks can train and test the deformation data. Furthermore, with the augment of the proportion of samples for testing, we obtain that MSE decreases under 2000 μN; however, it increases under 500 μN. This elucidates the deformation during the nanoscratching process to be stochastic.

2.3 Sparse identification

Generally, a dynamic system has the following form:

$$\dot{\mathbf{X}}(t) = \mathbf{f}(\mathbf{X}(t)), \tag{4}$$

where $\mathbf{X}(t) = (\mathbf{x}_1(t), \mathbf{x}_2(t), \ldots, \mathbf{x}_m(t))$, $\mathbf{x}_k(t) \in \mathbb{R}^n$ are the state of a system at time t, and the function $\mathbf{f}(\mathbf{X}(t))$ is the dynamic constraints that define the equations of motion of the system. As discussed in ref. [43], if we rewrite the system as:

$$\dot{\mathbf{X}}(t) = \mathbf{f}(\mathbf{X}(t)) = \Theta(\mathbf{X})\Xi, \tag{5}$$

where $\Xi = \{\xi_1, \xi_2, \ldots, \xi_n\}$ is the sparse matrix of coefficients, and $\Theta(\mathbf{X})$ is a basic library function with the combination of $\mathbf{X}$, including constant, polynomial, and trigonometric terms,

$$\Theta(\mathbf{X}) = \begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \\ 1 & \mathbf{X} & P(\mathbf{X}) & \sin(\mathbf{X}) & \cos(\mathbf{X}) & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{pmatrix}.$$

$P(\mathbf{X})$ is the polynomial of $\mathbf{X}$, e.g., quadratic polynomial can have C_n^2 combination

$$P(\mathbf{X}) = \{\mathbf{x}_1^2(t), \mathbf{x}_1(t)\mathbf{x}_2(t), \ldots, \mathbf{x}_n(t)\mathbf{x}_{n-1}(t), \mathbf{x}_n^2(t)\}.$$

Our aim is to formulate appropriate $\Theta(\mathbf{X})$ and sparse matrices Ξ. As per the phase space reconstruction theorem [54], the evolution of variate $(\mathbf{x}_k(t), \mathbf{x}_k(t+\tau), \mathbf{x}_k(t+2\tau), \cdots), k = 1, 2, \ldots, m$, is topology equivalent to the dynamic evolution of system. Unlike considering the derivative of the variable, we define operator $\mathfrak{R}$ to implement the variate phase space reconstruction. Then, the left-hand side derivative of system (4) can be replaced by the operator, $\mathfrak{R}(\mathbf{x}_k) = \mathbf{f}(\mathbf{x}_k)$, to discover the governing systems from a black-box. The related algorithms are presented in Algorithm 2.

Table 1 MSE of dynamical neural networks under 500 and 2000 μN-5 s

MSE (μN)	Training	Validation	Testing
500 (15%)	1.2255×10^{-1}	1.1744×10^{-1}	1.1266×10^{-1}
2000 (15%)	2.4581×10^{-1}	3.0303×10^{-1}	2.9091×10^{-1}
500 (25%)	1.4472×10^{-1}	1.4140×10^{-1}	1.4046×10^{-1}
2000 (25%)	9.4809×10^{-2}	9.3158×10^{-2}	9.9836×10^{-2}

Algorithm 2 Sparse identification

Input: State variable space, X, Sparsity threshold, λ;
Output: Sparse coefficient matrix, Ξ;
1. Solve the numerical derivative, $\mathrm{d}X/\mathrm{d}t$ based on X;
2. Construct the base function library, Θ based on X;
3. Find the sparse matrix by the least squares method, Ξ;
4. **for** $k = 1 : 10$ **do**
5. Find small coefficient, smallinds=(abs(Ξ) < λ);
6. Assign the small coefficient to 0, Ξ(smallinds) = 0;
7. **for** ind = $1 : n$, n is state space dimension **do**
8. Biginds =~ smallinds(:, inds);
9. Do the regression for the remaining terms and continue to solve the sparse matrix, Ξ(biginds, ind) = Θ(:, biginds) $\mathrm{d}X/\mathrm{d}t$(:, ind);
10. **end for**
11. **end for**
12. Return sparse matrix, Ξ.

In this algorithm, we first use the least squares algorithm to preliminarily estimate the sparse matrix Ξ. Then, we look for the small coefficient in the sparse matrix to determine whether the small coefficient is less than the sparsity threshold λ, and assign these coefficients to zero. The remaining non-zero coefficient matrix is substituted back into the original problem to form a new sparse solution problem. Then, the sparsity threshold condition λ is continuously used for filtering until all non-zero coefficients converge. The sparsity threshold condition is helping for solving sparsity, making it possible to quickly converge to the sparse solution after certain iteration steps.

In our study, for $Ni_{62}Nb_{38}$ in the nano-scratch process under 500 and 2000 μN-5 s, we attempt to extract a two-dimensional system, where $\mathbf{x}_1$ represents LD and $\mathbf{x}_2$ is LF. We consider reconstructed LF as input to obtain Ξ using the least squares method. Table 2 shows the corresponding 2-dimensional identify systems. As shown in Table 2, Systems 1 and 2 represent 500 μN-5 s and 2000 μN-5 s, respectively; moreover, $\mathfrak{R}_1$ and $\mathfrak{R}_2$ are the columns of the reconstructed LF in the phase space. The difference between columns of the governing system can be equally as one type of derivative, thus revealing an inner connection between plastic deformation and LF rate. This is consistent with the conclusion of ref. [17], which shows the chaotic attractor by LF and LF rate. Table 2 shows the system coefficient errors for repeated experiments. The fluctuation of the system coefficient is $\leq 0.14\%$ in repeated experiments under the same conditions, thus comfirming that coefficient is stable and repeatable. Figure 5 shows the curves of the identified system and original data. The blue line corresponds to the original data, and the red-dotted line corresponds to the model data. This shows that, the model data agree with the real value for the

Table 2 Sparse coefficients of identified systems with LF and LD as input vectors, and reconstructed LF as output vector. Systems 1 and 2 represent 500 μN-5 s and 2000 μN-5 s, respectively. The error from the repeated experiments under the same condition

Element	System 1		System 2	
	$\Re_1$	$\Re_2$	$\Re_1$	$\Re_2$
1	0	$0.3662\pm4\times10^{-5}$	0	$0.4812\pm4\times10^{-5}$
$\mathbf{x}_1$	0	$-0.0423\pm1\times10^{-5}$	0	$0.1007\pm1.4\times10^{-3}$
$\mathbf{x}_2$	1	$1.0000\pm8\times10^{-8}$	1	$1.0011\pm3\times10^{-9}$
$\mathbf{x}_1^2$	0	$0.0007\pm1\times10^{-5}$	0	$0.0093\pm4\times10^{-8}$
$\mathbf{x}_1\mathbf{x}_2$	0	$0.0008\pm1\times10^{-5}$	0	$-0.0010\pm1\times10^{-8}$
$\mathbf{x}_2^2$	0	-0.0002 ± 0	0	0.0000 ± 0

deformation process. The results demonstrate that a simple polynomial can describe the deformation dynamics. The concise models confirm that there exists a governing system for training data of neural networks that can capture system dynamics.

Based on the governing system, we can determine the evolution of LF over time. The curve of LF versus time comprises serration events induced by the stick-slip movements. From the system coefficient, we obtain that the combination of LF and LD (i.e., the right-hand side of the system) positively influences the serration events during the nanoscratching process under a loading force of 500 μN. Moreover, it has a negative impact under a loading force of 2000 μN. Furthermore, along with refs. [16, 55], shear branching is the main mechanism that dominates the scratching process of the nanoindenter, which forms a shear-band-structure-like state. The single-shear band sliding in repetitive patterns leads to considerable ductility, and the simultaneous nucleation of numerous shear bands and their interactions contribute to plasticity. Thus, controlling the governing system parameters may have the same effect on plasticity with changing extrinsic conditions.

As discussed in ref. [16], increasing the loading force gradually increses the amount of deformation on the surface underneath the nanoindenter tip with increasing indentation depth. Multiple clusters with low shear modulus independently develop in the contact area between the nanoindenter tip and surface, which are the fundamental units of shear deformation in MG. At a higher loading force, the motion of LF corresponds to the operation of shear branching during the nanoscratching process. At a smaller loading force, the adsorption of surface atoms on the nanoindenter tip leads to certain forces that influence the LF. The local properties of LF at a smaller loading force are more complex than those at a higher loading force, which is demonstrated by the results in Table 2. This shows that the system under the loading force of 500 μN has additional nonlinearity than that under the loading force of 2000 μN.

Using the governing system, we obtain an obvious correlation between LF and LD used to compare with certain experimental phenomena. For example, we may speculate shear-separation activity in the stick-slip behavior by continuously observing the fluctuating amplitude of the LF rate from the governing system. The larger fluctuating amplitude of the LF rate, shows a more significant shear-separation activity in the stick-slip behavior. Moreover, to confirm the basis function, we consider $Zr_{64.13}Cu_{15.75}Al_{10}Ni_{10.12}$ (at.%) BMG compressed at the temperature 183 K with a strain rate of 2.5×10^{-4} s^{-1} to test the basis function. We conducted compressive tests using a legacy dynamics and fatigue system model 8652 electric actuator (Instron) equipped with an environmental box. To exclude the influence of data-acquisition frequency on stress fluctuation sensitivity, we selected data acquisition frequencies of 0.5, 1.0, 10.0, and 100.0 points/s with an increase in strain rates.

Figure 6(a) shows the compressive nominal stress-time curve. Each stress drop is associated with the nucleation of a band of localized plastic deformation, which propagates along the sample under certain conditions. After the stress in-

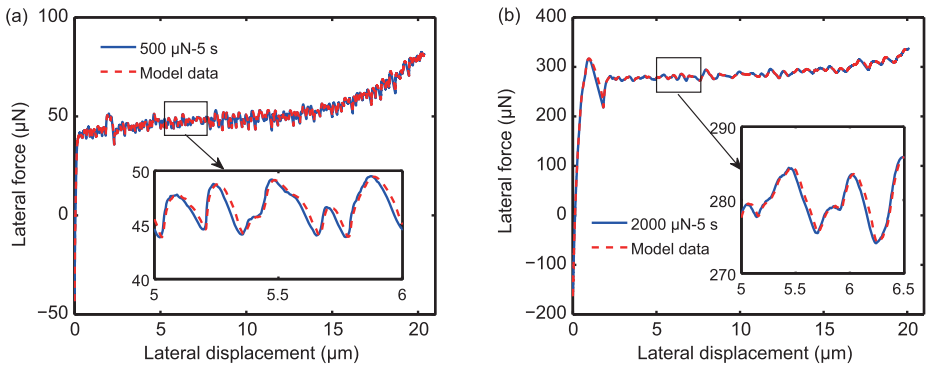

Figure 5 (Color online) Performance of the sparse identification of nonlinear dynamic algorithm. $Ni_{62}Nb_{38}$ MG film under (a) 500 μN-5 s, and (b) 2000 μN-5 s. The blue line corresponds to the real data, and the red-dotted line corresponds to the model data. The subfigure is a magnification of the region in the square box.

creases beyond the yield stress, the plastic deformation commences. Figure 6(b) and (c) show the correlation analysis between strain, stress, stress rate and deformation energy. As shown in Figure 6(b) and (c), the strain or energy is relatively closely connected with stress. This indicates strain or energy can be selected as input if we consider stress as the training data using dynamical neural networks. Figure 7(a) shows the testing result by considering the normal stress and normal strain as input and reconstructed stress as output vector. The blue line corresponds to the real data and the red line corresponds to the testing data. In this process, the number of the sample points is 19189, and the MSE of training, validation, and testing are 7.2513×10^{-2}, 6.9528×10^{-2}, 7.4081×10^{-2} respectively. If 50% of the samples are used to train the network and 25% of the samples are applied to test the data, the MSE reduces instead (6.6648×10^{-2}, 6.8206×10^{-2}, 6.9650×10^{-2}). Table 3 presents sparse coefficients of identified systems for $Zr_{64.13}Cu_{15.75}Al_{10}Ni_{10.12}$ (at.%) BMG, where $\mathbf{x}_1$ represents the normal strain in System 1 or energy in System 2, and $\mathbf{x}_2$ is the normal stress. Figure 7(b) shows the performance of the sparse identification of nonlinear dynamics algorithm. The subfigure is a magnification of the region in the square box. This indicates that the simple polynomial can depict deformation dynamics in compression experiments.

The MSE of the variable between real data and model data from sparse identification under three conditions (500 μN, 2000 μN and 183 K, 2.5×10^{-4} s^{-1}) are 2.439×10^{-1}, 7.578×10^{-1}, 3.966×10^{-1}, respectively. The results show that simple polynomial can have high accuracy and reliability for

Table 3 Sparse coefficients of identified systems for $Zr_{64.13}Cu_{15.75}Al_{10}Ni_{10.12}$ (at.%) BMG, where $\mathbf{x}_1$ represents the normal strain in System 1 or energy in System 2, and $\mathbf{x}_2$ is the normal stress

Element	System 1		System 2	
	$\mathfrak{R}_1$	$\mathfrak{R}_2$	$\mathfrak{R}_1$	$\mathfrak{R}_2$
1	0	–0.8530	0	0.5147
$\mathbf{x}_1$	0	2.5454	0	–0.0163
$\mathbf{x}_2$	1	0.9894	1	0.9995
$\mathbf{x}_1^2$	0	–0.2555	0	0.0000
$\mathbf{x}_1\mathbf{x}_2$	0	0.0011	0	0.0001
$\mathbf{x}_2^2$	0	0.0000	0	0.0000

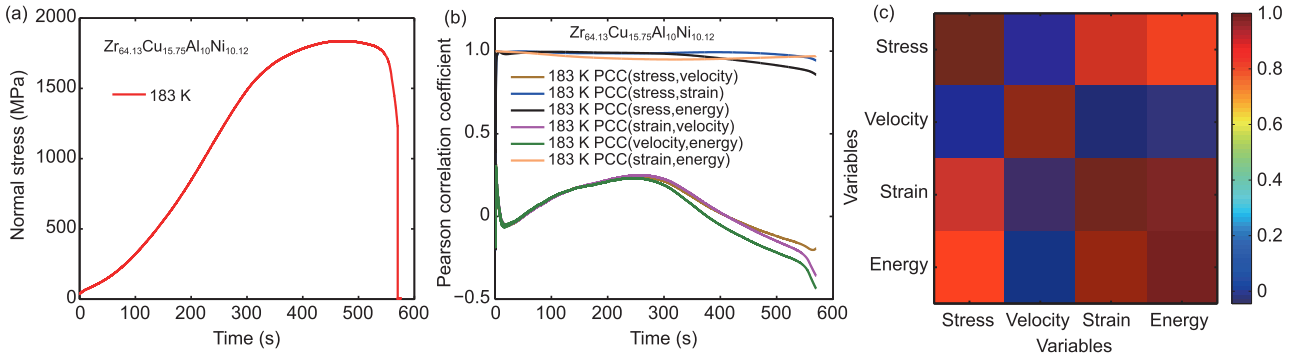

Figure 6 (Color online) (a) Stress-time curves for $Zr_{64.13}Cu_{15.75}Al_{10}Ni_{10.12}$ BMG deformed at 183 K with strain rate of 2.5×10^{-4} s^{-1}. (b), (c) Curves of Pearson correlation coefficient for the variables against time, and the correlation matrix, which clearly show the relationship between stress, strain, energy, and stress rate (denote as velocity in the figure).

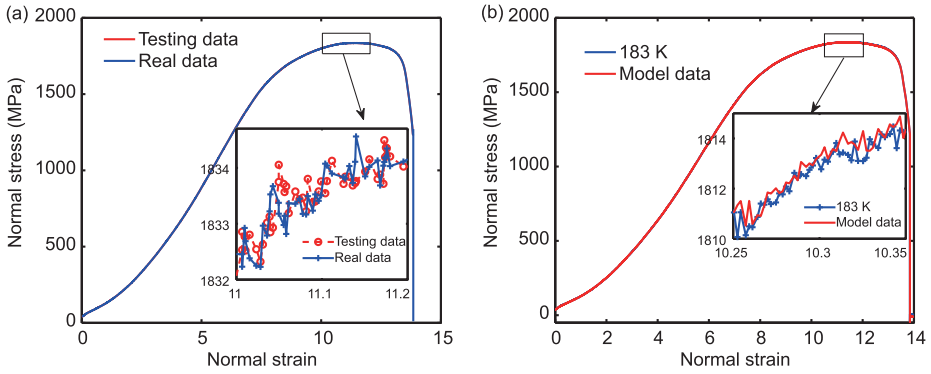

Figure 7 (Color online) $Zr_{64.13}Cu_{15.75}Al_{10}Ni_{10.12}$ BMG deformed at 183 K with strain rate of 2.5×10^{-4} s^{-1}. (a) Explore physical mechanism of deformation process by dynamical neural networks; (b) the performance of the sparse identification of nonlinear dynamic algorithm. The subfigure is a magnification of the region in the square box.

describing deformation dynamics. Naturally, if we consider all variables as input to obtain a 4-dimensional governing system, the system will be more complex with higher accuracy error. This method provides a principled approach to maximally leverage the data to capture the hidden system.

2.4 Prediction by governing system

In this section, by considering $Ni_{62}Nb_{38}$ (at.%) MG film as an example, we employ the governing system to predict the deformation process under 2000 μN-5 s and 500 μN-5 s. Denote the variable for LF and LD as $x(i)$ and $y(i)$, where $i = 1, 2, 3, \ldots, N$, N is the length of the data. We take $x(id : id + \tau - 1)$ and $y(id : id + \tau - 1)$ as initial value where τ is the time delay of the reconstructed phase space. Then, we substitute the initial value into the loop as per the governing system to obtain a new value and store it (see the corresponding algorithm in Algorithm 3). During the calculation for 2000 μN-5 s, *id* and step are considered as 2000 and 5600, respectively, in Figure 8(a). The red line corresponds to the real data, and the blue line corresponds to the predicted data from the governing equation. The prediction is consistent with the real value at the beginning, and then the curve emerges from separation as time goes on. This is because the largest Lyapunov exponent goes from negative to positive with time. The system dynamics behavior is unstable, stochastic, and sensitive to the variation of external conditions. As shown in Figure 8(b), we move the initial value back to $x(9712 : 9712 + \tau - 1)$ and $y(9712 : 9712 + \tau - 1)$, and present the prediction for the following points. We obtain that the trend of the predicted data is consistent with the trend of the real data, and the predicted value agrees with the real value.

Algorithm 3 Predicted algorithm by governing system

Input: The variable for the lateral force and lateral displacement, $x(i), y(i), i = 1, 2, 3, \ldots, N$ and sparse matrix, Ξ;

Output: Return predicted values;

1. Fix the initial value, $yb = y(id : id + \tau - 1, :)$, $xb = x(id : id + \tau - 1, :)$ and circulation step;

4. **for** $k = 1$: step **do**

5. tmp = $[xb, yb]$;

6. Θ = poolData(tmp, r, polyorder, usesine);

7. $yy = \Theta * \Xi$;

8. pre = $yy(:, 2)$;

9. $yb = [yb(2 : \text{end}); \text{pre}]$, $xb = [xb(2 : \text{end}); x(id + \tau - 1 + k, :)]$;

10. Pred = [Pred;pre];

11. **end for**

12. Return predicted values, Pred.

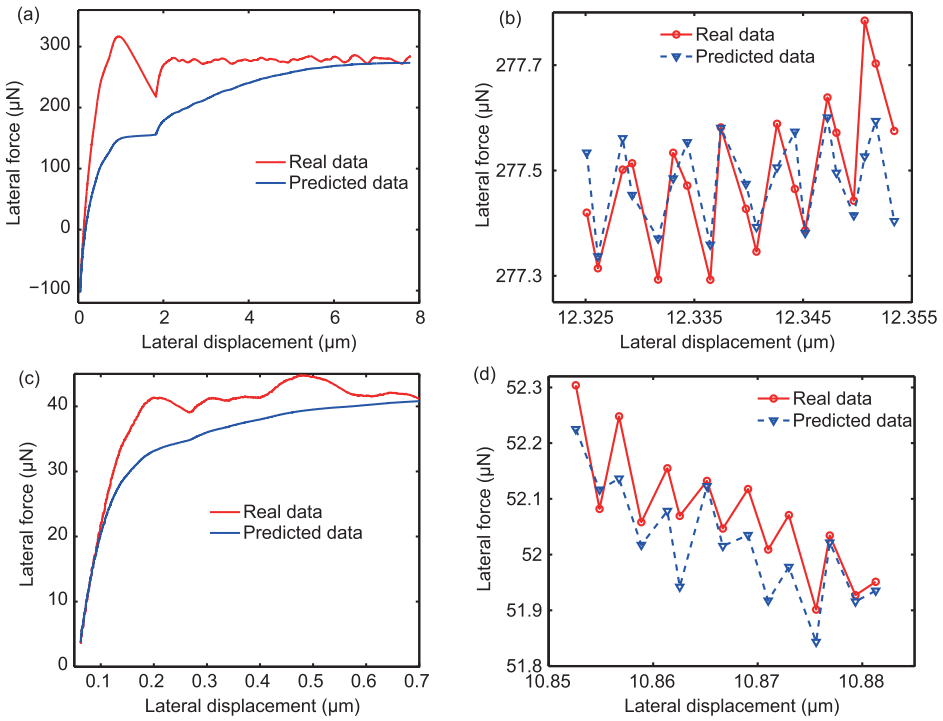

Figure 8 (Color online) Prediction based on the governing system for $Ni_{62}Nb_{38}$ under 2000 μN-5 s ((a), (b)) and 500 μN-5 s ((c), (d)). (a) Taking $x(2000 : 2000 + \tau - 1)$ and $y(2000 : 2000 + \tau - 1)$ as initial values, and step=5600; (b) taking $x(9712 : 9712 + \tau - 1)$ and $y(9712 : 9712 + \tau - 1)$ as initial values, and predicting the following points; (c) taking $x(1400 : 1400 + \tau - 1)$ and $y(1400 : 1400 + \tau - 1)$ as initial values, and step=1300; (d) taking $x(8893 : 8893 + \tau - 1)$ and $y(8893 : 8893 + \tau - 1)$ as initial values, and predicting the following points.

Furthermore, we extract the coefficient matrix of 500 μN-5 s from the first experiment to predict the real data of the second experiment. Setting $x(1400:1400+\tau-1)$, $y(1400:1400+\tau-1)$ as initial values, and step=1300, the result is plotted in Figure 8(c). We obtain that the curves overlap at the beginning, and then slowly separate from each other, which is similar to Figure 8(a). The result shows the reliability of the governing system in repeat experiments. As shown in Figure 8(d), we fix initial value as $x(8893:8893+\tau-1)$ and $y(8893:8893+\tau-1)$, and predict following points. This shows that the following steps prediction is very close to the actual value. Based on the above mentioned results, we obtain that the predicted value from governing system matches better at long-term time series under the stable state; however, it is efficient at short-term time series under unstable state.

3 Summary

As per the sparse identification algorithm [43], we set LF and LD as a combination of $\mathbf{X}(t)$ corresponding to neural network. Then, we employ the fourth-order central difference scheme to obtain the derivative. The ordinary differential equation, $\dot{\mathbf{X}}(t) = \Theta(\mathbf{X})\Xi$, can be derived using ODE45 algorithm [56] after calculating the base function $\Theta(X)$ and sparse matrix Ξ. However, additional studies are required because of the ineffectiveness of the identified system obtained. Therefore, we adopt an alternative method to select reconstruction as an operator to replace the fourth-order central difference scheme. The results indicate that this approach with sparse coefficient provides an outstanding description of the deformation behavior, which can yield substantial information on the nature of the underlying process. Therefore, the sparse dynamics can resolve drastically different phenomena with a small amount of information.

Considering the spatial distribution and temporal motion discontinuity, the deformation behaves as a complex dynamical system, characterized by instability, irregular distribution, catastrophic, and unpredictable motion. Thus, the research of deformation, including influencing factor, correlation, and inner mechanisms, is important. Unlike previous work, our target is to discover a governing system underlying a dynamical system from data measurements. Based on the results of correlation analysis, the evolutionary mechanism is explored by adopting dynamical neural networks. The general explicit equation is reported by the contribution of the sparse identification. The governing system in the nanoscratching process shows the response of LF and LD under varing loading forces, while the governing system under the compressive test describes the relationship between normal stress and normal strain. The excavating results show that the second-order polynomial can have high accuracy and reliability in describing system dynamics, including nanoscratching process and compression experiments. In particular, it can be conducted to predict the deformation process. Considering the field effects, the cause of dependencies from different driving factors (e.g., temperature and strain rate) is divergent. The subject includes the influence of different factors on deformation variables, and the applications of other material fields are worthy of further study.

This work was supported by the National Natural Science Foundation of China (Grant Nos. 52071298, and 11771407), the ZhongYuan Science and Technology Innovation Leadership Program (Grant No. 214200510010), the China Postdoctoral Science Foundation funded project (Grant No. 2019M651600), the Innovative Funds Plan of Henan University of Technology (Grant No. 2020ZKCJ09), and the Research Foundation for Advanced Talents of Henan University of Technology (Grant No. 2018BS027). The authors also would like to thank Miss Yutong Sun for improving the linguistic quality of this article.

1 F. Spaepen, Acta Metall. **25**, 407 (1977).
2 A. S. Argon, Acta Metall. **27**, 47 (1979).
3 M. L. Falk, and J. S. Langer, Phys. Rev. E **57**, 7192 (1998), arXiv: cond-mat/9712114.
4 A. L. Greer, Y. Q. Cheng, and E. Ma, Mater. Sci. Eng.-R-Rep. **74**, 71 (2013).
5 J. O. Krisponeit, S. Pitikaris, K. E. Avila, S. Küchemann, A. Krüger, and K. Samwer, Nat. Commun. **5**, 3616 (2014).
6 Y. Q. Cheng, and E. Ma, Phys. Rev. B **80**, 064104 (2009).
7 P. J. E. Forsyth, Acta Metall. **11**, 703 (1963).
8 J. D. Pribe, T. Siegmund, V. Tomar, and J. J. Kruzic, Int. J. Fatigue **120**, 283 (2019).
9 N. E. Dowling, S. Katakam, and R. Narayanasamy, *Mechanical Behavior of Materials: Engineering Methods for Deformation, Fracture, and Fatigue* (Pearson Education Limited, Boston, 2012).
10 W. R. Ashby, J. Gen. Psychol. **37**, 125 (1947).
11 P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. Lett. **59**, 381 (1987).
12 P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. A **38**, 364 (1988).
13 J. L. Ren, C. Chen, Z. Y. Liu, R. Li, and G. Wang, Phys. Rev. B **86**, 4093 (2012).
14 M. S. Bharathi, M. Lebyodkin, G. Ananthakrishna, C. Fressengeas, and L. P. Kubin, Phys. Rev. Lett. **87**, 165508 (2001), arXiv: cond-mat/0109189.
15 M. A. Lebyodkin, and T. A. Lebedkina, Phys. Rev. E **77**, 026111 (2008).
16 D. X. Han, G. Wang, J. L. Ren, L. P. Yu, J. Yi, I. Hussain, S. X. Song, H. Xu, K. C. Chan, and P. K. Liaw, Acta Mater. **136**, 49 (2017).
17 L. P. Yu, D. X. Han, J. L. Ren, X. X. Guo, S. K. Guan, and G. Wang, Sci. China-Phys. Mech. Astron. **63**, 277011 (2020).
18 A. M. Lyapunov, *The General Problem of the Stability of Motion* (Taylor & Francis, London, 1992).
19 T. Y. Li, and J. A. Yorke, Am. Math. Mon. **82**, 985 (1975).
20 J. L. Kaplan, and J. L. Yoke, *Chaotic Behavior of Multidimensional Difference Equation* (Springer, Berlin, Heidelberg, 1979).
21 W. L. Johnson, and K. Samwer, Phys. Rev. Lett. **95**, 195501 (2005).
22 B. A. Sun, S. Pauly, J. Hu, W. H. Wang, U. Kühn, and J. Eckert, Phys. Rev. Lett. **110**, 225501 (2013).
23 L. Li, E. R. Homer, and C. A. Schuh, Acta Mater. **61**, 3347 (2013).
24 E. G. Daub, D. Klaumünzer, and J. F. Löffler, Phys. Rev. E **90**, 062405 (2014).
25 B. Sun, L. Yu, G. Wang, X. Tong, C. Geng, J. Wang, J. Ren, and W.

Wang, Phys. Rev. B **101**, 224111 (2020), arXiv: 1909.11853.
26 M. L. Falk, and J. S. Langer, MRS Bull. **25**, 40 (2000).
27 J. S. Langer, Phys. Rev. E **77**, 021502 (2008), arXiv: 0712.0399.
28 C. H. Rycroft, and F. Gibou, J. Comput. Phys. **231**, 2155 (2012).
29 V. S. Subrahmanian, and S. Kumar, Science **355**, 489 (2017).
30 M. I. Jordan, and T. M. Mitchell, Science **349**, 255 (2015).
31 J. Xiong, S. Q. Shi, and T. Y. Zhang, Mater. Des. **187**, 108378 (2020).
32 T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning* (Springer-Verlag, New York, 2009).
33 C. Bishop, *Pattern Recognition and Machine Learning* (Springer-Verlag, New York, 2006).
34 A. E. Hoerl, and R. W. Kennard, Technometrics **12**, 55 (1970).
35 R. M. Neal, *Bayesian Learning for Neural Networks* (Springer, New York, 1996).
36 J. R. Quinlan, Mach. Learn. **1**, 81 (1986).
37 C. J. C. Burges, Data Min. Knowledge Discovery **2**, 121 (1998).
38 G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning* (Springer, New York, 2013).
39 C. E. Rasmussen, and C. K. I. Williams, *Gaussian Processes for Machine Learning* (MIT Press, Cambridge, 2006).
40 L. K. Jones, Ann. Statist. **20**, 608 (1992).
41 L. Breiman, Mach. Learn. **26**, 123 (1996).
42 D. H. Wolpert, Neural Networks **5**, 241 (1992).
43 S. L. Brunton, J. L. Proctor, and J. N. Kutz, Proc. Natl. Acad. Sci. USA **113**, 3932 (2016), arXiv: 1509.03580.
44 S. H. Rudy, S. L. Brunton, J. L. Proctor, and J. N. Kutz, Sci. Adv. **3**, e1602614 (2017), arXiv: 1609.06401.
45 H. Jaeger, and H. Haas, Science **304**, 78 (2004).
46 M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. van Nes, M. Rietkerk, and G. Sugihara, Nature **461**, 53 (2009).
47 V. Dakos, E. H. van Nes, R. Donangelo, H. Fort, and M. Scheffer, Theor. Ecol. **3**, 163 (2010).
48 S. H. Strogatz, *Nonlinear Dynamics and Chaos* (Perseus Books, Cambridge, 1994).
49 N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, Phys. Rev. Lett. **45**, 712 (1980).
50 P. Grassberger, and I. Procaccia, Phys. D-Nonlinear Phenom. **9**, 189 (1983).
51 A. M. Fraser, and H. L. Swinney, Phys. Rev. A **33**, 1134 (1986).
52 L. Cao, Phys. D-Nonlinear Phenom. **110**, 43 (1997).
53 A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, Phys. D-Nonlinear Phenom. **16**, 285 (1985).
54 F. Takens, Lect. Notes Math. **898**, 366 (1981).
55 D. Klaumünzer, R. Maaß, and J. F. Löffler, J. Mater. Res. **26**, 1453 (2011).
56 D. Houcque, Applications of MATLAB: Ordinary Differential Equations (ODE), Internal Communication, Robert R. McCormick School of Engineering and Applied Science, Northwestern University, Evanston, 2005.