Proceedings of Machine Learning Research vol 245:1–9, 2024 2024 International Conference on Machine Learning and Intelligent Computing

Polynomial Fitting Based on Integrable Deep Neural Networks for Landau-Energy of Ferroelectrics

Wenyu Zhang
School of Mechanical Power and Engineering, East China University of Science and Technology, Shanghai, 200237, China

Yabin Yan YANYABIN@ECUST.EDU.CN
School of Mechanical Power and Engineering, East China University of Science and Technology, Shanghai, 200237, China

Editors: Nianyin Zeng and Ram Bilas Pachori

Abstract

Fitting the Landau-Energy polynomial has always been challenging because it is difficult to directly obtain Landau-Energy data for coefficient fitting. One possible approach to address this problem is to handle the derivative of the Landau-Energy polynomial with respect to the second-order polarization (dielectric constant) to obtain relevant information about the Landau-Energy. This chapter will introduce a method based on integrable neural networks to obtain an approximate model for the Landau-Energy polynomial and its parameters.

Keywords: Ferroelectrics, Landau-Energy, Integrable Deep Neural Networks

1. Introduction

The study of Landau-Energy in the context of ferroelectric materials is crucial for advancing our understanding of phase transitions and their underlying mechanisms. The intricate relationship between temperature, polarization, and the dielectric properties of materials forms the foundation of Landau theory, which has been extensively applied to predict and analyze the behavior of ferroelectrics near phase transitions (Woo and Zheng, 2008; Yazawa et al., 2022; Lich et al., 2016). However, the traditional methodologies employed for fitting the Landau-Energy polynomial often struggle with the limited availability and variability of experimental data, which can impede the development of accurate predictive models. Recent advancements have seen the integration of various computational techniques to improve the precision and reliability of these fittings (Raissi et al., 2019; Teichert et al., 2019, 2020). However, these methods still face challenges regarding scalability and adaptability to complex ferroelectric behaviors.

Therefore, in current study, we extend the IDNN method to second-order partial derivatives and proposes the IDNN-PSO approach for parameter estimation in the ferroelectric Landau-Energy polynomial. Finally, the IDNN-PSO method is applied to real data for parameter estimation, and the fit of the Landau-Energy polynomial to the original dataset in terms of temperature-polarization and temperature-dielectric constant is evaluated.

2. Method

The core idea of solving partial differential equation (PDE) problems based on neural networks is to use the control equation, initial values, and boundary conditions to sample data points within the domain and at the domain boundaries, which form the training dataset. The neural network N is then trained with these data points to obtain optimized parameters. This means that after training, the neural network will satisfy the constraints of the control equation and boundary conditions within a certain range of error. When the model's training accuracy meets the requirements, the neural network can be regarded as an approximate solution or agency model for the PDE problem. Neural networks are used to approximate solutions under partial differential equation (PDE) constraints by fitting observational datasets. IDNN (Raissi et al., 2019) has been proposed to handle partial differential data using neural networks. This method achieves the differentiation of neural networks by applying the chain rule of forward propagation. The following are the symbols for neural networks, as shown in Table 1.

Table 1: Neural Network Representation Symbols

Symbol	Meaning
L	The number of layers in a neural network
M_l	The number of neurons in layer l
$f_l(\cdot)$	The activation function of layer l neurons
$W^l \in R^{M_l \times M_{l-1}}$	Weight matrix from layer $l-1$ to layer l
$b^l \in R^{M_l}$	Bias from layer $l-1$ to layer l
$z^l \in R^{M_l}$	Net input (net activity value) of layer l neurons
$a^l \in R^{M_l}$	Output of layer l neurons (activity value)

For the fully connected neural network, the activation value a_i^l of neuron i in the $l-th$ layer is given by:

$$a_i^l = f\left(z_i^l\right) \tag{1}$$

$$z_i^l = b_i^l + \sum_{j=1}^{m_{l-1}} W_{i,j}^l a_j^{l-1} \tag{2}$$

The activity value of the output layer of the neural network is,

$$Y_i = b_i^{n+1} + \sum_{j=1}^{m_n} W_{i,j}^{n+1} a_j^n \tag{3}$$

Based on dataset $D_f \{(x_1, y_1), \quad (x_2, y_2), \quad \ldots (x_n, y_n)\}$, parameter optimization of neural networks can be expressed as minimizing the loss function $loss\,(W, b\,|D_f\,)$ with weight matrix W and bias matrix b as independent variables,

$$\hat{W}, \hat{b} = \arg\min_{W,b} \mathrm{MSE}\left(Y(x, W, b)|_{x_\theta}, y_\theta\right) \tag{4}$$

Similarly, the optimization of partial differential datasets by IDNN can be expressed as,

$$\hat{W}, \hat{b} = \arg\min_{W,b} \sum_{k=1}^{n} \mathrm{MSE}\left(\left. \frac{\partial Y(x, W, b)}{\partial x_k} \right|_{x_\theta}, y_\theta \right) \tag{5}$$

From equation (5), it can be observed that IDNN's differentiation with respect to neural networks has an analytical form. Specifically, it follows the chain rule within the neural network. The differentiation of IDNN's neurons has some slight variations in form, where the activation unit can be considered as the partial derivative form of DNN's activation unit.

$$\frac{\partial a_i^l}{\partial x_k} = f'\left(z_i^l\right) \frac{\partial z_i^l}{\partial x_k} \tag{6}$$

$$a_i^l = f\left(z_i^l\right) \tag{7}$$

$$\frac{\partial z_i^l}{\partial x_k} = \sum_{j=1}^{m_{l-1}} W_{i,j}^l \frac{\partial a_j^{l-1}}{\partial x_k} \tag{8}$$

By rearranging equations (6-8), new activation unit α_i^l, $\beta_{k_i}^l$ can be obtained.

$$\beta_{k_i}^l = f'\left(z_i^l\right) \sum_{j=1}^{m_{l-1}} W_{i,j}^l \beta_{k_j}^{l-1} \tag{9}$$

$$\alpha_i^l = f\left(z_i^l\right) \tag{10}$$

Perform chain calculation on feedforward neural networks, and the partial differentiation of the neural network can be expressed as,

$$y_{i,k} = \frac{\partial Y_i}{\partial x_k} = \sum_{j=1}^{m_n} W_{i,j}^{n+1} \beta_{k_j}^n \tag{11}$$

$$Y_i = \sum_{j=1}^{m_n} W_{i,j}^{n+1} \alpha_j^n \tag{12}$$

Based on the above formula derivation, IDNN can use equations (5) and (9-12) to solve first-order PDE problems. Overall, IDNN is a partial differential variant of DNN, which shares weight matrices with DNN. During the parameter iteration of the neural network, both partial differential neurons and regular neurons participate in backpropagation optimization. Therefore, IDNN can recover the original function by training partial differential data of the original function and a small number of initial values, that is, realizing the integration process. However, since the above formulas can only handle first-order partial derivatives, we extend the IDNN algorithm by introducing second-order activity values.

$$\gamma_{(p,q)_i}^0 = 0 \tag{13}$$

$$\gamma_{(p,q)_i}^1 = f''\left(z_i^l\right) \cdot f'\left(z_i^1\right)^2 \cdot \left(W_{i,p}^1 \cdot W_{i,q}^1\right) \tag{14}$$

$$\gamma^{l}_{(p,q)_i} = f''(z^l_i)\left(\sum_{j=1}^{m_{l-1}} W^l_{i,j}\beta^{l-1}_{p_j}\right)\left(\sum_{j=1}^{m_{l-1}} W^l_{i,j}\beta^{l-1}_{q_j}\right) + f'\left(z^l_i\right)\sum_{j=1}^{m_{l-1}} W^l_{i,j}\gamma^{l-1}_{(p,q)_j} \tag{15}$$

Equations (13-15) represent the second partial derivative of the first-order partial differential activity values in IDNN. In the IDNN output layer, the second mixed partial derivative of DNN is given by:

$$y_{i,(p,q)} = \frac{\partial^2 Y_i}{\partial x_p \partial x_q} = \sum_{j=1}^{m_n} W^{n+1}_{i,j}\gamma^{n}_{(p,q)_j} \tag{16}$$

Therefore, in order to achieve the goal of restoring the original function through the convergence of the second-order IDNN, the following conditions are required: Firstly, the Hessian matrix $H(f)$ of the second derivative of the original function; Secondly, the Jacobian matrix $J(f)$ of the first partial derivative matrix of the original function; and finally, a few initial values $f(x_i)$ of the original function.

3. Polynomial Fitting based on Second-Order IDNN

The Landau-Energy theory reproduces the experimental observations of mineral phase transitions, and its accuracy is sufficient for many applications. Assuming a ferroelectric Landau-Energy polynomial as follows:

$$\begin{aligned} F_{\text{Land}}\ (T, P_1, P_2, P_3) = {} & a\,(T - T_0)\left(P_1^2 + P_2^2 + P_3^2\right) + a_{11}\left(P_1^4 + P_2^4 + P_3^4\right) \\ & + a_{12}\left(P_1^2P_2^2 + P_2^2P_3^2 + P_3^2P_1^2\right) + a_{111}\left(P_1^6 + P_2^6 + P_3^6\right) \\ & + a_{112}\left(P_1^4\left(P_2^2 + P_3^2\right) + P_2^4\left(P_1^2 + P_3^2\right) + P_3^4\left(P_1^2 + P_2^2\right)\right) \\ & + a_{123}P_1^2P_2^2P_3^2 + a_{1111}\left(P_1^8 + P_2^8 + P_3^8\right) \\ & + a_{1112}\left(P_1^6\left(P_2^2 + P_3^2\right) + P_2^6\left(P_2^2 + P_3^2\right) + P_3^6\left(P_2^2 + P_3^2\right)\right) \\ & + a_{1122}\left(P_1^4P_2^4 + P_2^4P_3^4 + P_3^4P_1^4\right) \\ & + a_{1123}\left(P_1^2P_2^2P_3^2\left(P_1^2 + P_2^2 + P_3^2\right)\right) \end{aligned} \tag{17}$$

$$a = \frac{1}{2\kappa_0 C_0} \tag{18}$$

Here, $T$is the instantaneous temperature, T_0 is the Curie-Weiss temperature, $\begin{pmatrix} P_1, & P_2, & P_3 \end{pmatrix}$ is the polarization vector, a_i is the dielectric coefficient, C_0 represents the Curie constant, and $\kappa_0$is the vacuum permittivity. Different forms of polarization vectors are determined for different phases of a ferroelectric body, including C-phase $(0,0,0)$, T-phase $(P,0,0)$, O-phase $(P/\sqrt{2}, P/\sqrt{2}, 0)$, and R-phase $(P/\sqrt{3}, P/\sqrt{3}, P/\sqrt{3})$. In the Landau-Energy polynomial fitting problem, the known conditions are the polarization intensity and dielectric constant corresponding to different temperatures as

$$P_s\,|_{T=T_i} = \arg\min_{P}\ (F_{Land}\,|_{T=T_i}) \tag{19}$$

In this case, there exists a partial derivative relationship between polarization intensity and Landau-Energy, given by equation (20-22).

$$\frac{\partial F_{Land}}{\partial P_i} |_{P=P_s} = 0 \tag{20}$$

$$x_{ij} = \varepsilon_0 \frac{\partial^2 F_{Land}}{\partial P_i \partial P_j} \tag{21}$$

$$\varepsilon_{ij} = \frac{1}{x_{ij}} \tag{22}$$

Where ε_{ij} represents the dielectric constant corresponding to polarization. According to the relationship (22) between the dielectric constant and Landau-Energy, the form of the dielectric constant can be determined for different phases. The main objective of using IDNN to fit F_{Land} is to approximate a agency model function similar to $F_{Land}(P,T)$. Although $F_{Land}(P,T)$ is discontinuous and non-differentiable at phase transition temperatures, IDNN can approximate the original function with a certain precision. The following introduces a method for constructing the Hessian matrix $H(f)$ and Jacobian matrix $J(f)$ based on the relationship between F_{Land} and ε_{ij}. To ensure algorithm convergence, data preprocessing is performed initially, including dimensionless scaling

$$P^* = P/P_0 \tag{23}$$

$$T^* = (T - T_0)/ |T_d - T_0| \tag{24}$$

Where P_0 and T_d can be set by oneself, and it is recommended to set them as the minimum value or median value in the dataset. T_0 represents the instantaneous temperature, which needs to be obtained from the data. In this case, the form of the agency model function becomes $F^*_{Land}(P^*, T^*)$. The relationship between the dimensionless agency model function and the original function in terms of the parameters is as follows:

$$\alpha_1 = (T_d - T_0) /2\kappa_0 C_0 \tag{25}$$

$$\alpha_0 = |\alpha_1| \tag{26}$$

$$a^* = 1 \tag{27}$$

$$a^*_{ij} = \left(a_{ij} \cdot P_0^2\right) /a_0 \tag{28}$$

$$a^*_{ijk} = \left(a_{ijk} \cdot P_0^4\right) /a_0 \tag{29}$$

$$a^*_{ijkl} = \left(a_{ijkl} \cdot P_0^6\right) /a_0 \tag{30}$$

Where κ_0 and C_0 represent the permittivity in vacuum and the Curie temperature, respectively, and are represented in the agency model as unknown constants. This dimensionless approach ensures that the data used for training the IDNN is within a reasonable range defined by the Sigmoid function, preventing the occurrence of vanishing or exploding gradients during the gradient iteration process. The $H(f)$ and $J(f)$ used for training the dimensionless agency model are as follows:

$$H\left(F^*_{\text{Land}}\right) = \begin{pmatrix} \frac{\partial^2 F^*_{\text{Land}}}{\partial P^{*2}} & \frac{\partial^2 F^*_{\text{Land}}}{\partial P^* \partial T^*} \\ \frac{\partial^2 F^*_{\text{Land}}}{\partial P^* \partial T^*} & \frac{\partial^2 F^*_{\text{Land}}}{\partial T^*} \end{pmatrix} \tag{31}$$

$$J\left(F^*_{\text{Land}}\right) = \begin{pmatrix} \frac{\partial F^*_{\text{Land}}}{\partial P^*} & \frac{\partial F^*_{\text{Land}}}{\partial T^*} \end{pmatrix} = \begin{pmatrix} 0 & \frac{\partial F^*_{\text{Land}}}{\partial T^*} \end{pmatrix} \tag{32}$$

The convergence condition of the IDNN also requires a small number of values from the original function. Here, the Landau Energy of the C phase is used to fulfill this condition, i.e.,

$$F_{Land|P=0} = 0 \tag{33}$$

Equations (31-33) represent all the necessary conditions for the convergence of IDNN.

The significance of the agency model lies in its ability to connect the independent variables P, T with the Landau Energy space, rather than the dielectric constant. This allows for easier fitting of Landau Energy polynomial parameters. Here, the Particle Swarm Optimization (PSO) algorithm is chosen for parameter fitting, primarily because PSO exhibits strong global search capabilities and is easy to implement. Additionally, in dealing with high-dimensional problems, the search process of PSO does not depend on the dimensionality of the problem, as the Landau Energy polynomial has 9 unknown coefficients. The form of the loss function in the optimization task is as follows:

$$N \quad S_N = \left\{ \hat{t}_i \mid \hat{t}_i \neq t_i \right\} \tag{34}$$

$$\text{Penalty} = \sum |\alpha_i| \tag{35}$$

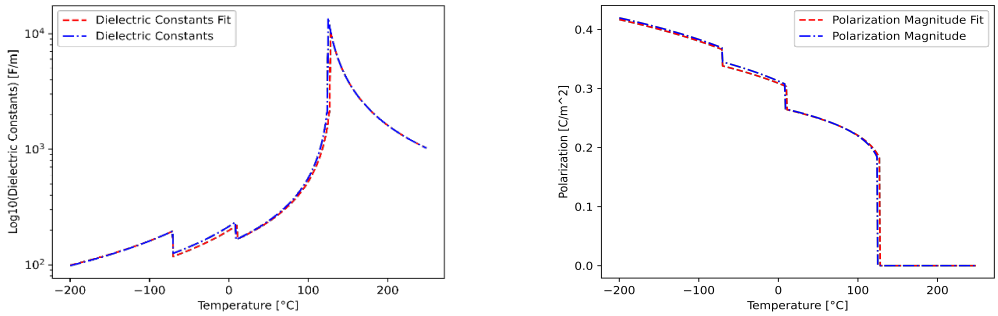
$$\begin{aligned} l\left(\alpha_1, \ldots, a_{1123} \mid w_1, \ldots, w_5\right) =& w_1 \Delta f + w_2 \Delta \frac{\partial^2 f}{\partial P^2} + w_3 \Delta \frac{\partial f}{\partial P} \\ &+ w_4 N + w_5 \bullet \text{Penalty} \end{aligned} \tag{36}$$

Equation (36) represents the polynomial of the loss function, which contains 5 terms: the loss term of the original function, the loss term of the second derivative, the loss term of the first derivative, the number of incorrectly predicted phase attributes of data points, and the penalty term. The values of the weights for each loss term are usually set to 1. The loss term of the original function is the main driving force for parameter fitting, while the first derivative and second derivative play an auxiliary role in structural convergence. Additionally, we note that the information of the original function values and derivatives alone is not sufficient to characterize the features of the Landau energy polynomial with respect to the phase. Therefore, the number of data points with incorrectly predicted phase attributes is added to the loss function for prediction. Furthermore, without any constraints, the optimization algorithm may produce extreme results in parameter optimization. To avoid the optimized parameters being located on the boundary of the parameter space, a penalty term is used to control the absolute value of each parameter. Finally, the parameters obtained by the optimization algorithm are still dimensionless and need to be scaled.

4. Verification of reliability

Using the parameters listed in Table 2, we randomly selected 100 sets of temperatures within the range of -200 200°C according to (1.21) to calculate the corresponding polarization and dielectric constants. Subsequently, we used them to train the IDNN surrogate model and fit the Landau-energy polynomial parameters using PSO optimization. Based on the fitted parameters, polarization and dielectric constants for the temperature range of interest were predicted and shown in Figure 1. Overall, the predicted results agree well with the ideal data. Although the objective function of the IDNN is not differentiable at the phase transition temperature, the modeling strategy of piecewise functions and the loss term of phase characteristics in the loss function have played their roles, and the fitted parameters still perform well near the phase transition temperature. Table 2 shows the

fitted parameters, which numerically differ from the ideal parameters. However, this is an inevitable result of finding a local optimum through global optimization with PSO.



(*a*) (*b*)

Figure 1: Comparison of the Fitted Parameters and Original Parameters for (a) Dielectric Constant (b) Polarization.

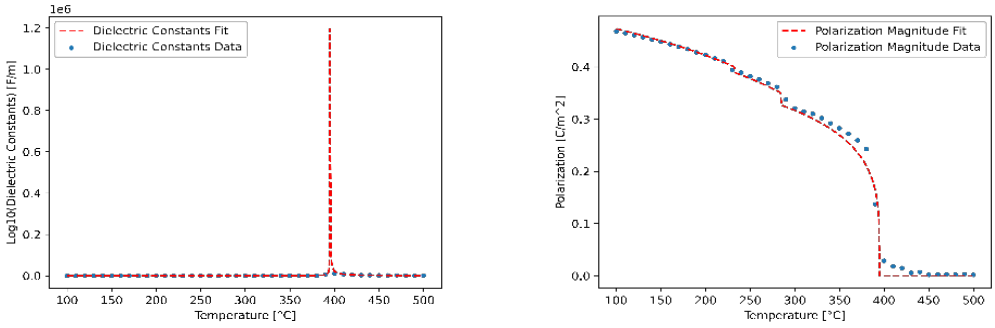
Table 2: Landau-Energy Polynomial Parameters in the Test Experiment

Parameter	value	fitting
a	412400	412400
T_0	115	115
a_{11}	-2.10×10^8	-2.60×10^8
a_{12}	7.97×10^8	8.24×10^8
a_{111}	1.29×10^9	2.24×10^9
a_{112}	-2.50×10^9	-3.30×10^9
a_{123}	-2.50×10^9	5.49×10^9
a_{1111}	3.86×10^{10}	3.45×10^{10}
a_{1112}	2.53×10^{10}	3.9×10^{10}
a_{1122}	1.64×10^{10}	3.9×10^{10}
a_{1123}	1.37×10^{10}	3.5×10^{10}

We also tested the parameter solution of BaTiO3 temperature polarization dielectric constant data generated by IDNN based on the effective Hamiltonian method (Figure 2, Table 2), and the model also showed good convergence under high noise data.

5. Conclusion

In summary, we have developed a method for solving the parameters of the ferroelectric Landau-Energy polynomial based on IDNN. Considering the mathematical relationship between Landau-Energy, dielectric properties, and temperature, we propose a second-order partial differential IDNN algorithm. By dimensionless scaling of the Landau-Energy polynomial, we adapted the temperature-dielectric and temperature-polarization datasets to meet the convergence conditions of the neural



(*a*) (*b*)

Figure 2: Comparison of the Fitted Parameters and Original BaTiO3 Parameters of Effective Hamiltonian for (a) Dielectric Constant; (b) Polarization.

network model. The convergence of the model demonstrates the feasibility of modeling the Landau-Energy polynomial surrogate model using the second-order IDNN algorithm. Subsequently, the surrogate model parameters were optimized using the PSO algorithm, and the temperature-polarization and temperature-dielectric constant datasets were calculated using the parameterized Landau-Energy polynomial. Observations showed that the surrogate model data maintained high consistency with the training dataset and reflected the phase transition characteristics of the training data. This indicates that the IDNN-PSO based Landau-Energy polynomial parameter estimation method has sufficient reliability in terms of convergence and accuracy, and possesses certain learning adaptability to the physical mechanisms of ferroelectric phase transitions.

References

Le Van Lich, Takahiro Shimada, Jie Wang, and Takayuki Kitamura. Instability criterion for ferroelectrics under mechanical/electric multi-fields: Ginzburg-landau theory based modeling. *Acta Materialia*, 112:1–10, 2016. doi: 10.1016/j.actamat.2016.03.081.

M. Raissi, P. Perdikaris, and G.E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019. doi: 10.1016/j.jcp.2018.10.045.

G.H. Teichert, A.R. Natarajan, A. Van der Ven, and K. Garikipati. Machine learning materials physics: Integrable deep neural networks enable scale bridging by learning free energy functions. *Computer Methods in Applied Mechanics and Engineering*, 353:201–216, 2019. doi: 10.1016/j.cma.2019.05.019.

G.H. Teichert, A.R. Natarajan, A. Van der Ven, and K. Garikipati. Scale bridging materials physics: Active learning workflows and integrable deep neural networks for free energy function representations in alloys. *Computer Methods in Applied Mechanics and Engineering*, 371:113281, 2020. doi: 10.1016/j.cma.2020.113281.

C. H. Woo and Yue Zheng. Depolarization in modeling nano-scale ferroelectrics using the landau free energy functional. *Applied Physics A*, 91(1):59–63, 2008. doi: 10.1007/s00339-007-4355-4.

Keisuke Yazawa, Andriy Zakutayev, and Geoff L. Brennecka. A Landau–Devonshire analysis of strain effects on ferroelectric $Al_{1-x}Sc_xN$. *Applied Physics Letters*, 121(4):042902, 07 2022. doi: 10.1063/5.0098979.