

---

# Revealing the Mechanism of Large-scale Gradient Systems Using a Neural Reduced Potential

---

**Shunya Tsuji**

The University of Electro-Communications  
t2230094@edu.cc.uec.ac.jp

**Ryo Murakami**

National Institute for Material Science  
MURAKAMI.Ryo@nims.go.jp

**Hayaru Shouno**

The University of Electro-Communications  
shouno@uec.ac.jp

**Yoh-ichi Mototake**

Hitotsubashi University  
y.mototake@r.hit-u.ac.jp

## Abstract

Constructing the reduction model of large-scale pattern dynamics is challenging. In this study, a framework is proposed to estimate a reduction model of the gradient system, often observed in various pattern dynamics, in a data-driven manner using a deep learning model inspired by Hamiltonian neural networks for video. Furthermore, the proposed framework verifies whether the reduction model is consistent with the phenomenon and contains useful properties. To demonstrate its usefulness, it is applied to the numerical calculation data of magnetic domain pattern formation. Consequently, the previous reduction model proposed for magnetic domain pattern dynamics is found to be insufficient to explain the phenomenon, and suggestions for possible directions for the improvement of the reduction model are provided.

## 1 Introduction

Various natural phenomena, such as crystal growth [9, 13], organism pattern formation [15], and magnetic pattern formation [7, 8, 10], can be modeled as large-scale gradient systems that follow the gradient of an energy function (potential function). The potential function depends only on the system state. In such gradient systems, both the potential function and its gradient play crucial roles.

To understand the macroscopic behavior of such large-scale systems, scientists have constructed not only microscopic models but also its reduction models for various unknown phenomena that reflect the representative features of the observed phenomena without contradicting them [1–3, 14]. However, the construction of reduction models typically requires extensive experimentation, considerable knowledge, and novel insights from experts, thereby rendering such modeling highly challenging. Chen et al. [4] developed a framework to support the construction of reduction models using machine learning techniques. Chen et al. used neural networks to extract low-dimensional reduced coordinate systems from the spatiotemporal data of large-scale systems; subsequently, they modeled time evolution in a reduced coordinate system. Generally, models obtained using neural networks are black boxes that cannot be interpreted by humans. Chen et al. evaluated a reduced coordinate system modeled by a neural network to extract information that could be used to guide physicists in constructing interpretable reduction models. However, Chen et al. interpreted only the reduced coordinate system, and did not focus on the interpretation of the time evolution model. In gradient systems, which are the focus of our study, time evolution is governed by the potential function; therefore, extracting interpretable information regarding the potential function of the reduction model is particularly essential.

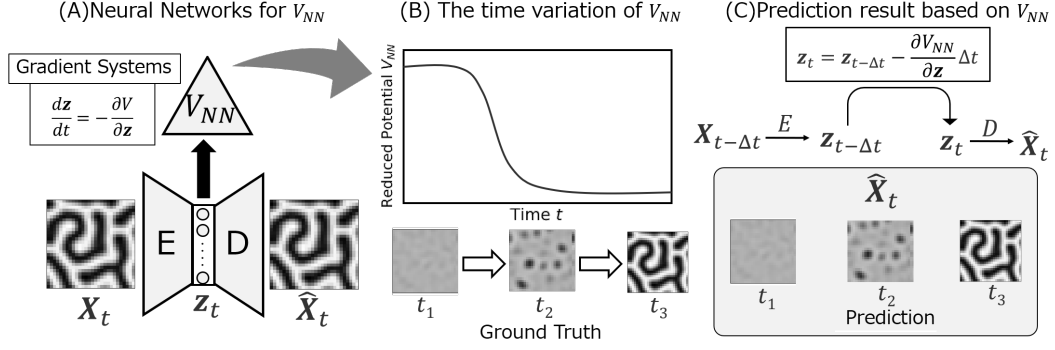


Figure 1: Overview diagram of the proposed framework in this study. First, using the network shown in (A), the reduced potential function  $V_{NN}$  is obtained, as shown in (B). Next, as illustrated in (C), the true time-series image data at a certain time is compared with the predicted time-series image data derived according to  $V_{NN}$  to validate its suitability as a potential function. Subsequently, the gradients of  $V_{NN}$  are analyzed to establish connections with the insights and properties of the phenomenon.

In this study, a framework to estimate a reduction model of the potential function (reduced potential function) from data for unknown phenomena that are expected to follow gradient systems is proposed. Furthermore, the proposed framework validates whether the reduced potential function is consistent with the observed phenomena and incorporates the useful properties of the phenomena.

When the reduced potential function is obtained, useful properties governing gradient systems, such as the potential gradient, can be extracted.

## 2 Proposed framework

In this study, a framework that estimates the reduced potential function from data for unknown phenomena expected to follow gradient systems is proposed. The framework further validates whether the model is consistent with the observed phenomena and incorporates its useful properties.

An overview of the proposed framework is presented in Figure 1. In this framework, as shown in Figure 1(A), firstly the reduced potential function of a gradient system is estimated by training a deep neural network  $V_{NN}$ , inspired by Hamiltonian neural networks (HNN) [5], on time-series image data of unknown phenomena. HNN is an excellent neural network, and it can estimate the energy function (Hamiltonian) from time-series image data of the phenomenon. It consists of an autoencoder (AE) [6], which estimates the spatial distribution (phase space) of the system positions and momenta from the time-series image data, and neural network, which models the relationship between the phase space and Hamiltonian. The details of the model shown in Figure 1(A) are discussed later in this section. Subsequently, as shown in Figure 1(B), the values of the reduced potential function for each time step of the time-series image data are obtained using the trained network. In gradient systems, the following is known.

$$\frac{d\mathbf{u}}{dt} = -\frac{\partial V}{\partial \mathbf{u}} \quad (1)$$

where  $\mathbf{u}$  represents the state of the system,  $t$  is time, and  $V$  is the potential function. Equation (1) indicates that by moving  $\mathbf{u}$  in the direction of  $-\frac{\partial V}{\partial \mathbf{u}}$ , the temporal evolution of the system can be obtained. Therefore, as shown in Figure 1(C), to verify whether  $V_{NN}$  captures the gradient information of the phenomenon with reasonable accuracy, the time-series image data perturbed by  $\Delta t$  according to the gradients of  $V_{NN}$  can be compared with the true time-series image data.

Next, we describe the details of the model shown in Figure 1(A) used to estimate the reduced potential function using the data. To model the reduced potential function, first, a latent space that explains the system is constructed using an AE. The AE used in this study is composed of an encoder network  $E$ , which transforms the time-series data  $X_t$  at a specific time  $t$  into a reduced vector  $\mathbf{z}_t$ , and decoder network  $D$ , which predicts  $X_t$  from  $\mathbf{z}_t$  as  $E(X_t) = \mathbf{z}_t$ ,  $D(\mathbf{z}_t) = \hat{X}_t$ . AE is trained to minimize the loss

function  $\mathcal{L}_{AE} = \|X - \hat{X}\|_2$ . Using such an AE, a space in which the time-series data  $X_t$  are represented as reduced vector  $\mathbf{z}_t$  can be created. The obtained reduced vector  $\mathbf{z}_t$  from the AE, which represents the time-series data  $X_t$  in a reduced form, can be considered as the system state  $\mathbf{u}_t$ ; thus, we treat  $\mathbf{z}_t = \mathbf{u}_t$ .

Notably, the neural network  $V_{NN}$  that constructs the reduced potential function in the latent space mapped by the AE can learn the values that are "similar" to the system's potential from the data using a loss function, based on Equation (1).  $V_{NN}$  learns the reduced potential function by minimizing the loss function represented by  $\mathcal{L}_V = \left\| \frac{d\mathbf{z}}{dt} + \frac{\partial V_{NN}}{\partial \mathbf{z}} \right\|_2$ , while assuming that  $\frac{d\mathbf{z}}{dt}$  is observable and considered as data. In this study, such "similar" reduced potential functions obtained by training the neural network such as  $V_{NN}$  are referred to as the Neural Reduced Potential.

Herein, the entire network used for estimating the reduced potential function in the gradient system was trained to minimize the combined loss function, which is the sum of  $\mathcal{L}_{AE}$  and  $\mathcal{L}_V$  weighted by  $\lambda$ .

$$\mathcal{L}_{all} = \|X - \hat{X}\|_2 + \lambda \left\| \frac{d\mathbf{z}}{dt} + \frac{\partial V_{NN}}{\partial \mathbf{z}} \right\|_2 \quad (2)$$

Here,  $\lambda = 0.1$ . Furthermore, note that the objective of this study was not to propose a method for predicting the dynamics of a system from observable data but to determine a potential function based on the data. Therefore, to ensure that the determined potential function represents the phenomenon more accurately, the input to  $V_{NN}$  includes both  $\mathbf{z}$ , which represents the system, and some of the physical parameters that explain it.

### 3 Demonstration

In this study, the effectiveness of the proposed method was validated by introducing the case applied in the proposed framework to the simulation data of the magnetic domain pattern dynamics occurring in magnetic materials. Magnetic materials are important industrial materials that are widely used in various devices, such as car motors and magnetic heads on hard disks. Understanding the energy possessed by magnetic materials is crucial for considering their suitability and performance as magnetic materials. The potential function of the magnetic domain pattern dynamics is often modeled as:  $\frac{\partial \phi(\mathbf{r})}{\partial t} = -\frac{\delta H}{\delta \phi(\mathbf{r})}$ , using the time-dependent Ginzburg–Landau (TDGL) equation [11]. Where

$H = \alpha \lambda(\mathbf{r}) \int d\mathbf{r} \left( -\frac{\phi(\mathbf{r})^2}{2} + \frac{\phi(\mathbf{r})^4}{4} \right) + \beta \int d\mathbf{r} \frac{|\nabla \phi(\mathbf{r})|^2}{2} + \gamma \int d\mathbf{r} d\mathbf{r}' \frac{\phi(\mathbf{r})\phi(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|^3} - h(t) \int d\mathbf{r} \phi(\mathbf{r})$ . The TDGL equation is already a 2D model that effectively captures this phenomenon. However, it is a high-dimensional model that describes the system using a scalar field  $\phi(\mathbf{r}) : \mathbb{R}^2 \rightarrow \mathbb{R}$  composed of the average of the vertical components of the spins in small grid-like regions, and the resulting patterns can be complex. Thus, extracting interpretable descriptors to understand the mechanisms of a system is challenging when using only this model. Therefore, we attempted to extract them from the reduced potential function acquired by applying the proposed approach to the magnetic domain pattern dynamics obtained through simulations using the TDGL equation. The simulation data used in this study were generated based on the study by Kudo et al. [10], as follows. First, a saturation magnetic field  $h_{init}$  was applied, and a constant decay rate  $v$  was assumed for the external magnetic field  $h(t)$ . Therefore,  $h(t)$  can be written as  $h(t) = h_{init} - vt$  ( $t < T_0$ ) or  $h(t) = 0$  ( $t \geq T_0$ ), where,  $T_0$  is the time at which  $h(t)$  first becomes 0 and is considered as  $T_0 = \frac{h_{init}}{v}$ . The simulation was performed up to the time  $t = 2T_0$ . Further,  $v = 10^{-2}$ , the magnetic interaction between neighboring spins was set as  $\beta = 2.0$ , the magnetic dipolar interaction was  $\gamma = 2.0/\pi$ ,  $\lambda(\mathbf{r}) \sim \mathcal{N}(0, 0.3^2)$ ,  $h_{init} = 1.5$ , and the magnitude of the anisotropy  $\alpha$  was varied within the range of [1.0, 4.0].

In the demonstration, the input  $X_t$  to the AE used in proposed approach is a matrix with  $\phi(\mathbf{r})$ , simulated based on TDGL equation, as elements. Furthermore, to ensure that the determined potential function represents the phenomenon more accurately, the input to  $V_{NN}$  includes not only  $\mathbf{z}$ , which represents the system, but also the magnitude of the anisotropy  $\alpha$  and the external magnetic field  $h(t)$ .

#### 3.1 Reduced potential function of magnetic domain pattern formation process

This section discusses the study by Mototake et al. [12], who attempted to construct reduced potential function in the typical manner using their novel insights. They conducted analysis on the magnetic domain pattern dynamics obtained through simulations using the TDGL equation and thereby

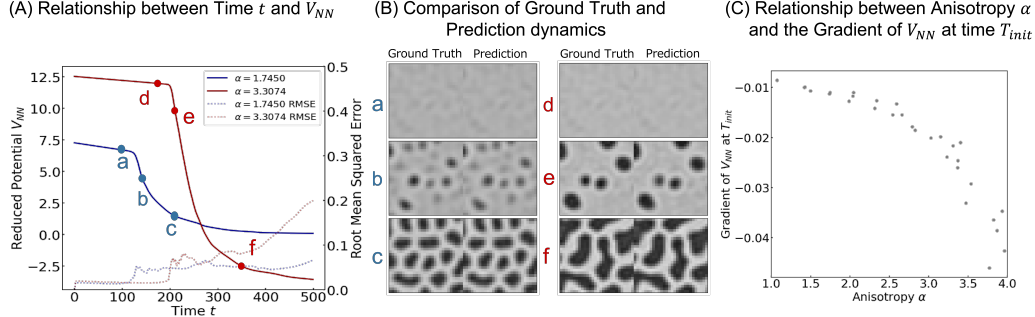


Figure 2: Results obtained by applying the proposed framework to the simulation data of the magnetic domain pattern dynamics. The solid lines in (A) represent the relationship between  $V_{NN}$  and time  $t$ , and the dotted lines represent the root mean square error of the true pattern images and the predicted pattern images obtained by varying true pattern image at the time  $t - 1$  according to the  $V_{NN}$ . Here, the red and blue lines depict the differences in the time evolution of  $V_{NN}$  owing to variations in the anisotropy  $\alpha$ . Additionally, the points labeled a–f on the left figure correspond to (B). The "Ground Truth" column in (B) represents the true pattern images at time  $t$  for points a–f, whereas the "Prediction" column shows the predicted pattern images. (C) represents the relationship between the gradient of  $V_{NN}$  and anisotropy  $\alpha$  at time  $T_{init}$ .

demonstrated that the final magnetic domain patterns can undergo transitions in the following order with respect to changes in anisotropy: labyrinth structure, island structure (or a mixture of labyrinth and island structures), and then labyrinth structure again. Furthermore, they confirmed that these two labyrinth structures undergo different formation processes. Based on these findings, they focused on the moment of occurrence of the negative magnetic domains and made significant assumptions, such as the formation of the same magnetic domain. Thus, they proposed a reduced potential function that can explain the two distinct labyrinth structures with different formation processes. This reduced potential function undergoes a qualitative transition beyond a certain threshold of anisotropy, particularly before and after the occurrence of negative domains. Furthermore, this reduced potential function successfully explains the differences in the formation processes of the two labyrinth structures based on such characteristics.

## 4 Result

In this study, next, the proposed approach was applied to the simulation data based on the TDGL equation of the magnetic domain pattern dynamics, and the obtained results are discussed as follows.

The solid lines in figure 2(A) depict the relationship between acquired Neural Reduced Potential  $V_{NN}$  and time, and the dotted lines depict the root mean square error of the true pattern images at time  $t$  and the predicted pattern images at time  $t$  obtained by varying true pattern image at the time  $t - 1$  according to the  $V_{NN}$ . Here, the blue and red lines represent different values of anisotropy  $\alpha$ . Figure 2(B) shows comparison of the true and predicted pattern images at each time step. The column labeled "Ground Truth" in Figure 2(B) corresponds to the true pattern images at each time  $t$ , whereas the "Prediction" column represents the predicted pattern images at time  $t$ . The rows labeled a–f correspond to points a–f plotted in Figure 2(A). The dotted lines in figure 2(A) and figure 2(B) show that the true and predicted pattern images at each time step are significantly similar. Therefore, the Neural Reduced Potential  $V_{NN}$ , acquired using the proposed framework, is the reduced potential function that adequately represents the properties of the potential in the gradient system.

Potential gradients are particularly important for gradient systems. Therefore, herein, valuable insights into the properties of this phenomenon were revealed using the gradient information of the reduced potential function  $V_{NN}$ . Furthermore, to verify the reduced potential function proposed by Mototake et al., focus was placed on the gradients of  $V_{NN}$  at time  $T_{init}$ , which indicates the onset of the negative magnetic domains. Time  $T_{init}$  was defined as  $T_{init} = \min \{t \mid \min \text{matrix}(X_t) < 0\}$ , where  $X_t$  represents the matrix of the magnetic domain pattern image at time  $t$  and  $\min \text{matrix}$  is a function that returns the minimum element of the matrix. Figure 2(C) illustrates the relationship between



anisotropy  $\alpha$  and the gradients of  $V_{\text{NN}}$  at time  $T_{\text{init}}$ . As shown in figure 2(C), the gradients of  $V_{\text{NN}}$  change continuously with the variations in anisotropy  $\alpha$ .

## 5 Discussion

Mototake et al. analyzed data and accordingly constructed a reduced potential function to effectively explain the two labyrinth structures in magnetic pattern-formation process. Evidently, the reduced potential function qualitatively transitioned with changes in anisotropy [12]. However, the analysis results obtained in this study (section 4) suggest that the gradient of  $V_{\text{NN}}$  with respect to changes in anisotropy  $\alpha$  varies continuously. Therefore, the data-driven reduced potential function  $V_{\text{NN}}$  does not exhibit discontinuous transitions with changes in anisotropy. Furthermore, the reduced potential function proposed by Mototake et al. does not consider the intermediate structure between the two labyrinth structures, namely the island structure (or a mixture of labyrinth and island structures), which exists during the formation process with varying anisotropy. The results obtained herein (section 4) imply that their model does not sufficiently explain this phenomenon. Thus, their proposed model must be modified to ensure that it considers the intermediate structure between the two labyrinth structures.

## Acknowledgments and Disclosure of Funding

The work was supported by JST PRESTO (JPMJPR212A), NEDO (Grant No. JPNP22100843-0), and JSPS KAKENHI (22K13979, 23H03460) and was partly supported by JST KAKENHI JP19H04982, and JST CREST JPMJCR1861, JAPAN.

## References

- [1] Edward H. Adelson and James R. Bergen. Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A*, 2(2):284–299, Feb 1985. doi: 10.1364/JOSAA.2.000284. URL <https://opg.optica.org/josaa/abstract.cfm?URI=josaa-2-2-284>.
- [2] David Bohm and David Pines. A collective description of electron interactions. i. magnetic interactions. *Phys. Rev.*, 82:625–634, Jun 1951. doi: 10.1103/PhysRev.82.625. URL <https://link.aps.org/doi/10.1103/PhysRev.82.625>.
- [3] John W. Cahn and John E. Hilliard. Free energy of a nonuniform system. i. interfacial free energy. *The Journal of Chemical Physics*, 28(2):258–267, 1958. doi: 10.1063/1.1744102. URL <https://doi.org/10.1063/1.1744102>.
- [4] Boyuan Chen, Kuang Huang, Sunand Raghupathi, Ishaan Chandratreya, Qiang Du, and Hod Lipson. Automated discovery of fundamental variables hidden in experimental data. *Nature Computational Science*, 2(7):433–442, 2022. URL <https://doi.org/10.1038/s43588-022-00281-6>.
- [5] Samuel Greydanus, Misko Dzamba, and Jason Yosinski. Hamiltonian neural networks. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32, pages 15379–15389. Curran Associates, Inc., 2019. URL <https://proceedings.neurips.cc/paper/2019/file/26cd8ecadce0d4efd6cc8a8725cbd1f8-Paper.pdf>.
- [6] G. E. Hinton and R. R. Salakhutdinov. Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507, 2006. doi: 10.1126/science.1127647. URL <https://www.science.org/doi/abs/10.1126/science.1127647>.
- [7] E. A. Jagla. Numerical simulations of two-dimensional magnetic domain patterns. *Phys. Rev. E*, 70:046204, Oct 2004. doi: 10.1103/PhysRevE.70.046204. URL <https://link.aps.org/doi/10.1103/PhysRevE.70.046204>.
- [8] E. A. Jagla. Hysteresis loops of magnetic thin films with perpendicular anisotropy. *Phys. Rev. B*, 72:094406, Sep 2005. doi: 10.1103/PhysRevB.72.094406. URL <https://link.aps.org/doi/10.1103/PhysRevB.72.094406>.

- [9] Ryo Kobayashi. Modeling and numerical simulations of dendritic crystal growth. *Physica D: Nonlinear Phenomena*, 63(3):410–423, 1993. ISSN 0167-2789. doi: [https://doi.org/10.1016/0167-2789\(93\)90120-P](https://doi.org/10.1016/0167-2789(93)90120-P). URL <https://www.sciencedirect.com/science/article/pii/016727899390120P>.
- [10] Kazue Kudo, Michinobu Mino, and Katsuhiro Nakamura. Magnetic domain patterns depending on the sweeping rate of magnetic fields. *Journal of the Physical Society of Japan*, 76(1), January 2007. ISSN 0031-9015. doi: 10.1143/JPSJ.76.013002. URL <https://doi.org/10.1143/JPSJ.76.013002>.
- [11] Lev Davidovich Landau and V L Ginzburg. On the theory of superconductivity. *Zh. Eksp. Teor. Fiz.*, 20:1064, 1950. URL <http://cds.cern.ch/record/486430>.
- [12] Yoh-ichi Mototake, Masaichiro Mizumaki, Kazue Kudo, and Kenji Fukumizu. Revealing the mechanism of magnetic domain formation by topological data analysis, 2022. URL <https://arxiv.org/abs/2204.12194>.
- [13] Ingo Steinbach. Phase-field models in materials science. *Modelling and Simulation in Materials Science and Engineering*, 17(7):073001, jul 2009. doi: 10.1088/0965-0393/17/7/073001. URL <https://dx.doi.org/10.1088/0965-0393/17/7/073001>.
- [14] Sin-itiro Tomonaga. Remarks on Bloch’s Method of Sound Waves applied to Many-Fermion Problems. *Progress of Theoretical Physics*, 5(4):544–569, 07 1950. ISSN 0033-068X. doi: 10.1143/ptp/5.4.544. URL <https://doi.org/10.1143/ptp/5.4.544>.
- [15] Alan Mathison Turing. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 237(641):37–72, 1952. URL <https://doi.org/10.1098/rstb.1952.0012>.