Disorder Identification in Hysteresis Data: Recognition Analysis of the Random-Bond–Random-Field Ising Model

O. S. Ovchinnikov,[1] S. Jesse,[2] P. Bintacchit,[3] S. Trolier-McKinstry,[3] and S. V. Kalinin[2,*]

[1]*Department of Physics and Astronomy, University of Tennessee, Knoxville, Tennessee 37996, USA*

[2]*Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA*

[3]*Department of Materials Science and Engineering and Materials Research Institute, Pennsylvania State University, University Park, Pennsylvania 16802, USA*

(Received 26 May 2009; revised manuscript received 31 August 2009; published 9 October 2009)

An approach for the direct identification of disorder type and strength in physical systems based on recognition analysis of hysteresis loop shape is developed. A large number of theoretical examples uniformly distributed in the parameter space of the system is generated and is decorrelated using principal component analysis (PCA). The PCA components are used to train a feed-forward neural network using the model parameters as targets. The trained network is used to analyze hysteresis loops for the investigated system. The approach is demonstrated using a 2D random-bond–random-field Ising model, and polarization switching in polycrystalline ferroelectric capacitors.

DOI: 10.1103/PhysRevLett.103.157203 PACS numbers: 75.10.Hk, 61.43.Bn, 75.60.−d, 84.35.+i

Random and disordered systems are ubiquitously present in physics and materials science [1,2], and in fact constitute the vast majority of real world materials. Correspondingly, understanding of the emergence of macroscopic deterministic properties from local dynamics is one of the key challenges in condensed matter physics. The development of physics has given rise to a number of well-established statistical models describing local dynamics of a system of interacting spins. Depending on the character of the spins, interactions, and dimensionality of the system, as described by corresponding Hamiltonians, a number of universal behaviors emerge [3]. Remarkably, even the simplest models have defied analytical description and can be studied only numerically [4].

One of the crucial tasks in the physics of disordered systems is the identification of the model and disorder types in real materials, necessary to elucidate the microscopic mechanisms underlying macroscopic behaviors. In most cases, this information is based on matching the known information on temperature and field dependence of the ground state to a phase diagram of the corresponding model. In addition, scattering techniques can be used to elucidate the time and temperature dependence of the corresponding correlation functions. Recently, significant attention was directed towards disorder identification from hysteresis loops and first-order reversal curves (i.e., families of minor loops) [5–8] However, this approach lacks the capacity to refine the model parameters based on experimental input.

Here, we develop an approach for the *quantitative* analysis of hysteresis loops to extract the type and strength of disorder present in the system, and demonstrate it for an example of a random-bond–random-field (RB-RF) 2D Ising model. This neural network approach operates in a manner reminiscent of ''recognition'' in human brain, combining the universality of associative thinking with the precision of a mathematical model.

As an example, we choose the generalized random-bond–random-field Ising model. The Hamiltonian is defined as

$$H(H) = \sum_{i,j} J_{ij} S_i S_j + \sum_i (h_i + H) S_i, \qquad (1)$$

where $S_i = \pm 1$ are the local spins, J_{ij} are the nearest-neighbor interactions, h_i are the random fields, and H is the external field. The nearest-neighbor interactions have a Gaussian distribution with an average, J_0, and width, δJ, i.e., $P(J_{ij}) = (\delta J\sqrt{2\pi})^{-1} \exp(-(J_{ij} - J_0)^2/\delta J^2)$. The random-field components are assumed to have a Gaussian distribution with width δh, i.e., $P(h_i) = (\delta h\sqrt{2\pi})^{-1} \exp(-h_i^2/\delta h^2)$. For $\delta h = 0$ and $\delta J = 0$, the free energy Eq. (1) corresponds to the simple ferromagnetic ($J_0 > 0$) or antiferromagentic ($J_0 < 0$) Ising model. For $\delta h = 0$ and $J_0 = 0$ Eq. (1) corresponds to Edwards-Anderson spin-glass model. Finally, for $\delta J = 0$, Eq. (1) corresponds to a random-field Ising model. The dynamics of the system described by Eq. (1) has been extensively studied as a function of dimensionality and disorder type [2–4].

For given model parameters (J_0, δJ, δh), the evolution of the system is studied on a 2D (10×10) field using Glauber dynamics. The simulation yields a single hysteresis loop $S(H)$ as a function of model parameters, (J_0, δJ, δh). To extract the model parameters from the hysteresis curve, we utilize the nonparametric deconvolution method based on a combination of principal value decomposition with neural network interpolation [9,10] as shown in Fig. 1(a). The family of theoretical hysteresis loops $S_H(J_0, \delta J, \delta h)$ is generated in the parameter interval $J_0 \in (J_{\min}, J_{\max})$, $\delta J \in (\delta J_{\min}, \delta J_{\max})$, and $\delta h \in (\delta h_{\min}, \delta h_{\max})$,

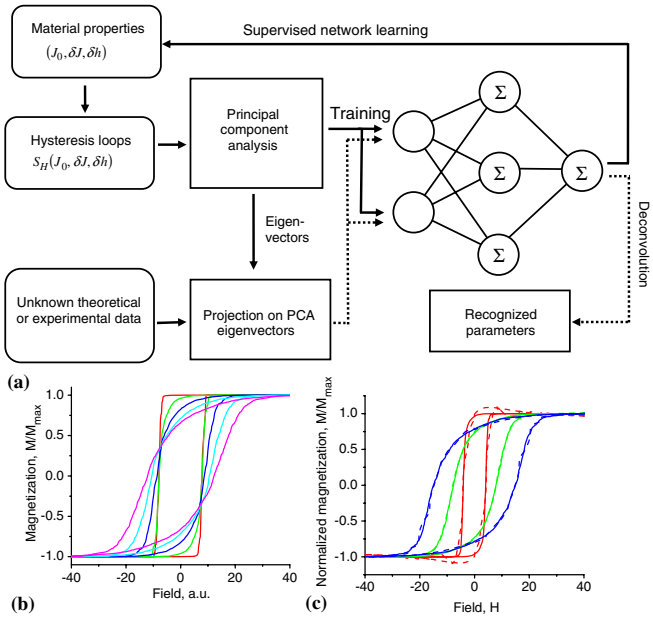

FIG. 1 (color online). (a) Schematics of neural network recognition analysis. (b) Hysteresis loops for random-bond model, $\delta h = 0$, for $J_0 = 2$ and $\delta J = 1, 2, 3, 4, 5$. (c) Reconstructed hysteresis loop using the 1st 6 eigenvalues for $(J_0, \delta J) = (1, 1)$ (red), (1.5, 3) (green), and (3, 5) (blue). Shown are the original hysteresis loop (solid) and the PCA reconstruction (dashed).

where $\delta J_{\min}$ and $\delta h_{\min}$ are positively defined. Shown in Fig. 1(b) is the example of hysteresis loops in random-bond case, $\delta h = 0$, illustrating systematic change in loop shape as a function of model parameters.

The family of hysteresis loops in a given subset of parameter space is decorrelated using principal component analysis (PCA) [11]. In PCA, the set of N hysteresis loops containing P points is represented as a superposition of the eigenvectors w_j, $S_m(H_j) = a_{mk} w_k(H_j)$, where $a_{mk} = a_k(J_0, \delta J, \delta h)$ are referred to as loading coefficients and $S_m(H_j) \equiv S(J_0, \delta J, \delta h, H_j)$ is the magnetization value for field H_j (separate for forward and reverse curves). The eigenvectors $w_k(H)$ and the corresponding eigenvalues λ_k are found from the covariance matrix, $\mathbf{C} = \mathbf{S}\mathbf{S}^T$, where $\mathbf{S}$ is the matrix of all experimental data points S_{mj}. The rows of $\mathbf{S}$ correspond to parameter variation (m defines total number of hysteresis loops in a set), and columns correspond to field points, $j = 1, \ldots, P$.

The eigenvectors $w_k(H)$ are orthonormal and are chosen such that the corresponding eigenvalues are placed in descending order, $\lambda_1 > \lambda_2 > \ldots$. In other words, the first eigenvector $w_1(H)$ contains the most information (defined as the largest variation) within the data set, the second contains the most "informative" response after subtraction of the first, and so on. This process is highly reminiscent of Gram-Schmidt orthogonalization, but with the norm defined in a statistical sense. Mathematically, the eigenvalues and corresponding eigenvectors are determined through singular value decomposition of the $\mathbf{S}$ matrix. The number of significant values, m, can be chosen based on the overall shape of $\lambda_k(k)$ dependence or on correlations in the parameter or real space (for SPM data) [12]. The comparison between the hysteresis loops and PCA reconstruction is shown in Fig. 1(c), illustrating that even for a small number of PCA components the hysteresis loop shape can be reproduced.

The determined loading vector is used to train a feed-forward neural network (NN) using the set of $a_k(J_0, \delta J, \delta h)$ as inputs and J_0, δJ, δh as desired outputs. A trained neural network acts as a universal interpolator that establishes a relationship between the hysteresis loop described as a superposition of linearly independent $w_k(H_j)$ with coefficients $a_k(J_0, \delta J, \delta h)$, and model parameters J_0, δJ, δh. On the analysis stage, the same set of $w_k(H_j)$ is used to project unknown experimental or theoretical hysteresis loops $S_{\exp}(H)$ on to a set of $\alpha_{i_{\exp}}$ values. The $\alpha_{i_{\exp}}$ are then fed into the trained neural net to extract model-dependent J_0, δJ, δh. This recognition approach thus solves the inverse problem of Eq. (1), i.e., identification of disorder from the hysteresis loop shape.

The recognition analysis was performed on a several two parametric and a full RB-RF Ising model, when all three parameters J_0, δJ and δh were varied. Overall, 726 curves corresponding to $J_0 = 0, \ldots, 5$ with step 0.5, $\delta J = 0, \ldots, 5$ with step 1, and $\delta h = 0, \ldots, 20$ with a step of 2 was generated. The principal components were used to train a (10, 24, 2) feed-forward neural network with a tansig transfer function in the hidden layer and a linear transfer function in the output layer [13]. For optimal fitting, the overall error decreased by ~1.5–2 orders of magnitude compared to a random input. After training, the validity of network training was checked using the simulated array, and a comparison of the network output and model parameters is shown in Fig. 2.

The number of relevant principal components and the structure of the neural network was optimized using the PCA component maps and convergence behavior, respectively. The recognition results for the J_0 are shown in Fig. 2. The reconstructed J_0 values show a relatively broad scatter around the actual values (standard deviation of ~0.5, corresponding to relative error of ~10%) [Fig. 2(b)]. However, note that the reconstructed values represent the collective effect of all possible δJ and δh values on the reconstruction.

The structure of corresponding error surfaces, defined as $Er(P) = (P_{\text{sim}} - P_{\text{real}})/(P_{\text{sim}} + P_{\text{real}}) \times 100\%$, is shown in Fig. 2(c) and 2(d). Generally, the maximal relative error is observed for extreme parameter values, and the δJ error is independent of the actual δJ. For the chosen parameter space, the reconstruction error is ± 0.5 for the δJ changing from 0 to 5. As a further example of recognition analysis, several 2 parameter Ising models, including a frustrated spin-glass model ($J_0 = 0$ and varying δJ and δh), a random-bond Ising model ($\delta h = 0$) and a random-field

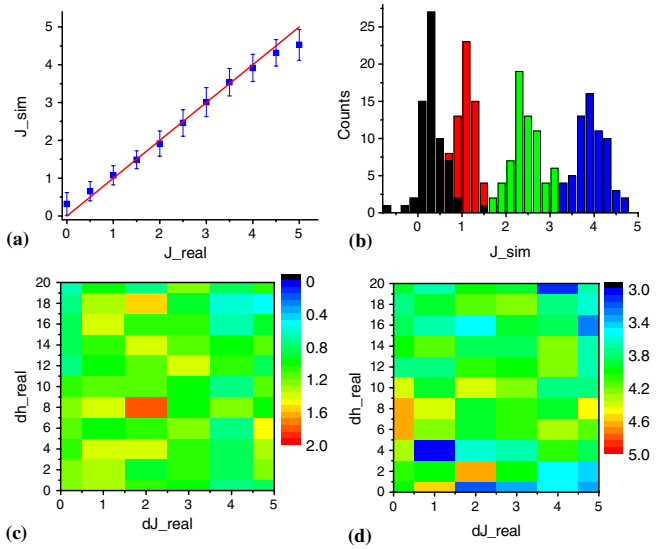

FIG. 2 (color online). Neural network reconstruction of 3—parameter Ising model for width-normalized data. (a) Reconstruction plot for J_0. Error bars are defined as the standard deviation of reconstructed values. (b) Histogram for recognized J_0 values for actual $J_0 = 0$, 1, 2.5, and 4. The distribution of reconstructed J_0 for (c) actual $J_0 = 1$ and (d) actual $J_0 = 4$ as a function of actual random bond, δJ, and random field, δh.

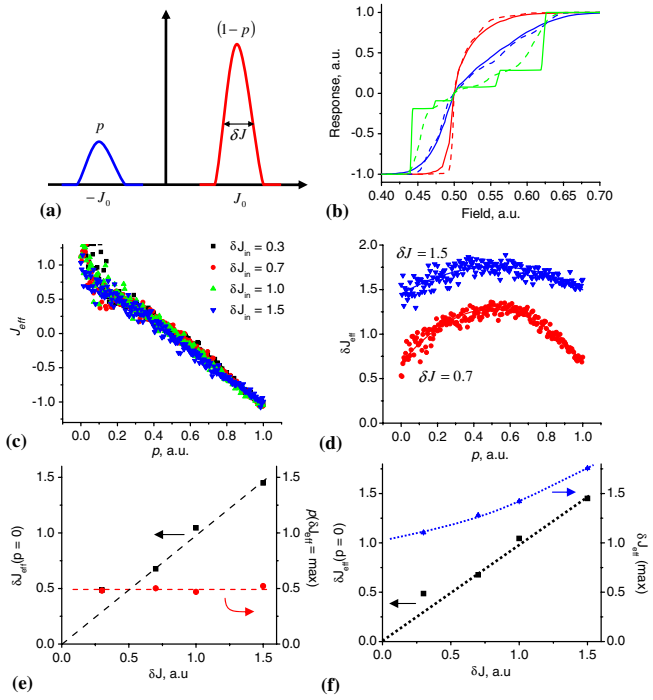

FIG. 3 (color online). (a) Distribution of exchange integrals in the p model. (b) Modeled (solid line) and reconstructed (dotted line) hysteresis loops for $\delta J = 0.3$ and $p = 0.2$ (red), 0.5 (blue), and 1 (green). (c) Reconstructed exchange integral $J_{\mathrm{eff}}(p)$ and (d) reconstructed RB disorder $\delta J_{\mathrm{eff}}(p)$ vs flip probability, p. (e), (f) Dependence of model parameters on the RB disorder δJ.

Ising model ($\delta J = 0$) were studied, illustrating reconstruction errors of ~3–5% over 2D parameter spaces (see supplementary material) [14]

To establish the applicability of the proposed recognition method to an unknown statistical physics model, we have studied the behavior of the p-model as shown in Fig. 3(a). In this model, the RF-RB Ising model is defined, and the random bonds are flipped from positive to negative value with probability, p. Here, we chose J_0 equal to 1 and δJ of 0.1, 0.3, 0.7, 1, and 1.5. The resulting distribution of exchange integrals is shown in Fig. 3(a). The probability of a flip was varied from 0 to 1 with a 0.05 step. The resulting hysteresis loops were then analyzed by a neural network that had been trained using a family with J_0 spanning -2 to 2 and δJ spanning 0 to 2.

The results of the fitting for J_0 and δJ are shown in Fig. 3(b). The recognition fitting allows reconstruction of the hysteresis loops in the interval of J from strong ferromagnetic to antiferromagnetic through the spin-glass state (except for $p = 1$, for which sharp jumps in response are observed). This indicated that hysteretic behavior of the p model is equivalent to an RB Ising model with Gaussian distribution of RB components with some effective exchange integral $J_{\mathrm{eff}}(p)$ and distribution width $\delta J_{\mathrm{eff}}(p)$.

The effective exchange integral $J_{\mathrm{eff}}(p)$ decreases linearly with the percent chance of flip almost independently of δJ (except for $\delta J = 0$, in which the choice of Glauber dynamics leads to unphysically sharp features in the loop) [Fig. 3(c)]. As expected, $J_{\mathrm{eff}} \approx J_0(1 - 2p)$. The $\delta J_{\mathrm{eff}}(p)$ depends on flip probability, p, approximately as $\delta J_{\mathrm{eff}} = a + b(p - p_0)^2$ [Fig. 3(d)]. The RB disorder is minimal and equal to the δJ_0 at the edges of the interval and maximal for $p_0 \approx 0.5$. The dependence of p_0 (flip probability corresponding to maximum disorder) and $\delta J_{\mathrm{eff}}(p = 0)$ on δJ is shown in Fig. 3(e), and demonstrates the expected behavior of $\delta J_{\mathrm{eff}}(p = 0) \approx \delta J$ and $p_0 \approx 0.5$. Finally, the maximum disorder $\delta J_{\mathrm{eff}}(p = p_0)$ is compared to $\delta J_{\mathrm{eff}}(p = 0)$ in Fig. 3(f), illustrating the role of bond flip on the effective model parameters.

To explore the applicability of this approach to the experimental data, we have chosen a polycrystalline ferroelectric capacitor as a model system. The capacitor used in this study is a high density $PbZr_{0.52}Ti_{0.48}O_3$ (PZT) films prepared by chemical solution deposition, and has mixed {001} and {111} orientation with columnar grains [15]. Spatially resolved studies of similar materials without top electrodes have demonstrated that typically the polarization switching occurs uniformly within the grains [16] and only 180° switching is observed [17]. These considerations suggest that the salient features of system behavior can be approximated by random-field random-bond 2D Ising model, with the domains inside individual columnar grains playing the role of Ising spins interacting through short range electrostatic interactions, random electric fields due to charged defects and interfaces playing the role of random fields, and variability of grain-grain coupling and

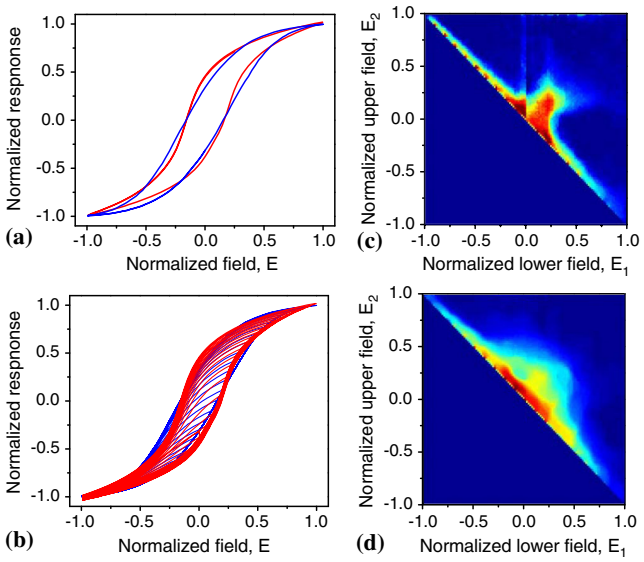

FIG. 4 (color online). (a) Comparison of the experimental hysteresis loop for ferroelectric capacitor (red) and corresponding fit for RBRF 2D Ising model (blue). (b) Comparison of the first-order reversal curve families. Preisach density maps for (c) capacitor data and (d) Ising model fit.

compositional inhomogeneities playing the role of random-bond component.

The macroscopic $P - E$ hysteresis loops and a full set of the first-order reversal curves (i.e., the family of minor hysteresis loops starting from a polarized state) are shown in Figs. 4(a) and 4(b). The corresponding Preisach densities [18] are shown in Figs. 4(c) and 4(d). The use of the recognition algorithm developed above has yielded model parameters as $J_{\mathrm{cap}} = 1.4$, $\delta J_{\mathrm{cap}} = 4.8$ and $\delta h_{\mathrm{cap}} = 20.4$. To verify the network fit, the direct search of minimal error over the training set yields $J_{\mathrm{cap}} = 2$, and $\delta J_{\mathrm{cap}} = 4$ and $\delta h_{\mathrm{cap}} = 20$, i.e., an identical result (within the error due to the discrete nature of training set). Note the close agreement between the experimental loop and the fit, as well as good agreement between the corresponding first-order reversal curves and Preisach densities that offer a powerful tool for comparison of hysteresis behavior. Hence, despite the fact that the description of capacitor by Ising model is nonideal, the overall features, including the distribution of the reversible components and fraction of irreversible polarization are similar, suggesting a good agreement between the two.

To summarize, a universal approach for the analysis of disorder type and strength in physical systems based on the analysis of hysteresis loops is developed. The approach is based on the neural network recognition of the decorrelated data set, where the network was first trained using a large number of examples uniformly distributed in the parameter space of the system. The operation of this approach is demonstrated for a 2D random-bond–random-field Ising model, and generally allows for high-precision (~2%–5% error) for 2, and ~20%–30% error for 3 free parameters. The approach is universal and can be applied to other statistical physical models as well as types of collected information, and thus represents an alternative to traditional functional fitting using least mean square or other algorithms.

The work was supported by the ORNL LDRD program (S. V. K, S. J.) and the ORNL HERE program (O. O). Funding for work at Penn State was supplied by the Center for Dielectric Studies, the National Security Science and Engineering Faculty program, and a Royal Thai Government grant (P. B).

*Corresponding author.
sergei2@ornl.gov

[1] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, Cambridge, England, 2000).
[2] J. P. Sethna, *Statistical Mechanics: Entropy, Order Parameters, and Complexity* (Oxford University Press, New York, 2006).
[3] K. Binder and A. P. Young, Rev. Mod. Phys. **58**, 801 (1986).
[4] K. Binder, Rep. Prog. Phys. **60**, 487 (1997).
[5] H. G. Katzgraber, G. T. Zimanyi, Phys. Rev. B **74**, 020405 (2006).
[6] I. D. Mayergoyz, Phys. Rev. Lett. **56**, 1518 (1986).
[7] J. G. Ramírez, A. Sharoni, Y. Dubi, M. E. Gomez, and I. K. Schuller, arXiv:0811.2206.
[8] H. G. Katzgraber, D. Herisson, and M. Östh *et al.*, Phys. Rev. B **76**, 092408 (2007).
[9] Z. Chen, S. Haykin, J. J. Eggermont, and S. Becker, *Correlative Learning: A Basis for Brain and Adaptive Systems* (Wiley, Hoboken, 2007).
[10] S. Haykin, *Neural Networks: A Comprehensive Foundation* (MacMillan, London, 2008).
[11] I. T. Jolliffe, *Principal Component Analysis* (Springer, New York, 2002).
[12] S. Jesse and S. V. Kalinin, Nanotechnology **20**, 085714 (2009).
[13] S. Samarasinghe, *Neural Networks for Applied Sciences and Engineering* (Auerbach, Boca Raton, 2006).
[14] See EPAPS Document No. E-PRLTAO-103-028943 for supplementary information on the recognition analysis for the 2D parameter set. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.
[15] K. Seal, S. Jesse, S. V. Kalinin, and I. Fujii *et al.*, Phys. Rev. Lett. **103**, 057601 (2009).
[16] V. Nagarajan, S. Aggarwal, and A. Gruverman *et al.*, Appl. Phys. Lett. **86**, 262910 (2005).
[17] F. Xu, S. Trolier-McKinstry, and W. Ren *et al.*, J. Appl. Phys. **89**, 1336 (2001).
[18] F. Preisach, Z. Phys. **94**, 277 (1935).