To be submitted to Journal of Applied Physics

Reconstruction and uncertainty quantification of lattice Hamiltonian model parameters from observations of microscopic degrees of freedom

Mani Valleti,[1,1] L. Vlcek,[2,3] Maxim Ziatdinov,[4] Rama K. Vasudevan,[4] and Sergei V. Kalinin[4,2]

[1] Bredesen Center for Interdisciplinary Research, University of Tennessee, Knoxville, TN 37996, USA

[2] Joint Institute for Computational Sciences, University of Tennessee, Knoxville, Oak Ridge, TN 37831, USA

[3] Materials Science and Technology Division, Oak Ridge National Laboratory, Oak Ridge TN 37831, USA

[4] The Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, TN 37831

Abstract

The emergence of scanning probe and electron beam imaging techniques have allowed quantitative studies of atomic structure and minute details of electronic and vibrational structure on the level of individual atomic units. These microscopic descriptors in turn can be associated with local symmetry breaking phenomena, representing the stochastic manifestation of the underpinning generative physical model. Here, we explore the reconstruction of exchange integrals in the Hamiltonian for a lattice model with two competing interactions from observations of microscopic degrees of freedom and establish the uncertainties and reliability of such analysis in a broad parameter-temperature space. In contrast to other approaches, we specifically specify a loss function inherent to thermodynamic systems and utilize it to estimate uncertainty in simulated

---
[1] Corresponding author, svalleti@vols.utk.edu
[2] Corresponding author, sergei2@ornl.gov

realizations of different models. As an ancillary task, we develop a machine learning approach based on histogram clustering to predict phase diagrams efficiently using a reduced descriptor space. We further demonstrate that reconstruction is possible well above the phase transition and in the regions of the parameter space when the macroscopic ground state of the system is poorly defined due to frustrated interactions. This suggests that this approach can be applied to the traditionally complex problems of condensed matter physics such as ferroelectric relaxors and morphotropic phase boundary systems, spin and cluster glasses, and quantum systems once the local descriptors linked to the relevant physical behaviors are known.

1. Introduction

Phase transitions between dissimilar phases underpin multiple areas of condensed matter physics and materials science. From an applications viewpoint, phase transitions are intricately involved in the process of material formation in virtually all technologies from ceramics to metals to semiconductors. Electrical, gating, and chemical control of phase transitions have emerged as a promising pathway for low-voltage electronics in tunneling barriers and field effect devices.[1, 2] Beyond condensed matter physics, phase transitions in Bose-Einstein condensates allow new classes of quantum information technology devices.[3]

The ubiquity of the phase transitions in virtually all areas of applied and fundamental science have made them one of the central areas for theoretical exploration. Among the existing paradigms for exploring phase transitions, a special place is held by the lattice models such as Ising[4], Heisenberg[4], Kitaev[5], etc. Here, the system is represented as a set of interacting spins on a spatial lattice. The spins represent the individual degrees of freedom of the material and can be magnetic spins, electrical dipoles, atomic species, etc. The lattice defines the geometry of the interactions. Depending on the conservation laws, character of the spins, interactions, and geometry, these lattice models can represent an extremely broad range of phenomena from ordering in metal alloys to ferroic transitions to surface adsorbate dynamics.

The universality of the lattice model approach has spurred intensive analytic and numerical investigations of these models.[6, 7] These studies mainly seek to explore the phase diagrams of the lattice models, i.e. correspondence between the global variables such as temperature and field and preponderant behaviors and patterns for spin arrangement. Similarly of interest are global thermodynamics parameters such as magnetization, heat capacity, and susceptibility. The spatial arrangements of the spins are traditionally explored in the form of the ground state pattern, as well as correlation functions representing the distributions of the defects and disorder, i.e. over which length scales the system is ordered[8]. Notably, these behaviors allow for straightforward exploration only for systems with well-defined ground states, whereas the presence of the competing interactions leads to frustrated and poorly defined ground states. Nonetheless, the combination of efficient sampling strategies and high-performance computing have allowed exploration even complex multidimensional models.

Over recent years, the emergence of machine learning (ML) tools and readily availability of information-compression algorithms, such as principal and independent component analysis,

support vector machines, as well as deep learning tools such as variational autoencoders have provided new impetus within the field.[9-12] Even though the applications of ML tools to lattice models are relatively recent, the field has experienced an exponential growth in the last 3 years, opening pathways to exploration of exotic quantum phases, off-lattice models, etc.[13-16]

However, experimental selection of the appropriate lattice model corresponding to a specific phase transition and determination of its parameters, termed the "inverse Ising problem," presents a more difficult problem[17-20]; an in-depth view of that field is explored in an article by Nguyen et al.[21] Traditionally, the symmetry and type of order parameter are determined based on macroscopic measurements and scattering studies, whereas numerical values of exchange integrals are obtained based on the matching between theoretical predictions and macroscopic measurements of thermodynamic parameters. However, these studies typically require high quality samples for which the corresponding critical behaviors can be reliably determined. Similarly, in the materials with high disorder, non-ergodic or non-thermalized ground states determination of appropriate lattice models can be highly non-trivial.

The development of high-resolution imaging tools such as scanning transmission electron microscopy and scanning probe microscopy have allowed direct insight into the atomic-scale structure of materials.[22-24] Beyond atomically resolved and element specific imaging, these techniques provide information on the minute symmetry-breaking distortions, allowing for direct mapping of order parameter fields such as polarization,[25-29] octahedra tilts,[30-33] strains,[34] and more complex emergent phenomena.[35] Correspondingly, of interest is the extraction of the parameters of the underpinning physics model from the atomically resolved observations. As one approach, this can be achieved using the direct matching of the mesoscopic order parameter field to Ginzburg-Landau type models.[36, 37] Alternatively, the mesoscopic descriptors such as hysteresis loops or relaxation curves can be used as proxy identifiers for local behavior.[38] Finally, it has been proposed that direct observation of the mesoscopic degrees of freedom can be directly compared to lattice model via statistical distance minimization[39-44]. Indeed, naively expecting to minimize loss between theoretical and experimental datasets without the appropriate metric for thermodynamic systems is incorrect and will lead to incorrect reconstructions.[45] Previously, we have reported the principles of the statistical distance minimization,[41, 42] as well as applications for establishing of the parameters of the non-ideal solid solutions in layered superconductors,[43] information fusion between surface and depth-resolved chemical information in manganites,[44] and

pair- and triple-atom model comparison for segregation in layered chalcogenides[40]. Here, we explore the reconstruction of the Ising model parameters in the presence of the competing interactions with statistical distance minimization and explore the sensitivity of the reconstruction to the incomplete knowledge of the model, as well as deriving uncertainty bounds on the reconstructions. The guidelines for experimentally driven model selection are proposed.

2. Model and phase diagram

We explore the behavior of the Ising model with the nearest neighbor and next nearest neighbor interactions. Depending on the sign and magnitude of the exchange integrals, this model can give rise to a rich phase diagram containing ferromagnetic, antiferromagnetic, paramagnetic, and frustrated phases. Here, first we apply the ML tools for rapid simulation of the associate phase diagram based on the local statistics.

2.1. Model and simulation

The classical Ising Hamiltonian model realized on N^2 -evenly spaced lattice sites of a square lattice with nearest neighbor (NN) and next nearest neighbor (NNN) interactions is the chosen model for this study. The Hamiltonian of a given configuration σ in the absence of external magnetic field is given by equation (1).

$$H(\sigma_i) = - \sum_{<i,j>} (J_1 + \Delta J_1)\sigma_i\sigma_j - \sum_{<i,k>} (J_2 + \Delta J_2)\sigma_i\sigma_k \tag{1}$$

where J_1 and J_2 are the exchange integrals corresponding to the nearest neighbor interactions and next nearest neighbor interactions respectively, ΔJ_1 and ΔJ_2 are the corresponding disorders in exchange integrals, $<i,j>$ and $<i,k>$ are the sums over all nearest neighbors and next nearest neighbors respectively.

A Monte Carlo (MC) simulation was performed on a square lattice with $N = 40$ using the Metropolis algorithm. The system was equilibrated for $1000 \times N^2$ Monte Carlo steps and the data is acquired on the next $1000 \times N^2$ Monte Carlo steps. A spin configuration reversal of each lattice site was attempted at every step. The flip was accepted when the energy of the resultant configuration was lower than the previous step. Else, the probability of spin flip was estimated

using the Boltzmann distribution given by equation (2), where $\beta = 1/k_B T$ is the inverse temperature, with T the absolute temperature and k_B the Boltzmann constant, and denominator is the partition function.

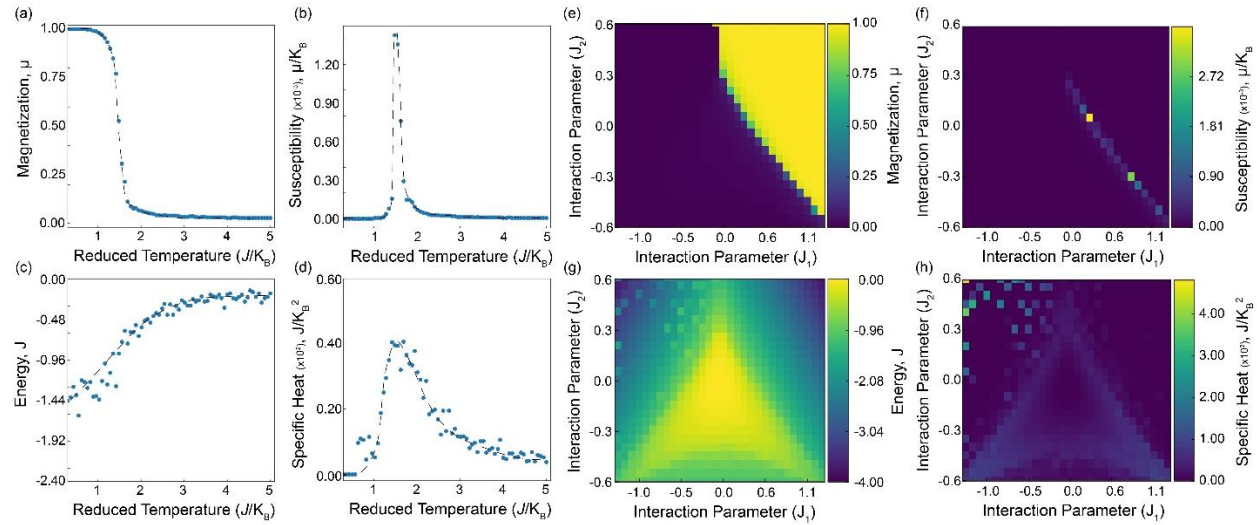$$P_\beta(\sigma_i) = \frac{e^{-\beta H(\sigma_i)}}{\sum_j e^{-\beta H(\sigma_j)}}$$

(2)

Figure 1. (a) Magnetization, (b) Susceptibility, (c) Energy, and (d) Specific Heat curves as a function of reduced temperature for the traditional 2D Ising model on a square lattice with NN interactions only ($J_1 = 0.65$ and $J_2 = 0$). (e) Magnetization, (f) Susceptibility, (g) Energy, and (h) Specific Heat surfaces of the square lattice Ising model at a reduced temperature $T_r = 0.96$ as function of exchange integrals corresponding to nearest (J_1) and next nearest neighbors in the units of $k_B T$.

Macroscopic properties as a function of temperature and a function of exchange integrals are explored. Setting J_2 and ΔJ_2 to zero gives the traditional 2D Ising model with only the nearest neighbor (NN) interactions. Macroscopic properties of such model as a function of reduced temperature as shown in Fig. 1(a-d). Clearly all four properties delineate the phase transition from ferromagnetism to paramagnetic region. Phase transitions as a function of exchange integrals at a constant temperature are in shown in Fig. 1 (e-f). Energy per site is illustrated the in Fig. 1g. The corresponding magnetization is shown in Fig. 1e, clearly highlighting the ferromagnetic region and the transition between the ferromagnetic and paramagnetic states (diagonal line). The specific heat is shown in the Fig. 1h, delineating the primary phase transitions between ferromagnetic, antiferromagnetic, paramagnetic, and frustrated states. Finally, susceptibility is shown in Fig. 1f,

showing the divergence across ferromagnetic – paramagnetic phase transition. Overall, Fig. 1 provides a traditional approach for mapping the phase diagram of lattice models based on the macroscopic thermodynamic parameter analysis.

2.2. Machine learning analysis of the phase diagram

An alternative approach for mapping phase diagrams in simulated lattice models is based on the machine learning and statistical analysis of the local spin configurations and correlation between spins. For example, multiple works showed that using a supervised learning techniques such as convolutional neural networks, which are traditionally used for image recognition tasks, one can identify a phase transition and map phase boundaries in the classical and quantum lattice models using 2-dimensional and 3-dimensional "images" of the spin configurations around the critical points as a training set[9-11]. At the same time, the more conventional multivariate statistics and probabilistic learning tools such as kernelized principal component analysis and variational autoencoder allowed learning a phase transition and the associated order parameter in the classical Ising model from local spin configurations in an unsupervised manner[46, 47].

Traditional descriptors of such microstates are spin configurations of all the lattice points[48]. In this scenario, the number of rows (R) in the feature matrix are $R = N_{sim}* N_{mc}$, where N_{sim} is the number of distinct sets of simulation parameters and N_{mc} are the number of Monte Carlo steps considered for analysis. While columns (C) are the number of features used for the analysis. Spin configurations of the entire lattice at each simulation are used as attributes. The first few principal components corresponding to such matrix closely resemble the macroscopic properties' (magnetization, energy, specific heat and susceptibility) trends as a function of reduced temperature[48].

Here, we propose the approach where phase classification is performed based on the relative frequencies of local configurations at a given set of simulation parameters. These frequencies of local configurations describe the microstates of the system exhaustively. For a given case, these are averaged over all the microstates generated by the Monte Carlo simulations. The local configurations considered for further analysis are described in Fig 2e. The relative frequencies of local configurations of ferromagnetic, anti-ferromagnetic, paramagnetic and the frustrated systems are shown in Fig 2a-d. In this case, the number of elements of the feature matrix are 5400 as opposed to 10^9 in the study referred. Monte Carlo simulations are performed at a

reduced temperature of $T_r = 0.96$ and at 900 evenly spaced points on a grid space of exchange integrals $J_1 = [-1.25, 1.25]$ and $J_2 = [-0.75, 0.75]$ in the units of k_BT. From here on all the values of exchange integrals are reported in the units of k_BT. Phase diagrams of complex systems can be mapped using statistical analysis of the local structures values.
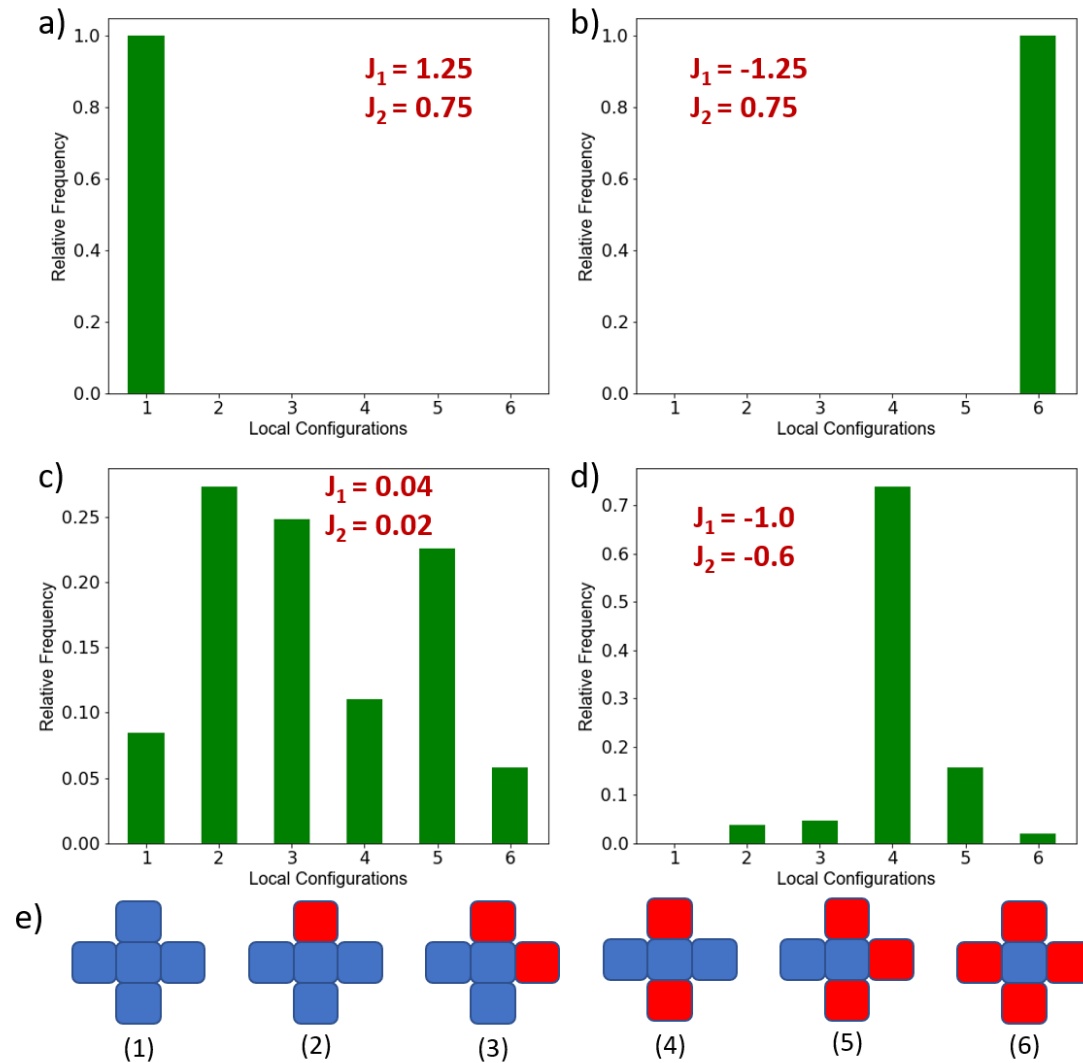

Figure 2. Representative relative frequencies of local configurations at reduced temperature $T_r = 0.96$ for (a) paramagnetic, (b) anti-ferromagnetic, (c) paramagnetic and (d) frustrated regimes of phase diagram. (e) Exhaustive set of six local configurations which serve as descriptors for a given microstate. Representative configurations for each regime are shown in Fig. 4d in the same order.

To establish the phase diagram of the system over the parameter space, we use simple k-means clustering approach. A similar analysis was proposed by Canabarro et al. in which the pairwise correlations among all spins are used as the descriptors[49]. This algorithm segregates the input data into k- clusters based on minimizing the variance within each cluster[50]. The number of clusters is *apriori* unknown but can be estimated based on the quality of separation and the target output. Here, when the feature matrix is segregated into 4 clusters using the k-Means clustering technique, the resulting clusters (Fig. 3a) exactly resembled the energy (Fig. 1a) and specific heat (Fig. 1c) phase diagram. In Fig.3a, phases – 1,2,3,4 represent the ferromagnetic, anti-ferromagnetic, frustrated, and paramagnetic regions respectively. Representative configurations for the four phases are plotted in Fig. 3d. Ferromagnetic and anti-ferromagnetic transitions to paramagnetic phase can clearly be observed in Fig. 3a. It is also observed that for k = 2, the clustering plot reproduced the magnetization and susceptibility phase diagrams. The clusters are also plotted in two-dimensional space of first two principal components[50] (Fig. 3b) and as a dendrogram (Fig. 3c) to show the hierarchy between the clusters. Fig. 3b illustrates the separability of the phases and their general configurations in reduced (principal component) space. This provides information on which phases are separated by continuous and discontinuous transitions. Dendrogram (Fig. 3b) shows the intra-cluster distances between different clusters in the 'histogram space'. The subsequent increase of the clusters number leads to the emergence of additional regions concentrated at the boundary between frustrated and paramagnetic states; however, primary ferromagnetic, antiferromagnetic, and frustrated regions remain invariant.
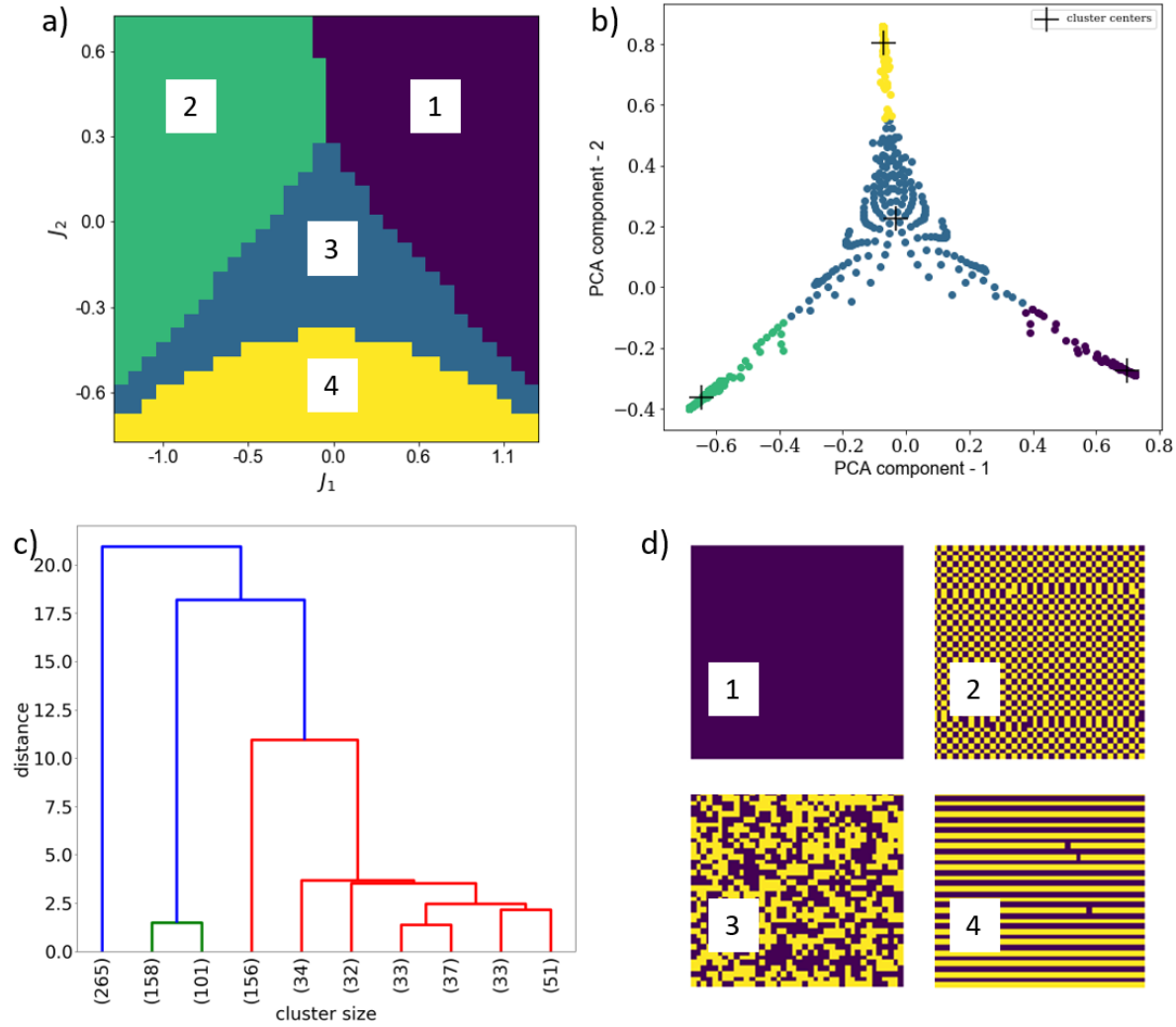
Figure 3. (a) Four clusters generated as a result of k-Means algorithm when relative frequencies of local configurations are used as descriptors of macro states. (b) Clusters and their centers plotted in 2-D space of first two principal components. (c) Dendrogram generated using hierarchical clustering. (d) Representative configurations that correspond to the four clusters formed.

We note that classically, exploration of the lattice models relies on the macroscopic thermodynamic descriptors such as those shown in Fig. 1, as well as structure factors and correlation functions. Here, we rapidly map the relevant phase diagram using the comparison of the local motifs. In addition to the obvious and previously reported applications for exploration of the phase diagrams of the lattice models, these observations suggest that the statistics of local spin configurations contain the information on the microscopic interactions in Eq. (1).

3. Results and Discussion

3.1. Reconstruction of interaction from microscopic degrees of freedom

Here, we extend the statistical distance minimization method to reconstruct the interaction in the Ising model within the full parameter space. During the reconstruction, we have access to the spin configurations (synthetic experimental data), but not the values of Ising model parameters (exchange integrals). Correspondingly, the goal of the reconstruction is to determine the exchange integrals (parameters of the generative physics model) based on the observations of local configurations and establish the associated uncertainties.

To determine the unknown model parameters, we run the Monte Carlo simulations at a given set of simulation parameters to generate the spin configurations of the base (pseudo-experimental or synthetic experimental data) case. Here, twenty snapshots from the last 400 steps of the MC simulations are used as experimental observables for the base case. To reconstruct the exchange integrals of a given base case, simulations were carried out on the entire grid space of the exchange integrals. The local configurations of the base case are then compared to the simulated cases using the statistical distance metric[51]. This metric is employed to quantify the distinction between pair thermodynamics systems and is given by

$$s = \arccos\left(\sum_{i=1}^{l} \sqrt{p_i}\sqrt{q_i}\right) \tag{3}$$

where p_i and q_i are the probabilities of finding a local configuration i in the base case and model simulations respectively, and the summation runs over all possible (6 in this case) local configurations. The probabilities of the local configurations are determined as the maximum likelihood estimates from the experimental (p) and simulated (q) histograms. Statistical hypothesis testing indicates that as the measured distance metric reduces to zero, the target and model system measurements become indistinguishable. A unique characteristic of this metric, which sets it apart from popular measures such as Kullback-Leibler divergence, is that it automatically incorporates the statistical weights of the collected data, thus allowing for separation of weak signals (viz. thermal fluctuations) from the statistical noise inherent in samples of a thermodynamic system.

Extracting these fluctuations from the data is critical for improving the predictive capabilities of the optimized model, as they encode the system's response to changes in external conditions.

Since sampling in any real experiment is limited, the key question of exactly how many samples are required to achieve a certain tolerance on the inferred parameters is an important one. Towards this aim, we consider an averaging method utilizing the samples at hand. In this case, we divided the parent data set of 400 images (last 400 MC steps of synthetic experimental data) into 20 sets of uncorrelated images. Interaction parameters are reconstructed for each subset at a given reduced temperature, and mean, standard deviation and the probability density function of the reconstructed values are reported.

3.2. Reconstruction of the full J_1 - J_2 model

First, we explore the reconstruction of the model parameters for the full model. The reconstruction results for all four quadrants of the phase diagram is shown in Figures 4 and 5. For all values of J_1 and J_2, the reconstruction shows that the statistical distance minimization yields an unbiased reconstruction of the model parameters. Here, two base cases ($J_1 = +0.75$, $J_2 = +0.35$ and $J_1 = -0.75$, $J_2 = +0.35$) were reconstructed that have a well-defined ground state where the nearest and next nearest neighbor interactions are not frustrated. The uncertainty is quantified using the experimental averaging technique discussed above.

The reconstruction fails below the phase transition region as there is no information in the spin configurations. Exchange integrals were reconstructed with high certainty at temperatures just above the phase transitions, where the configurations are strictly defined by the interaction parameters. The uncertainty of reconstruction increases with increasing temperature due to random spin configurations. Still, the reconstruction is possible for temperatures that are an order greater than the transition temperature before the uncertainty impedes the interpretation.
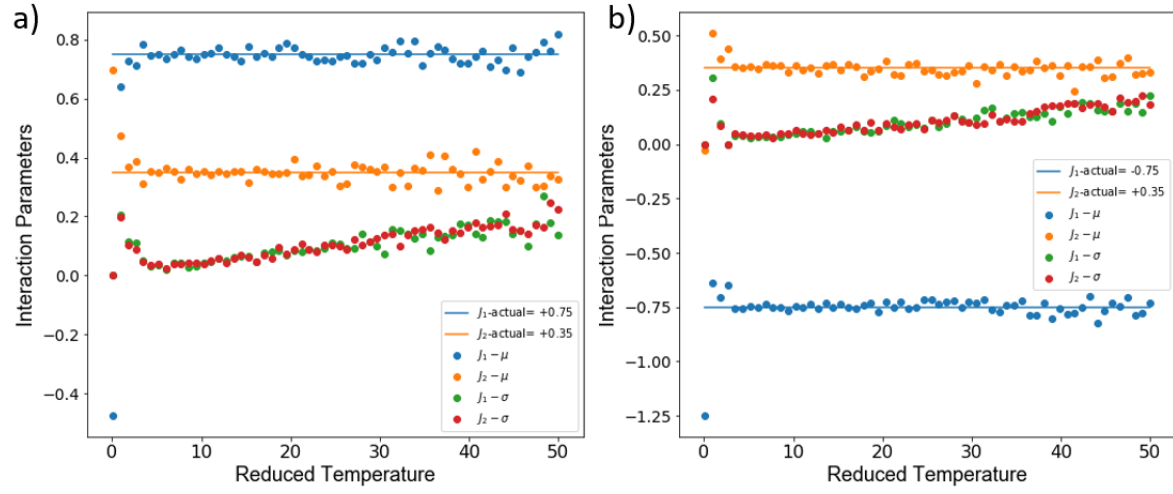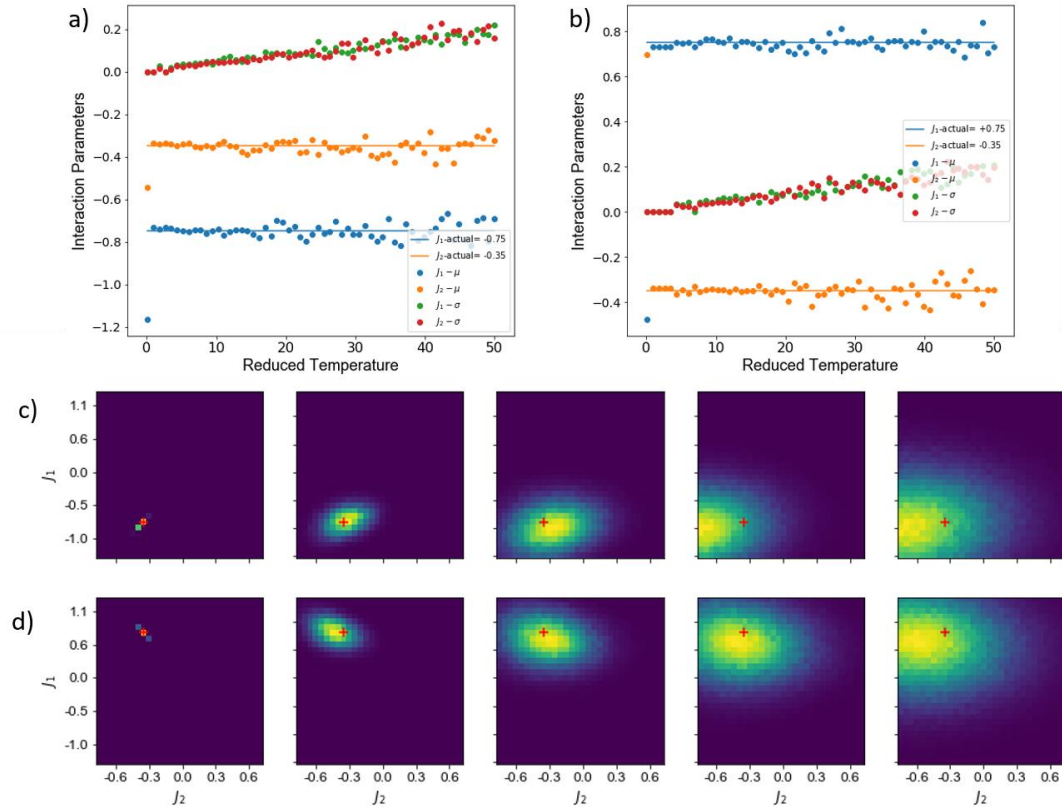
Figure 4. Reconstructed values (mean, μ) of the exchange integrals and their uncertainties (standard deviation, σ) in the (a) first ($J_1 > 0$, $J_2 > 0$) and (b) second ($J_1 < 0$, $J_2 > 0$) quadrants, where there are well defined ground states, as a function of reduced temperature.

For the next two cases ($J_1 = -0.75$, $J_2 = -0.35$ and $J_1 = +0.75$, $J_2 = -0.35$), the nearest and next nearest neighbor interactions are incompatible and lead to the frustrated ground states. The ground states of these systems are not well defined. Proposed reconstruction technique works even in this scenario (of no well-defined ground states). Since there are no apparent phase-transitions in these regions, the reconstruction works even at low temperatures (Fig. 5). High temperature behavior for these cases shows similar behavior to that described for the first two cases.
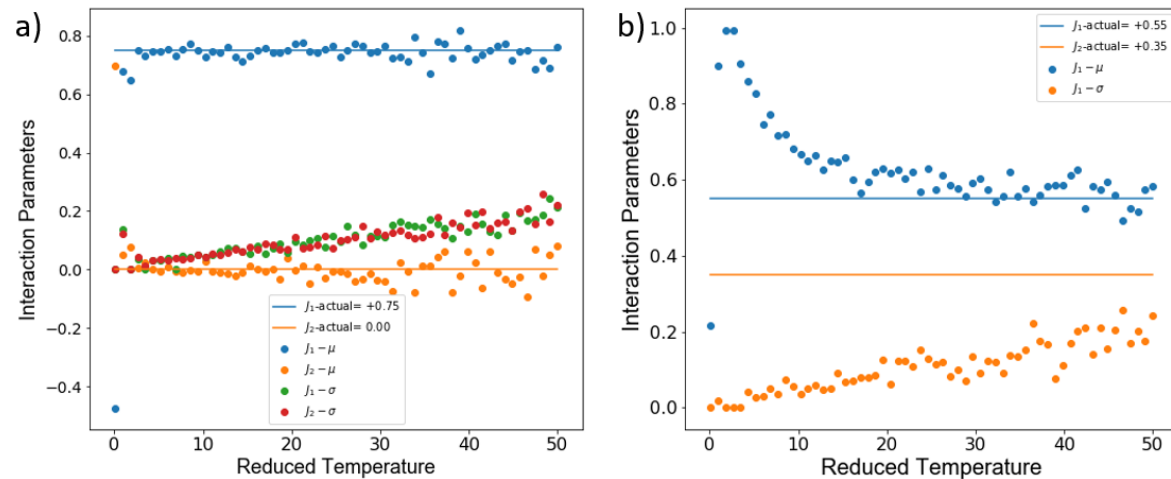
Figure 5. Reconstructed values (mean, μ) of the exchange integrals and their uncertainties (standard deviation, σ) in the (a) third ($J_1 < 0$, $J_2 < 0$) and (b) fourth ($J_1 > 0$, $J_2 < 0$) quadrants, where there are well defined ground states, as a function of reduced temperature. Likelihoods of the reconstructed values in the exchange integrals space of the (c) third and (d) fourth quadrants. The values of the exchange integrals used to produce to the pseudo experimental data is shown by the cross in each plot (c and d).

To complement the analysis of the uncertainties via the corresponding point estimates of mean and dispersion and the marginalized posterior distribution for individual exchange integral values, we also analyze corresponding joint distributions to perform uncertainty quantification (UQ) in the 2D parameter space. In the limit of large sample, where each sample corresponds to an individual local configuration, the likelihood function attains the shape of the normal distribution centered around the limiting distribution with variance equal to 1/4 as a result of the variance equalizing transformation built into the probability space with metric s. The log-likelihood of a sample at a distance s from the limiting probability distribution is then proportional

to $-2ns^2$, where n is the number of samples (individual local configurations)[51, 52]. As the distance s increases the log-likelihood decreases and it does so with a linear dependence on the number of samples n. The likelihood over the exchange integral space is shown in Fig. 5c-d at different points of reduced temperature. The increase in uncertainty with temperature can be visualized as the increase in spread of likelihood.

3.3. Reconstruction for overdetermined and underdetermined model

In this section, we explore the cases where the pseudo experimental (base) case is guided only by the interactions between the nearest neighbors ($J_1 = 0.75$) and is reconstructed by simulations incorporating both nearest and next nearest neighbors (J_1 and J_2). This case is equivalent to the traditional 2D Ising model with NN interactions only. The reconstructions follow a similar trend as discussed in the previous cases (Fig. 5a). It works with very high certainty post the transition temperature and the uncertainty increases with temperature. Interestingly, for this case the zero value of the J_2 can be reliably determined.
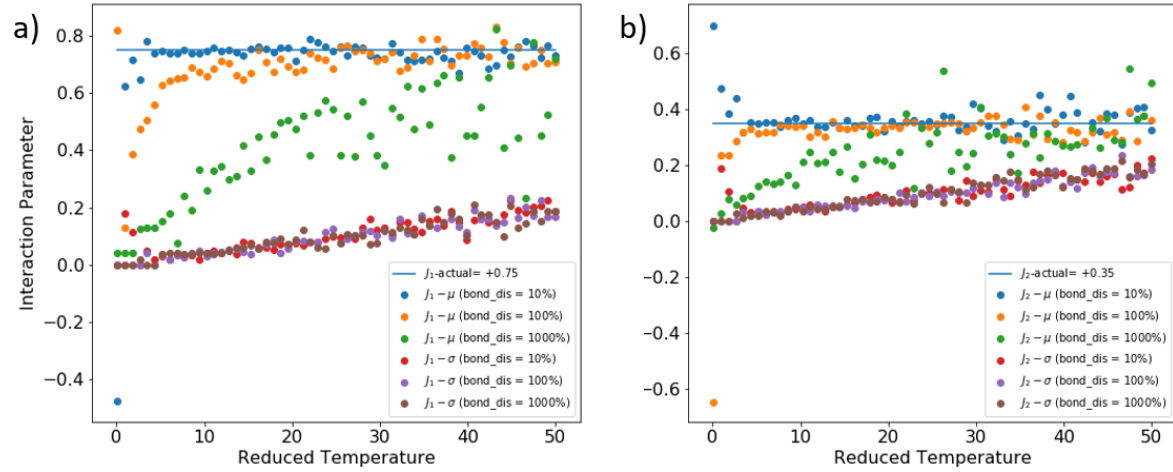

Figure 6. Reconstructed values (mean, μ) of the exchange integrals and their uncertainties (standard deviation, σ) in the (a) overdetermined and (b) underdetermined cases as a function of reduced temperature.

We further explore the reconstruction of an underdetermined model (Figure 6b) where the experimental image is guided by both nearest and next nearest neighbors ($J_1 = +0.75$, $J_2 = +0.35$)

while it is modeled using the simulations with only nearest neighbor interactions (J_1). The model attempts to fit the data by overestimating the exchange integral, but the estimated value is different at different temperatures. It can be inferred from the reconstruction curves that a model with only the nearest neighbor interactions will not be able to reproduce the experimental results that are governed by both nearest and next-nearest neighbors. The likelihood for both these cases in the exchange integral space is shown in Fig. 7a-b respectively.
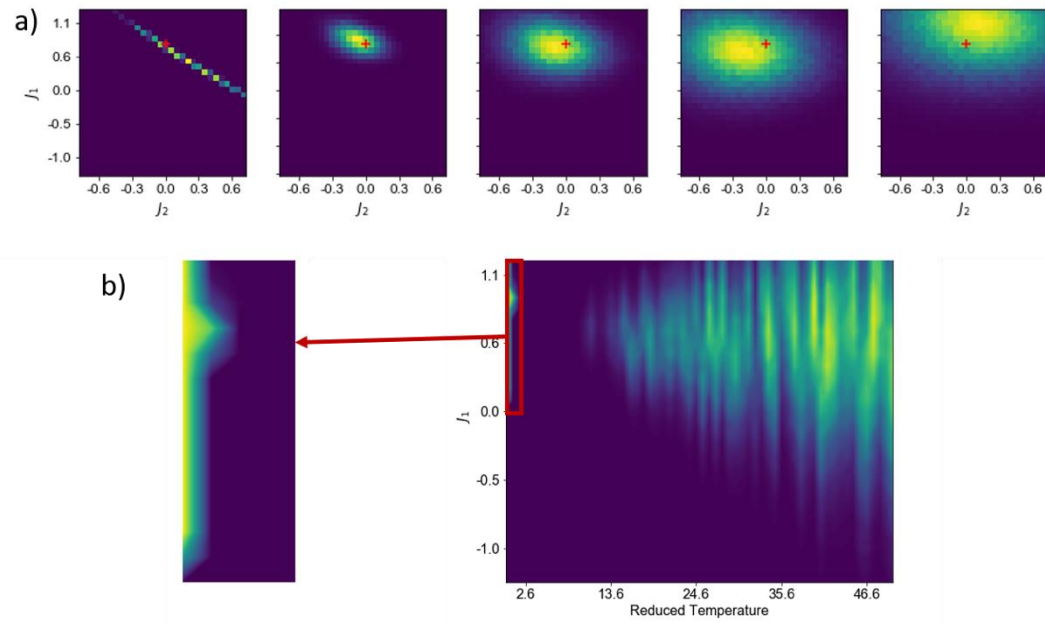
Figure 7. Likelihoods of the reconstructed values in the exchange integrals space of (a) overdetermined and (b) underdetermined cases. The values of the exchange integrals used to produce to the pseudo experimental data in the overdetermined case is shown by cross in (a) and the value of exchange integrals used in underdetermined case are $J_1 = +0.55$ and $J_2 = +0.35$.

3.4. Bond-disorder effects on the uncertainty quantification

Reconstruction until this section has been performed on a square lattice in which the values of exchange integrals in the pseudo experimental data set were constant over the entire lattice. In this section, we explore the reconstruction when a bond-disorder is introduced. To execute this, every lattice site is associated with a different value of exchange integral $J+\Delta J$, where ΔJ is normally distributed with zero mean and a specified standard deviation. The disorder is added to

both exchange integrals J_1 and J_2. The standard deviation is referred to as disorder for the rest of the section. The bond disorder was only introduced in the pseudo-experiment case while the simulation cases were run at $\Delta J = 0$.

Figure 8. Reconstructed values (mean, μ) of the exchange integrals and their uncertainties (standard deviation, σ) corresponding to the (a) nearest neighbor interactions, and (b) next nearest neighbor interactions as a function of reduced temperature and noise.

Reconstructions are performed for disorder levels of 10%, 100% and 1000% of the values of exchange integrals and the results are shown in Fig. 8. The reconstruction works well at 10%. With 100% disorder, the reconstruction starts to fail at the low temperatures since at these temperatures the values of exchange integrals have a very strong effect on configuration. The configurations get more random with increase in temperature and this is represented by increase in uncertainty in reconstructions. At 1000% disorder the reconstruction completely fails, and the proposed methods cannot be applied at these high levels of disorder.
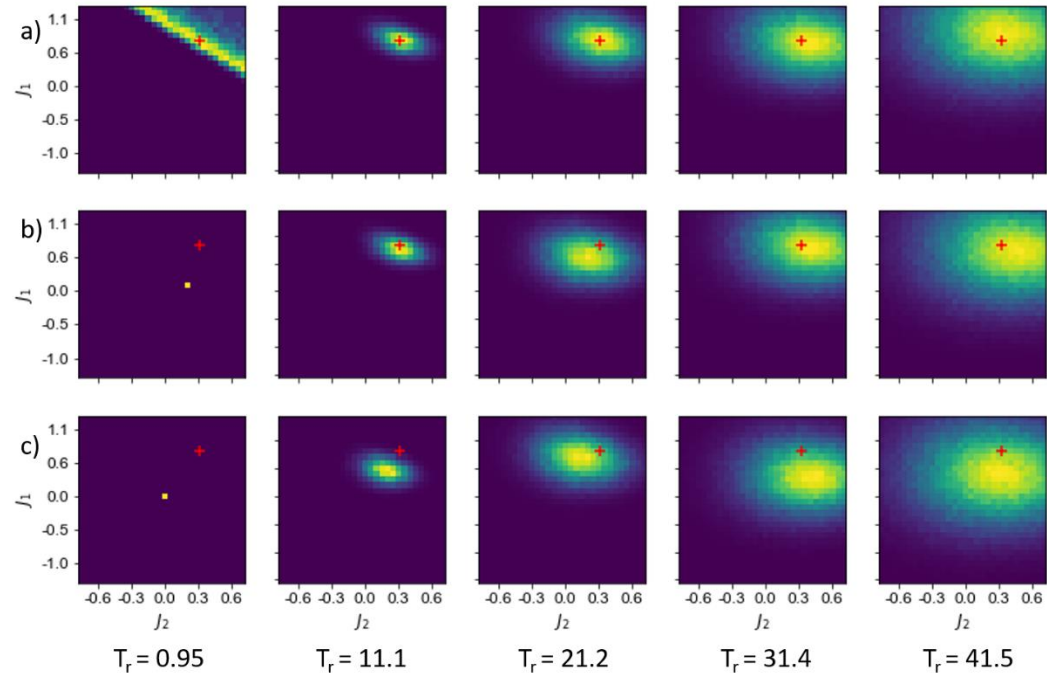
Figure 9. Likelihoods of the reconstructed values in the exchange integrals space at disorder levels of (a) 10%, (b) 100% and (c) 1000% at different points of reduced temperature. The values of the exchange integrals used to produce to the pseudo experimental data is marked by the cross in each plot.

The likelihood over the exchange integral space is shown in Fig. 9a-c at different points of reduced temperature. Note that for temperatures close to transition temperature and low disorder, the recovered exchange integrals are strongly correlated, suggesting that the system behavior can be well defined by single values of exchange integrals. For higher disorder level reconstruction is impossible.

Above the transition, the reconstruction yields unbiased estimates of exchange integrals for low and intermediate disorder, whereas for high disorder the reconstructed integrals are below the true value. This trend continues for higher temperatures. Finally, in all cases the temperature increase yields broader distribution of posterior values, rendering reconstruction less reliable. Further, the effects of lattice size on reconstruction and uncertainty quantification are discussed in supplementary material.

4. Conclusions

To summarize, here we have developed an approach for the reconstruction of the microscopic parameters of lattice model with two competing interactions from the observations of the microscopic degrees of freedom. We established the uncertainties and reliability of such analysis in a broad parameter-temperature space. Below the phase transition the reconstruction is impossible and can lead to the biased estimates of the parameters. However, these situations can be readily identified from experiment as the presence of large single-phase regions and correspondingly extremely poor sampling of the configuration space. At the same time, the reconstruction is possible well above the phase transition (1-2 order of magnitude) and in the regions of the parameter space when the macroscopic ground state of the system is poorly defined due to frustrated interactions. Similarly, the reconstruction is robust with respect to the frozen disorder. Interestingly, bond disorder tends to affect reconstruction close to transition temperatures leading to biased estimates of exchange integrals, whereas at high temperatures unbiased estimates with large uncertainty emerge.

As an ancillary task, we have developed a machine learning approach based on histogram clustering to predict phase diagrams efficiently using a reduced descriptor space. We illustrated that the phase evolution in the vicinity of regions comprising frustrated phase gives rise to the hierarchy of local configurations.

The explored approach can be applied to traditionally complex problems in condensed matter physics such as ferroelectric relaxors and morphotropic phase boundary systems, spin and cluster glasses, quantum systems once the local descriptors linked to the relevant physical behaviors are known. Correspondingly, of interest becomes the role of precision in determination of local descriptors on parameter reconstruction, necessitating the development of appropriate Bayesian frameworks to link the experimental data and reconstructions.

Supplementary Material

See supplementary material for discussion of effect of lattice size on reconstruction and uncertainty quantification.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.
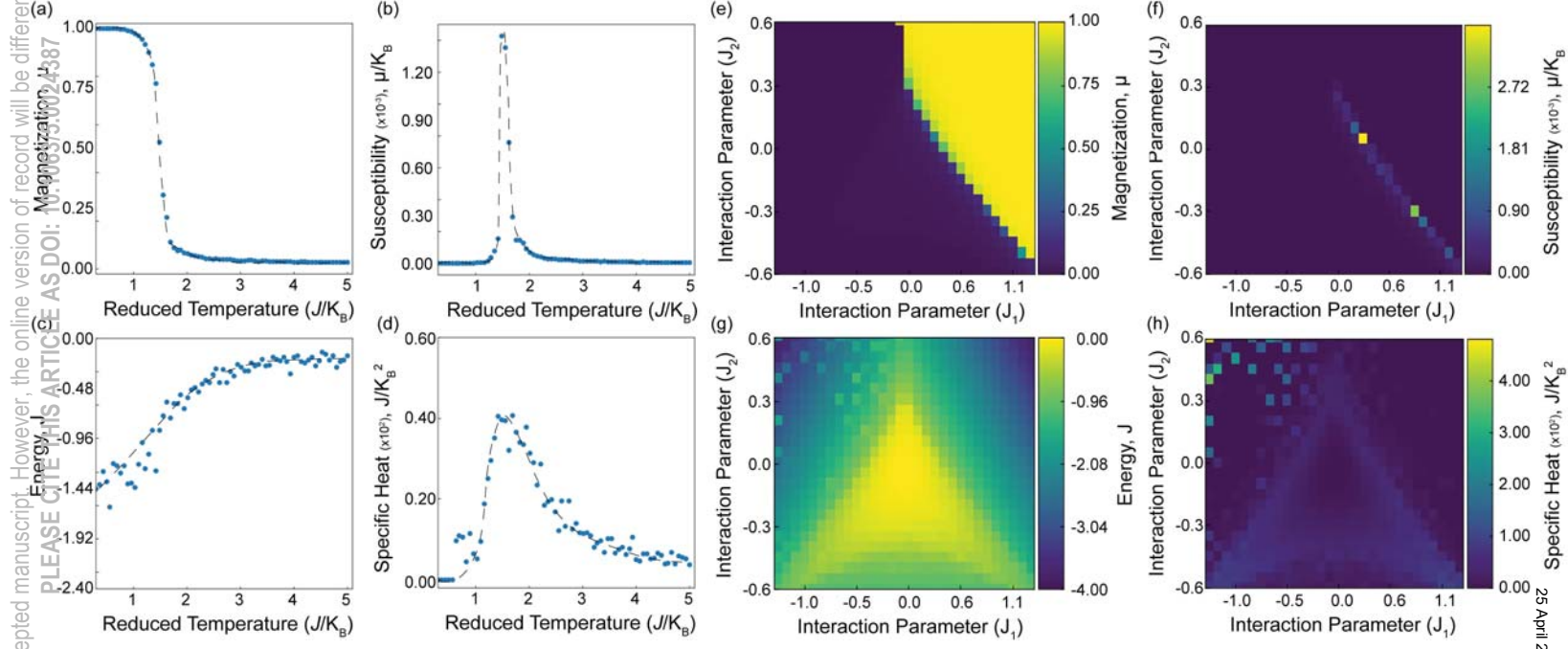
References

1.   D. A. Chenet, O. B. Aslan, P. Y. Huang, C. Fan, A. M. van der Zande, T. F. Heinz and J. C. Hone, Nano Lett. **15** (9), 5667-5672 (2015).

2.   K.-A. N. Duerloo, Y. Li and E. J. Reed, Nature Communications **5**, 4214 (2014).

3.   Y. Kohama, H. Ishikawa, A. Matsuo, K. Kindo, N. Shannon and Z. Hiroi, Proceedings of the National Academy of Sciences **116** (22), 10686 (2019).

4.   L. D. Landau and E. M. Lifshitz, *Statistical Physics*. (Elsevier Science, 2013).

5.   A. Kitaev, Annals of Physics **321** (1), 2-111 (2006).

6.   S. Cocco, R. Monasson, L. Posani, S. Rosay and J. Tubiana, Physica a-Statistical Mechanics and Its Applications **504**, 45-76 (2018).

7.   S. N. Datta and S. Hansda, Chem. Phys. Lett. **621**, 102-108 (2015).

8.   D. Chandler and D. Wu, *Introduction to Modern Statistical Mechanics*. (Oxford University Press, 1987).

9.   J. Carrasquilla and R. G. Melko, Nature Physics **13** (5), 431-434 (2017).

10.   K. Ch'ng, J. Carrasquilla, R. G. Melko and E. Khatami, Physical Review X **7** (3), 031038 (2017).

11.   Evert P. L. van Nieuwenburg, Y.-H. Liu and Sebastian D. Huber, Nature Physics **13** (5), 435-439 (2017).

12.   A. L. Ferguson, Journal of Physics: Condensed Matter **30** (4), 043002 (2018).

13.   B. S. Rem, N. Käming, M. Tarnowski, L. Asteria, N. Fläschner, C. Becker, K. Sengstock and C. Weitenberg, Nature Physics **15** (9), 917-920 (2019).

14.   A. Bohrdt, C. S. Chiu, G. Ji, M. Xu, D. Greif, M. Greiner, E. Demler, F. Grusdt and M. Knap, Nature Physics **15** (9), 921-924 (2019).

15.   J. F. Rodriguez-Nieva and M. S. Scheurer, Nature Physics **15** (8), 790-795 (2019).

16.   Y. Ming, C.-T. Lin, S. D. Bartlett and W.-W. Zhang, npj Computational Materials **5** (1), 88 (2019).

17.   E. Aurell and M. J. P. r. l. Ekeberg, **108** (9), 090201 (2012).

18.   C. Donner and M. J. P. R. E. Opper, **96** (6), 062104 (2017).

19.   H.-L. Zeng, M. Alava, E. Aurell, J. Hertz and Y. J. P. r. l. Roudi, **110** (21), 210601 (2013).

20.   A. Y. Lokhov, M. Vuffray, S. Misra and M. J. S. a. Chertkov, **4** (3), e1700791 (2018).

21.   H. C. Nguyen, R. Zecchina and J. J. A. i. P. Berg, **66** (3), 197-261 (2017).

22.   S. J. Pennycook and P. D. Nellist, (Springer, New York, 2011).

23.   O. L. Krivanek, M. F. Chisholm, V. Nicolosi, T. J. Pennycook, G. J. Corbin, N. Dellby, M. F. Murfitt, C. S. Own, Z. S. Szilagyi, M. P. Oxley, S. T. Pantelides and S. J. Pennycook, Nature **464** (7288), 571-574 (2010).

24.   P. Y. Huang, S. Kurasch, J. S. Alden, A. Shekhawat, A. A. Alemi, P. L. McEuen, J. P. Sethna, U. Kaiser and D. A. Muller, Science **342** (6155), 224-227 (2013).

25.   C. L. Jia, V. Nagarajan, J. Q. He, L. Houben, T. Zhao, R. Ramesh, K. Urban and R. Waser, Nat. Mater. **6** (1), 64-69 (2007).

26.   C. L. Jia, S. B. Mi, K. Urban, I. Vrejoiu, M. Alexe and D. Hesse, Physical Review Letters **102** (11) (2009).

27.   M. F. Chisholm, W. D. Luo, M. P. Oxley, S. T. Pantelides and H. N. Lee, Physical Review Letters **105** (19) (2010).

28.   A. K. Yadav, C. T. Nelson, S. L. Hsu, Z. Hong, J. D. Clarkson, C. M. Schlepuetz, A. R. Damodaran, P. Shafer, E. Arenholz, L. R. Dedon, D. Chen, A. Vishwanath, A. M. Minor, L. Q. Chen, J. F. Scott, L. W. Martin and R. Ramesh, Nature **530** (7589), 198-+ (2016).
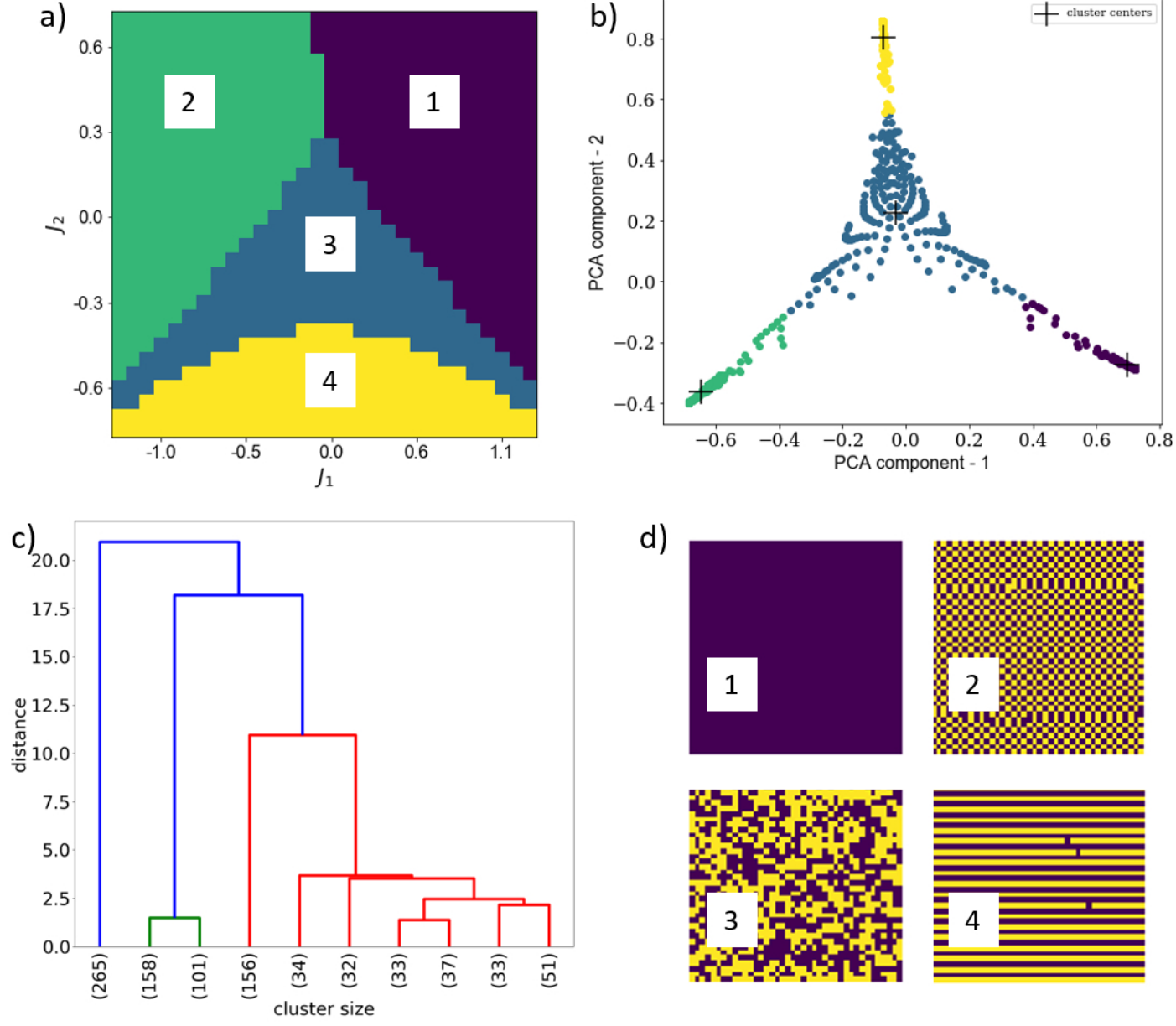
29.   S. Das, Y. L. Tang, Z. Hong, M. A. P. Goncalves, M. R. McCarter, C. Klewe, K. X. Nguyen, F. Gomez-Ortiz, P. Shafer, E. Arenholz, V. A. Stoica, S. L. Hsu, B. Wang, C. Ophus, J. F. Liu, C. T. Nelson, S. Saremi, B. Prasad, A. B. Mei, D. G. Schlom, J. Iniguez, P. Garcia-Fernandez, D. A. Muller, L. Q. Chen, J. Junquera, L. W. Martin and R. Ramesh, Nature **568** (7752), 368-+ (2019).

30.   C. L. Jia, S. B. Mi, M. Faley, U. Poppe, J. Schubert and K. Urban, Phys. Rev. B **79** (8) (2009).

31.   A. Borisevich, O. S. Ovchinnikov, H. J. Chang, M. P. Oxley, P. Yu, J. Seidel, E. A. Eliseev, A. N. Morozovska, R. Ramesh, S. J. Pennycook and S. V. Kalinin, ACS Nano **4** (10), 6071-6079 (2010).

32.   A. Y. Borisevich, H. J. Chang, M. Huijben, M. P. Oxley, S. Okamoto, M. K. Niranjan, J. D. Burton, E. Y. Tsymbal, Y. H. Chu, P. Yu, R. Ramesh, S. V. Kalinin and S. J. Pennycook, Physical Review Letters **105** (8) (2010).

33.   Q. He, R. Ishikawa, A. R. Lupini, L. Qiao, E. J. Moon, O. Ovchinnikov, S. J. May, M. D. Biegalski and A. Y. Borisevich, ACS Nano **9** (8), 8412-8419 (2015).

34.   Y. M. Kim, J. He, M. D. Biegalski, H. Ambaye, V. Lauter, H. M. Christen, S. T. Pantelides, S. J. Pennycook, S. V. Kalinin and A. Y. Borisevich, Nat. Mater. **11** (10), 888-894 (2012).

35.   E. A. Eliseev, S. V. Kalinin, Y. J. Gu, M. D. Glinchuk, V. Khist, A. Borisevich, V. Gopalan, L. Q. Chen and A. N. Morozovska, Phys. Rev. B **88** (22) (2013).

36.   A. Y. Borisevich, A. N. Morozovska, Y. M. Kim, D. Leonard, M. P. Oxley, M. D. Biegalski, E. A. Eliseev and S. V. Kalinin, Physical Review Letters **109** (6) (2012).

37.   Q. Li, C. T. Nelson, S. L. Hsu, A. R. Damodaran, L. L. Li, A. K. Yadav, M. McCarter, L. W. Martin, R. Ramesh and S. V. Kalinin, Nat. Commun. **8** (2017).

38.   O. S. Ovchinnikov, S. Jesse, P. Bintacchit, S. Trolier-McKinstry and S. V. Kalinin, Physical Review Letters **103** (15) (2009).

39.   L. Vlcek, S. Yang, M. Ziatdinov, S. Kalinin and R. Vasudevan, Microscopy and Microanalysis **25** (S2), 130-131 (2019).

40.   L. Vlcek, A. Maksov, M. Pan, R. K. Vasudevan and S. V. Kalinin, ACS Nano **11** (10), 10313-10320 (2017).

41.   L. Vlcek, R. K. Vasudevan, S. Jesse and S. V. Kalinin, Journal of Chemical Theory and Computation **13** (11), 5179-5194 (2017).

42.   L. Vlcek, W. W. Sun and P. R. C. Kent, J. Chem. Phys. **147** (16) (2017).

43.   L. Vlcek, A. Maksov, M. H. Pan, R. K. Vasudevan and S. V. Kahnin, ACS Nano **11** (10), 10313-10320 (2017).

44.   L. Vlcek, M. Ziatdinov, A. Maksov, A. Tselev, A. P. Baddorf, S. V. Kalinin and R. K. Vasudevan, ACS Nano **13** (1), 718-727 (2019).

45.   L. Vlcek, R. K. Vasudevan, S. Jesse and S. V. Kalinin, Journal of Chemical Theory and Computation (2017).

46.   S. J. Wetzel, Physical Review E **96** (2), 022140 (2017).

47.   J. M. Sanchez, F. Ducastelle and D. Gratias, Physica A: Statistical Mechanics and its Applications **128** (1), 334-350 (1984).

48.   W. Hu, R. R. P. Singh and R. T. Scalettar, Physical Review E **95** (6), 062122 (2017).

49.   A. Canabarro, F. F. Fanchini, A. L. Malvezzi, R. Pereira and R. Chaves, Physical Review B **100** (4), 045129 (2019).

50.   G. James, D. Witten, T. Hastie and R. Tibshirani, *An Introduction to Statistical Learning: with Applications in R*. (Springer New York, 2013).

51.   W. K. Wootters, Physical Review D **23** (2), 357-362 (1981).

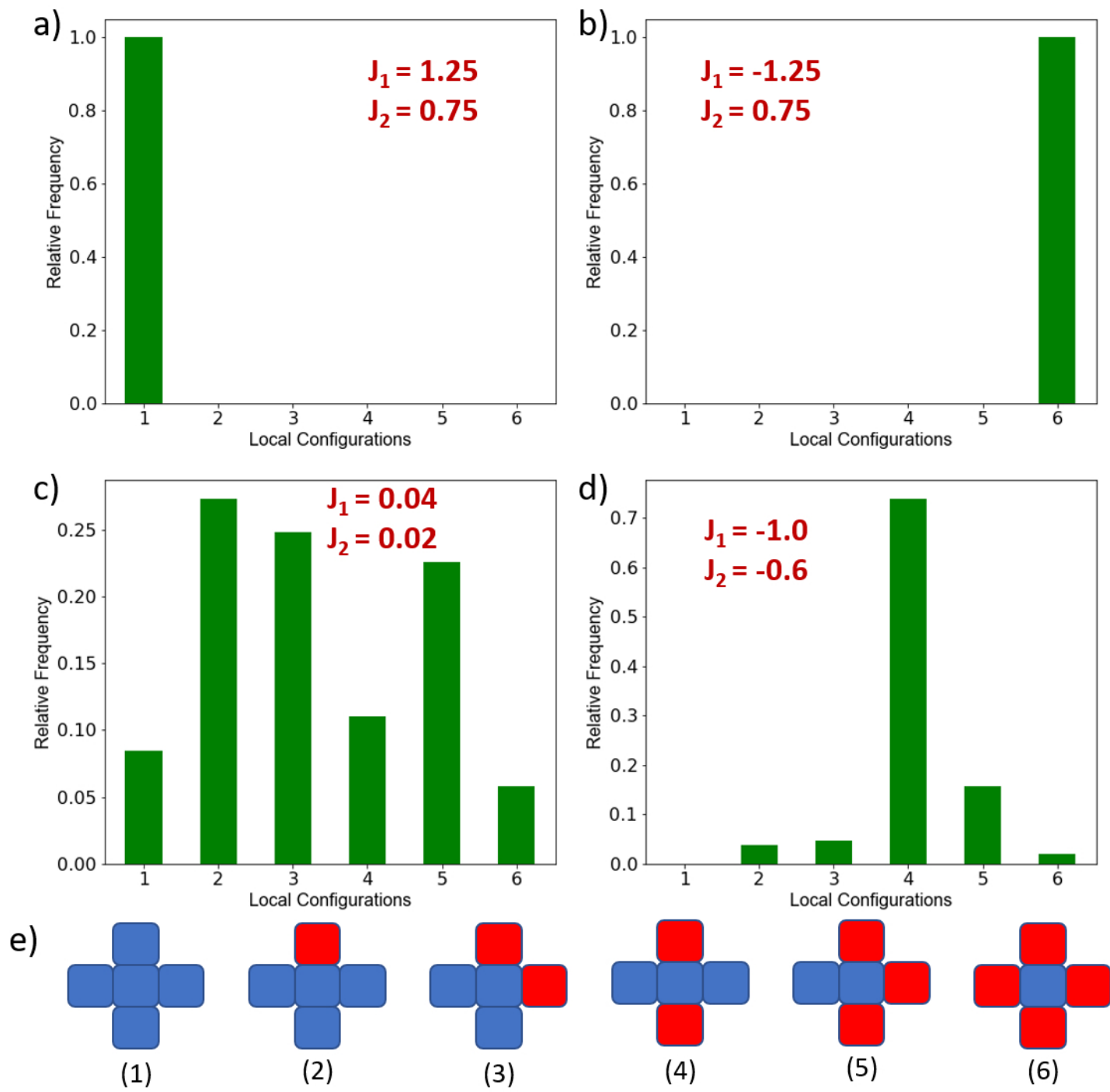52.   A. Bhattacharyya, Sankhyā: The Indian Journal of Statistics (1933-1960) **7** (4), 401-406 (1946).

25 April 2024 09:26:01

25 April 2024 09:26:01

25 April 2024 09:26:01