

Predicting parameters of a model cuprate superconductor using machine learning

V. A. Ulitko, D. N. Yasinskaya, S. A. Bezzubin, A. A. Koshelev,
Y. D. Panov

Ural Federal University, 19 Mira street, Ekaterinburg, 620002, Russia

Abstract

The computational complexity of calculating phase diagrams for multi-parameter models significantly limits the ability to select parameters that correspond to experimental data. This work presents a machine learning method for solving the inverse problem – forecasting the parameters of a model Hamiltonian for a cuprate superconductor based on its phase diagram. A comparative study of three deep learning architectures was conducted: VGG, ResNet, and U-Net. The latter was adapted for regression tasks and demonstrated the best performance. Training the U-Net model was performed on an extensive dataset of phase diagrams calculated within the mean-field approximation, followed by validation on data obtained using a semi-classical heat bath algorithm for Monte Carlo simulations. It is shown that the model accurately predicts all considered Hamiltonian parameters, and areas of low prediction accuracy correspond to regions of parametric insensitivity in the phase diagrams. This allows for the extraction of physically interpretable patterns and validation of the significance of parameters for the system. The results confirm the promising potential of applying machine learning to analyze complex physical models in condensed matter physics.

Keywords:

machine learning, U-Net, computer modeling, parameter regression, phase diagrams, cuprate high-temperature superconductors

Email address: vasiliy.ulitko@urfu.ru (V. A. Ulitko)

Introduction

Modern microscopic models of complex materials with non-trivial properties are typically multi-parametric, and the direct calculation of their phase diagrams is computationally expensive. The problem of computational complexity becomes especially critical when solving the inverse problem, which involves finding the model parameters that best reproduce an experimentally observed effect or a phase diagram.

An example of such a multi-parametric system is the previously proposed [1, 2, 3] pseudospin model of high-temperature superconducting (HTSC) cuprate. For this model, we derived equations for the critical temperatures of various ordered phases in mean-field approximation (MFA) [4] and constructed temperature phase diagrams [5, 6] – dependencies of the phase state on temperature and concentration. The possibility of simulating this system using the classical Monte Carlo method was also discussed in [7, 8].

Due to the large number of independent model parameters, solving the inverse problem of finding a parameter set that reproduces the experimental phase diagrams for such a system is extremely challenging. In this regard, it is feasible to explore the possibility of applying machine learning to reduce the required number of phase diagram calculations.

Previously, machine learning methods have been used to predict phase boundaries on experimentally and theoretically determined phase diagrams [9, 10], to identify stable phases and predict phase diagrams [11, 12, 13, 14], and to study phase transitions [15, 16, 17]. Specialized neural networks have been developed for analyzing images of magnetic domains to quantitatively estimate Hamiltonian parameters [18], while convolutional networks have successfully constructed detailed phase diagrams for systems with skyrmions and other topological phases [19, 20]. In the context of HTSC research, machine learning has been successfully applied to predict critical temperatures of various systems [21, 22, 23, 24]. In this work, we developed and tested a machine learning method based on a U-Net convolutional neural network to solve the inverse problem: predicting the parameters of a model HTSC cuprate from its theoretical phase diagram. The primary goal of this work is to overcome the computational complexity of calculating phase diagrams and to develop an efficient method for predicting model parameters. This research focuses not only on demonstrating the principal feasibility of applying machine learning but also on analyzing the physical interpretability of the obtained results.

The article is organized as follows. The first section introduces the physical model of the HTSC cuprate and the modeling methodology. The second section compares different learning architectures and discusses their features. The third section presents the results of numerical experiments conducted using the U-Net architecture to predict model parameters.

Physical model and modeling methods

The rich phase diagram of HTSC cuprates makes the direct statistical modeling of these systems within an effective pseudospin framework highly relevant [3].

The fundamental structural elements of HTSC cuprates are the CuO_2 planes, composed of square CuO_4 clusters connected via shared oxygen atoms. Within the effective pseudospin model for cuprates, the basis of the local Hilbert space for a CuO_4 cluster is chosen as a quartet of states $|1Msm\rangle$. In the effective pseudospin model for cuprates, the basis of the local Hilbert space formed by the quartet of states $|1Msm\rangle$. The projections $M = -1, 0, +1$ of pseudospin $S = 1$ correspond to three multi-electronic states of the cluster $[\text{CuO}_4]^{7-,6-,5-}$. States with $M = \pm 1$ are spin singlets ($s = 0$), while the state with $M = 0$ is a spin doublet with $s = 1/2$.

The Hamiltonian of the model cuprate can be written as follows:

$$\hat{H} = \hat{H}_{ch} + \hat{H}_{exc} + \hat{H}_{tr}^{(1)} + \hat{H}_{tr}^{(2)} - \mu \sum_i \hat{S}_{z,i}. \quad (1)$$

The first term

$$\hat{H}_{ch} = \Delta \sum_i \hat{S}_{z,i}^2 + V \sum_{\langle ij \rangle} \hat{S}_{z,i} \hat{S}_{z,j} \quad (2)$$

describes local (with $\Delta = U/2$, where U is the correlation parameter) and non-local ($V > 0$) density-density correlations. Here, the operator $\hat{S}_z$ represents the z-component of the pseudospin $S = 1$. The summation runs over N sites of a square lattice, and $\langle ij \rangle$ denotes nearest neighbors. The operator

$$\hat{H}_{exc} = Js^2 \sum_{\langle ij \rangle} \hat{\sigma}_i \hat{\sigma}_j \quad (3)$$

describes the antiferromagnetic ($J > 0$) Heisenberg exchange interaction for $[\text{CuO}]_4^{6-}$ centers, where operators $\hat{\sigma} = \hat{P}_0 \hat{\mathbf{s}}/s$ account for the spin density on

the site, with $\hat{P}_0 = 1 - \hat{S}_z^2$, and $\hat{\mathbf{s}}$ is the spin operator with $s = 1/2$ [25]. The operator

$$\begin{aligned} \hat{H}_{tr}^{(1)} = & - \sum_{\langle ij \rangle m} \left[t_p \hat{P}_{m,i}^+ \hat{P}_{m,j} + t_n \hat{N}_{m,i}^+ \hat{N}_{m,j} \right. \\ & \left. + \frac{t_{pn}}{2} (\hat{P}_{m,i}^+ \hat{N}_{m,j} + \hat{N}_{m,i}^+ \hat{P}_{m,j}) \right] + h.c. \end{aligned} \quad (4)$$

describes three types of correlated single-particle transport [1, 2]: terms proportional to t_p (t_n) describe the transport of $[\text{CuO}]_4^{5-}$ ($[\text{CuO}]_4^{7-}$) centers across the $[\text{CuO}]_4^{6-}$ lattice; terms proportional to t_{pn} correspond to disproportionation and recombination processes: $[\text{CuO}]_4^{6-} + [\text{CuO}]_4^{6-} \longleftrightarrow [\text{CuO}]_4^{7-} + [\text{CuO}]_4^{5-}$. The spin projection $m = \pm 1/2$ refers to the $s = 1/2$ spin states; the orbital parts of the operators $\hat{P}$ and $\hat{N}$ are expressed via the $S = 1$ pseudospin operators: $\hat{P}^+ = (\hat{S}_+ + \hat{T}_+)/2$, $\hat{N}^+ = (\hat{S}_+ - \hat{T}_+)/2$, $\hat{T}_+ = \hat{S}_z \hat{S}_+ + \hat{S}_+ \hat{S}_z$, $\hat{S}_+ = (\hat{S}_x + i\hat{S}_y)/\sqrt{2}$. The operator

$$\hat{H}_{tr}^{(2)} = -t_b \sum_{\langle ij \rangle} \left(\hat{S}_{+i}^2 \hat{S}_{-j}^2 + \hat{S}_{+j}^2 \hat{S}_{-i}^2 \right) \quad (5)$$

describes two-particle transport (transport of a local composite boson [1, 2]): $[\text{CuO}]_4^{5-} + [\text{CuO}]_4^{7-} \leftrightarrow [\text{CuO}]_4^{7-} + [\text{CuO}]_4^{5-}$. The last sum with the chemical potential μ allows us to take into account the condition of constant charge density:

$$n = \left\langle \sum_i \hat{S}_{z,i} \right\rangle / N = \text{const.} \quad (6)$$

In the mean-field approximation (MFA), equations for the critical temperatures of ordered phases were derived in [4], considering cases where only one of the possible order parameters is non-zero. The charge-ordered (CO) phase is characterized by a non-zero order parameter $L = (S_{z,A} - S_{z,B})/2$, where $S_{z,A}$ and $S_{z,B}$ are the statistical averages of the $\hat{S}_z$ operator on sublattices A and B , which describe a checkerboard subdivision of the square lattice constructed from $[\text{CuO}]_4$ clusters. Analogous to the model of local (hard-core) bosons [26], the phase with a non-zero average $\langle \hat{S}_+^2 \rangle$ can be termed a Bose superfluid (BS) phase. For the antiferromagnetic (AFM) phase, the order parameter is the magnitude of the antiferromagnetic vector $\mathbf{l} = (\boldsymbol{\sigma}_A - \boldsymbol{\sigma}_B)/2$,

where σ_A and σ_B are the statistical averages of the spin operators ($\hat{\sigma}$) on sublattices A and B . In phases with non-zero order parameters $\langle \hat{P}_m^+ \rangle$ (P phase) and $\langle \hat{N}_m^+ \rangle$ (N phase), correlated single-particle transfer of hole (P) or electron (N) type is realized.

Numerous experimental data indicate the key role of the competition between different phases and the formation of complex phase states in HTSC cuprates [27]. Investigating these issues with analytical methods remains challenging; however, within the effective pseudospin model, it is possible to address this problem using numerical statistical modeling techniques.

For numerical modeling, we use the heat bath Monte Carlo algorithm [28], which we previously applied to model orthonickelates [29]. A brief description of the algorithm for the HTSC cuprate model (1) is provided below; details can be found in the preprint [30]. The system's wave function is written as a product of the wave functions on individual clusters:

$$|\Psi\rangle = \prod_{\nu} |\psi_{\nu}\rangle, \quad |\psi_{\nu}\rangle = \sum_{Mm} a_{Mm}^{\nu} |1Msm\rangle. \quad (7)$$

This allows for the construction of the Hamiltonian for a single cluster ν by averaging over the states of all other clusters $\mu \neq \nu$:

$$\hat{H}_{\nu} = \langle \tilde{\Psi}_{\nu} | \hat{H} | \tilde{\Psi}_{\nu} \rangle, \quad |\tilde{\Psi}_{\nu}\rangle = \prod_{\gamma \neq \nu} |\psi_{\gamma}\rangle. \quad (8)$$

Subsequently, the cluster's state is updated according to the heat bath algorithm. For this, an eigenvalue problem is solved:

$$\hat{H}_{\nu} |\psi_{\nu,k}\rangle = \varepsilon_k |\psi_{\nu,k}\rangle, \quad k = 1, \dots, 4, \quad (9)$$

and based on the obtained spectrum $\{\varepsilon_k\}$, a distribution function is constructed:

$$F(n) = \sum_{k=1}^n p_k, \quad p_k = \frac{e^{-\beta\varepsilon_k}}{\sum_i e^{-\beta\varepsilon_i}}. \quad (10)$$

Based on the inequality $F(n-1) < \xi \leq F(n)$, where ξ is a random variable uniformly distributed over the unit interval, a new state $|\psi_{\nu,n}\rangle$ for cluster ν is determined. Next, utilizing the wave function (7) corresponding to the current system state, we compute the necessary quantum-mechanical averages and collect statistics over MC steps in the standard manner.

The structure of $\hat{H}_\nu$ corresponds to the Hamiltonian of the ideal system in MFA; however, the heat bath algorithm explicitly accounts for the inhomogeneous nature of the molecular fields, which are determined by the current state of the system. This allows for a more accurate treatment of temperature fluctuations and results in lower critical temperatures of phase transitions compared to MFA predictions. Additionally, numerical calculations reveal the instability of homogeneous AFM and P phases and the existence of phase-separation (PS) regions. The ability to determine the phase boundaries of this region is a distinctive feature of the employed MC method.

A comparison of the results from MFA and numerical modeling using the heat bath algorithm is shown in Figure 1. A limitation of the algorithm is the high computational cost of an elementary MC step, associated with the necessity of solving an eigenvalue problem at each step.

In the simulation, the parameter J in the Hamiltonian (1) was set to 1. Consequently, the physical model under consideration possesses six independent parameters. To simplify the prediction task, we restricted the analysis to four parameters: Δ , V , t_b , and t_p , given the weak influence of the parameters t_n and t_{pn} on the phase diagram [5].

Selecting a machine learning model

Predicting model parameters from spatial distributions is a classic regression problem. To address this, we initially selected two deep learning architectures known for handling complex spatial dependencies: VGG [31] and ResNet [32].

First, we employed the VGG16BN architecture, a modification of VGG16 incorporating batch normalization (BN) layers to stabilize training and accelerate convergence. Its sequential application of 3×3 convolutions efficiently captures local spatial features within an image. Each convolutional block is followed by a pooling layer, which progressively reduces the spatial dimensions while increasing the number of filters. With 16 weight layers (13 convolutional and 3 fully connected), VGG16 requires a large number of parameters, making it computationally expensive. Furthermore, the aggregation of spatial information in fully connected layers can disrupt fine-grained details, potentially hindering accurate regression. To mitigate these limitations, we also evaluated two variants of the ResNet architecture: ResNet18 and ResNet50. A defining feature of ResNet is the use of skip connections,

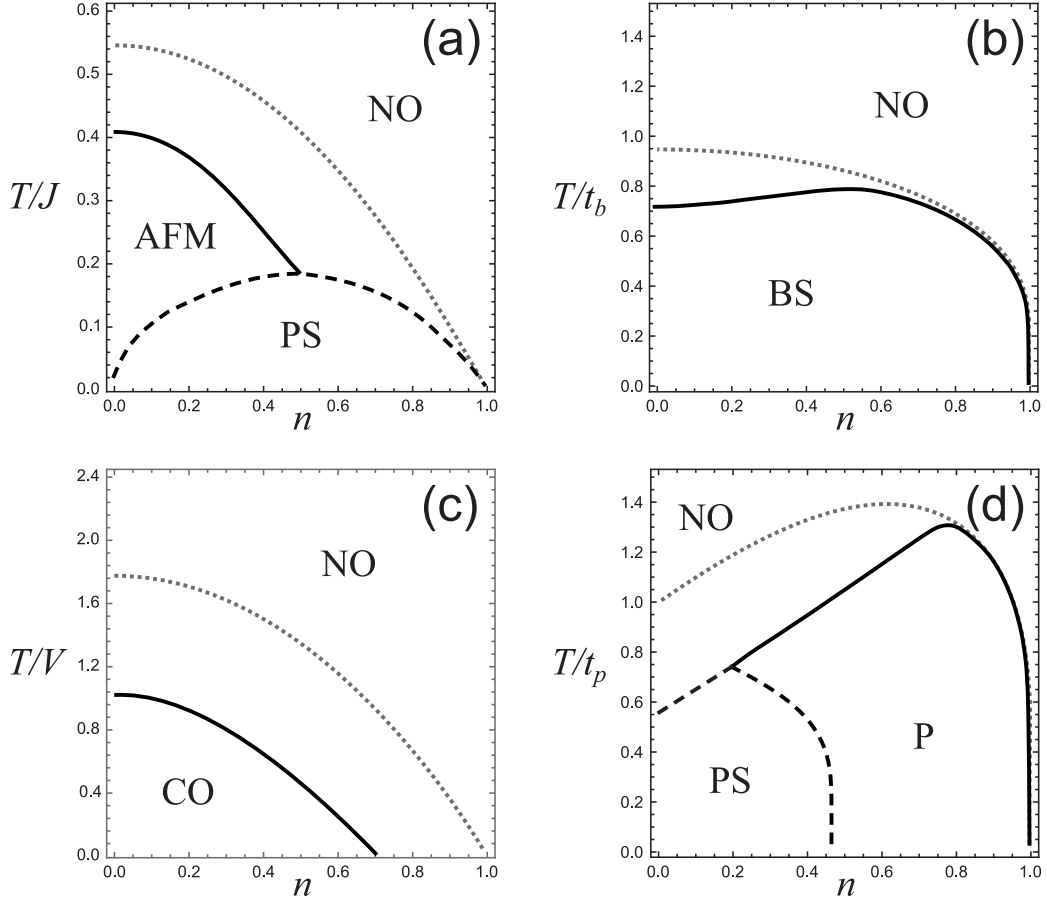


Figure 1: Comparison of the results from MFA and numerical modeling using the heat bath algorithm. The boundary between the non-ordered (NO) and corresponding ordered phases is shown by a dashed line in MFA, and by a solid line in the heat bath algorithm. The boundary of the PS region is indicated by a dotted line. The values of the non-zero model parameters are: $\Delta = 0.1$, (a) $J = 1$; (b) $t_b = 1$; (c) $V = 1$; (d) $t_p = 1$.

which alleviate the vanishing gradient problem and enable the effective training of deep networks. The fundamental component, the residual block, adds the input directly to the output of the convolutional layers. ResNet18 is a comparatively lightweight model suitable for problems with limited computational resources, while ResNet50 offers higher performance through bottleneck blocks that optimize depth and parameter efficiency.

Preliminary results indicated that the VGG and ResNet models performed poorly for our specific task. We attribute this to the fact that these models were originally designed for classification, which often does not require the preservation of detailed spatial context. Consequently, we turned to the U-Net architecture, adapting it for the regression task.

U-Net was originally developed for biomedical image segmentation tasks, where pixel-level accuracy is critical [33, 34]. Over time, its application has expanded beyond medicine, as evidenced by successful implementations in condensed matter physics [35, 36, 37, 38]. Previous work also includes examples of U-Net adapted for regression tasks [39, 40, 35, 38].

In this study, we propose a modified U-Net architecture combined with a two-stage training scheme. The operating diagram is shown in Figure 2. In the first stage, a model with an U-Net-like architecture is pre-trained to operate as an autoencoder, allowing for the identification of common features and patterns in phase diagrams. In the second stage, transfer learning is performed using frozen encoder weights, followed by further training of the regression head with L2 regularization. For this stage, the standard U-Net output layers are replaced with several fully connected layers that transform the feature maps into the four numerical values of the Hamiltonian parameters (1). The proposed modification also eliminates skip connections from the original architecture to prevent excessive model complexity and focus on the most informative features essential for the regression task.

The dataset was partitioned with 60% for training, 20% for validation, and the remaining 20% for testing. To comprehensively evaluate the predictive accuracy for the four parameter (Δ , V , t_b , t_p), we employed a set of metrics including MAE (Mean Absolute Error), RMSE (Root Mean Squared Error), and R^2 (coefficient of determination). Table 1 presents the combined results for the U-Net, VGG, and ResNet models, evaluated using the R^2 metric on phase diagrams generated with MFA. The dataset size was consistent across all models, consisting of 10,000 phase diagrams.

The results for R^2 convincingly demonstrate that the U-Net-based architecture is optimal for this task. Although VGG16BN achieves comparable

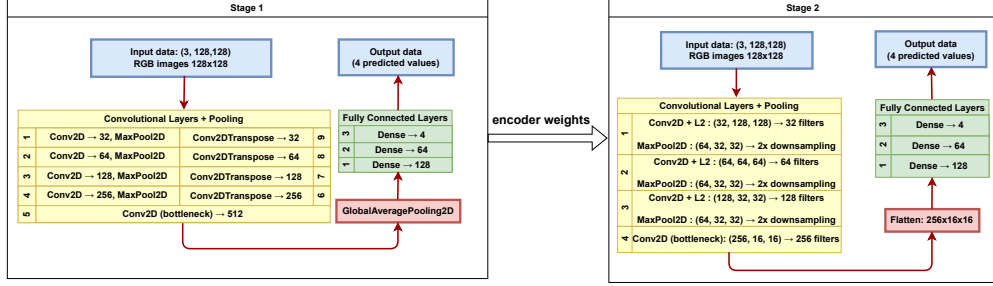


Figure 2: Two-stage neural network operation diagram.

Model \ Parameter	Δ	V	t_b	t_p	Total	Time, s
U-Net	0.4972	0.8714	0.6988	0.8818	0.7373	210
VGG16BN	0.5176	0.9132	0.4807	0.7960	0.6769	3287
ResNet18	0.0164	0.5982	0.5253	0.6937	0.4584	112
ResNet50	0.0012	-0.2436	0.3474	0.4048	0.1274	320

Table 1: Aggregate results of machine learning of different models by the R^2 metric and training time for different model parameters

prediction quality for certain individual parameters (notably V and t_p), U-Net outperforms it on the overall metric while requiring approximately 15 times less training time. Furthermore, the U-Net architecture exhibits better generalization ability with limited training data compared to the more parameter-intensive VGG model. This is a critical advantage when working with smaller datasets typical of resource-intensive phase diagram construction methods.

Results for the U-Net architecture

This section presents an analysis of the results of predicting four Hamiltonian model parameters (Δ , V , t_b , and t_p) using the U-Net convolutional neural network.

Two independent methods (MFA and the heat bath algorithm) were employed to obtain a diverse sample of phase diagrams in the temperature-density variables. Each diagram corresponds to a specific set of model parameters. Primary attention was paid to training on the diagrams obtained using MFA, as this approach enables the rapid generation of extensive datasets.

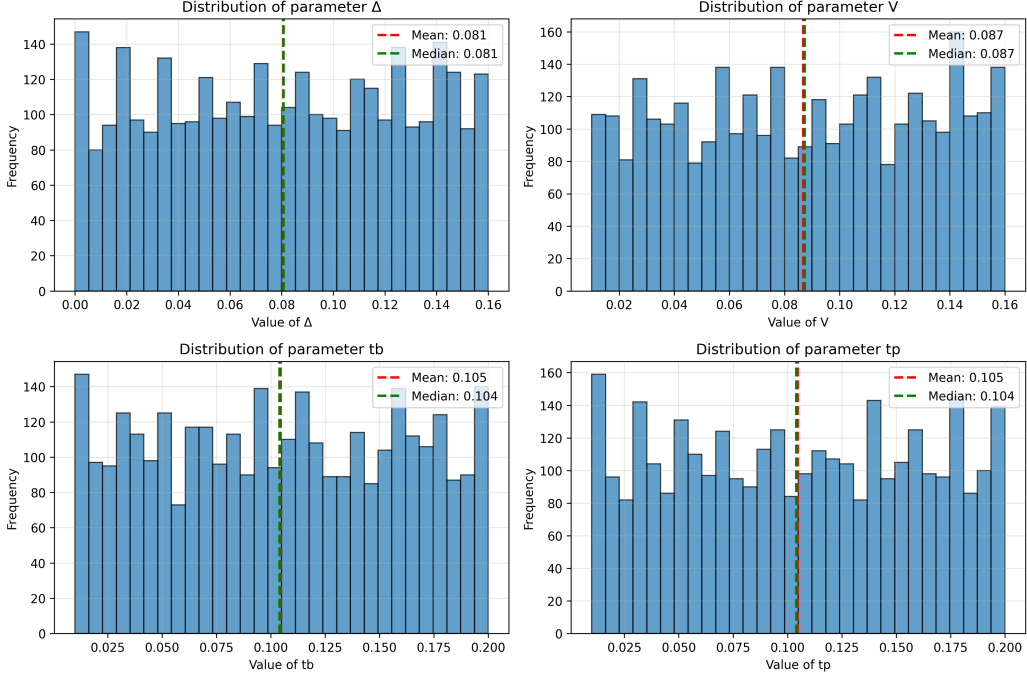


Figure 3: The distribution histograms of the parameters Δ , V , t_b and t_p for MFA demonstrate uniform generation of calculation points

Calculating a single phase diagram using MFA is at least two orders of magnitude faster than numerical calculation using the heat bath algorithm with similar computing power. The larger MFA dataset facilitates a detailed analysis of the influence of individual parameters on phase diagrams. Furthermore, the results of validating the approach on data obtained using the heat bath algorithm with a smaller training dataset are also presented.

Prediction results for MFA phase diagrams

To ensure that the calculation points in the neural network training dataset are uniformly distributed and to avoid data imbalance, we randomly generated new parameter sets based on the previously calculated points in the parameter space. As illustrated in Figure 3, the distributions of all four parameters are characterized by close mean and median, indicating the absence of systematic bias and ensuring data balance. Comparative analysis shows that random uniform parameter distributions yield a 20–30% increase in prediction accuracy compared to equidistant sampling methods.

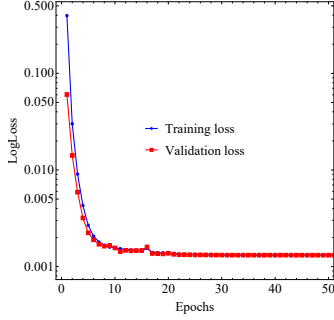


Figure 4: U-Net loss curve for MFA phase diagrams on a logarithmic scale

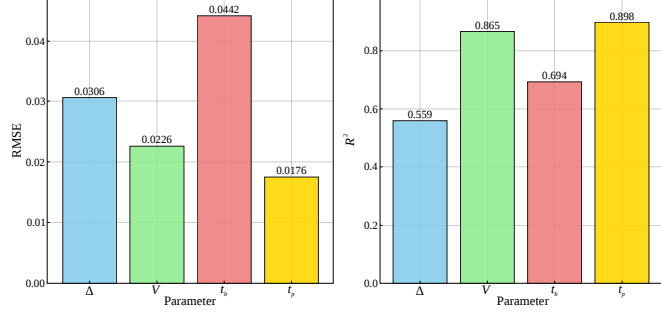


Figure 5: Bar charts of RMSE and R^2 metrics for each parameter Δ , V , t_b , t_p for MFA

Figure 4 shows the loss function curves during training and validation on a logarithmic scale. These curves converge rapidly and stabilize at very low levels, indicating a stable optimization process and the absence of overfitting.

A qualitative assessment of the prediction accuracy for each parameter is presented in Figure 5, which shows bar charts of the key metrics RMSE and R^2 for each predicted parameter. Notably, there is significant variability in prediction accuracy between different parameters, indicating various degrees of parameters influence on phase diagrams.

For visual evaluation of the correlation between the predicted and expected parameter values, scatter plots are shown in Figure 6. The red line indicates the linear dependence, representing the line of best fit. The blue line denotes the main trend according to the coefficient of determination R^2 . Areas showing no correlation between the predicted and expected parameter values are outlined in black.

For parameters Δ and t_p , the points are mostly cluster along the red ideal predictive line across the entire range of parameter variations, which confirm the high quality of the regression. Scatter plots for t_b and V reveal heterogeneity with near-zero correlation outside the general trend in regions of small parameter values. These regions are highlighted with black ovals in Figure 6.

This effect does not necessarily indicate mistakes in the machine learning model. Several factors could account for this. First, there might be an imbalance in the training dataset. Although the scatter plots display hori-

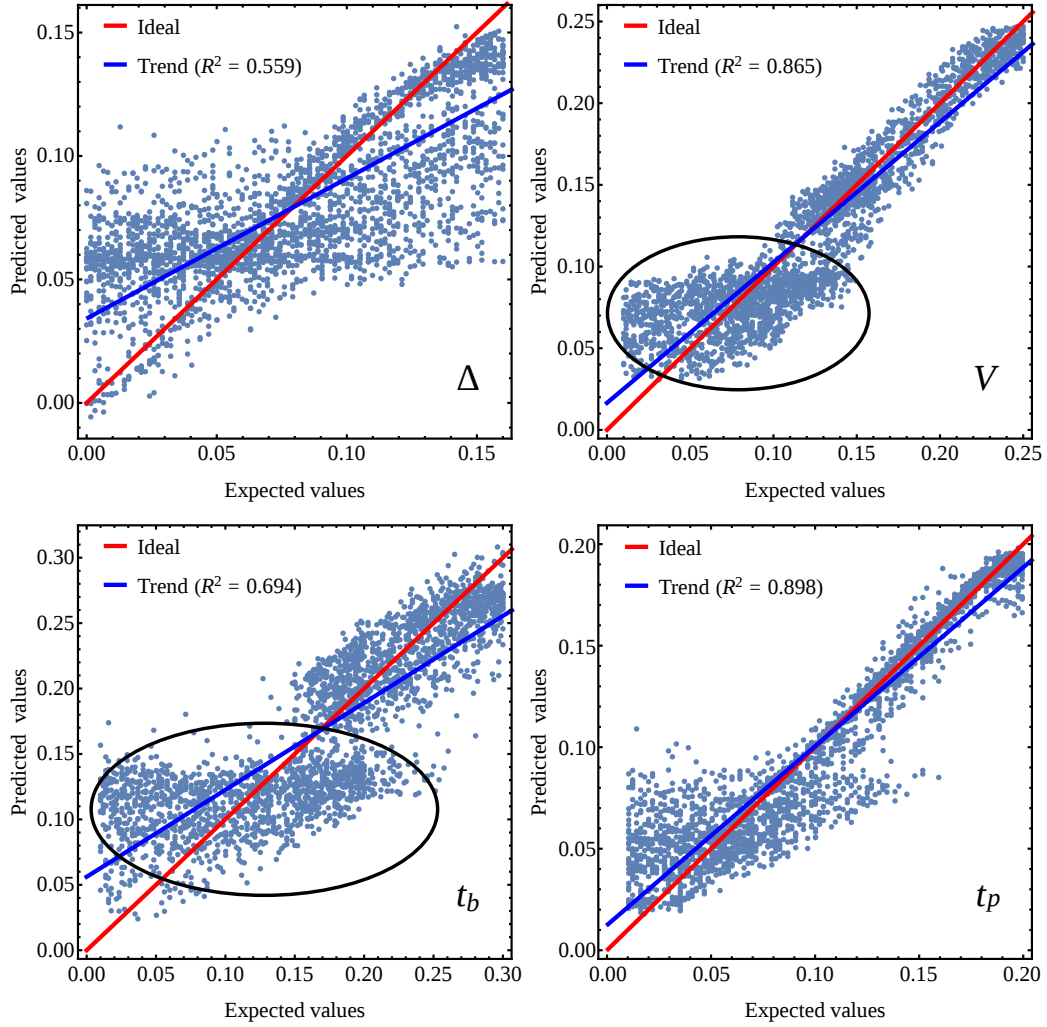


Figure 6: Scatter plots show correlation between predicted and expected values for each parameter for MFA phase diagrams. Areas where there is no correlation and the neural network can not predict the parameters are outlined in black

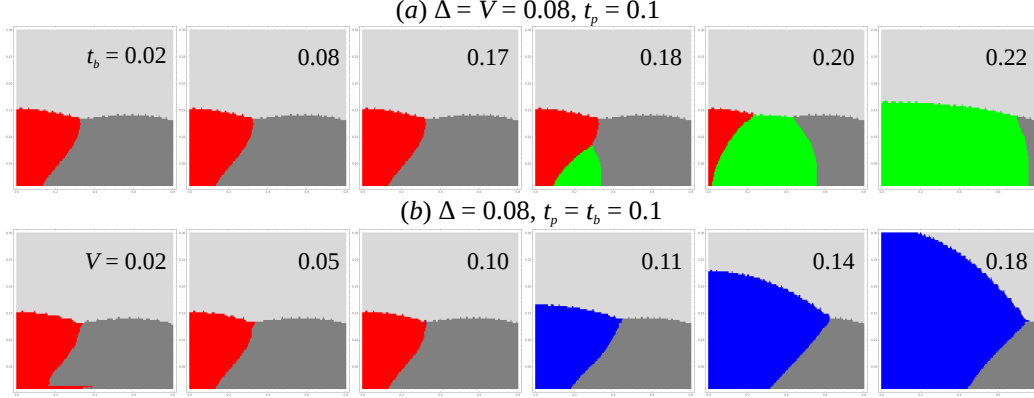


Figure 7: Sensitivity analysis of the T - n phase diagrams to parameter variations: (a) a series of diagrams plotted for fixed Δ , V , and varying $t_b \in [0.02 - 0.22]$; (b) a series of diagrams plotted for fixed Δ , t_b , and varying $V \in [0.02 - 0.18]$. Different colors represent different homogeneous phases. Light grey indicates the NO phase, red represents AFM, green represents CO, blue represents BS, and dark grey represents the Fermi liquid phase – an analogue of the P-phase with simultaneously non-zero order parameters $\langle \hat{P}_m^+ \rangle$ and $\langle \hat{N}_m^+ \rangle$.

zontal regions at certain values of t_b and V , the distribution of data points for these parameters was shown to be approximately normal (see Figure 3), which rules out this explanation.

A second potential reason is an inherent limitation of the physical model: if the parameters t_b or V are weakly correlated with the form of phase diagrams, the model may be unable to learn these relationships effectively. Furthermore, degeneracy of the phase diagrams is possible in the highlighted regions, since different parameter combinations can lead to statistically indistinguishable configurations. This can impose fundamental limits on predictive accuracy. To test this hypothesis, we fixed the mean values of three parameters and sequentially varied the fourth one under investigation.

Series of phase diagrams were calculated with sequential variation of the target parameters: to analyze the influence of parameter t_b in the interval from 0.02 to 0.22 with fixed values of parameters Δ , V , and t_p (Figure 7(a)); to analyze the influence of parameter V in the interval from 0.02 to 0.18 with fixed values of parameters Δ , t_b , and t_p (Figure 7(b)).

The results of the computational experiment confirmed that the phase diagrams remains nearly identical throughout the entire ranges of zero correlation between the predicted and expected values of parameters. This means

that variations of t_b and V within these limits does not lead to qualitative or substantial quantitative changes in the phase diagram. This observation elucidates the fundamental limitation of the prediction accuracy: the neural network can not reconstruct parameters that do not significantly affect the resulting phase diagram image. However, as demonstrated in Figure 7, beyond the critical values of parameters ($t_b \geq 0.18$, $V \geq 0.11$), qualitative changes in phase diagrams begin to manifest, which correlates with the emergence and strengthening of the statistical relationship in the scatter plots in Figure 6.

This suggests that the machine learning model acts not merely as a “black box”, but as a tool capable of analyzing the physical properties of the underlying physical model. The neural network’s inability to predict certain parameters indicates their weak influence on the target variable (the phase diagram) within the considered range. Conversely, it successfully identifies parameter intervals where the parameter becomes important and significantly affects the form of the diagram.

Low prediction accuracy in specific ranges indicates parametric degeneracy within the system, whereas high accuracy (error margins of 1–3%) is achieved in regions where the parameter significantly influences the morphological characteristics of the phase diagram.

Prediction results for the heat bath algorithm phase diagrams

Using a more precise, but computationally intensive heat bath algorithm, we calculated a dataset of 1,200 phase diagrams. Figure 8 presents the correlation scatter plots, which demonstrate excellent coefficients of determination R^2 for each parameter, as well as the presence of a significant region of low t_b values with zero correlation between the predicted and expected values.

Subsequent parametric analysis again revealed that, within the considered range of values, the phase diagrams also exhibited weak sensitivity to this parameter. This indicates that, within a more rigorous approach, the effect of the t_b parameter on the observed phase diagrams within the same range becomes less pronounced or is masked by other effects. Consequently, the primary source of prediction errors is again not a deficiency of the machine learning model but rather the physical insignificance of the parameter within the specified range. This underscores the importance of accurately defining the boundaries of parameter space during planning stages of computational experiments.

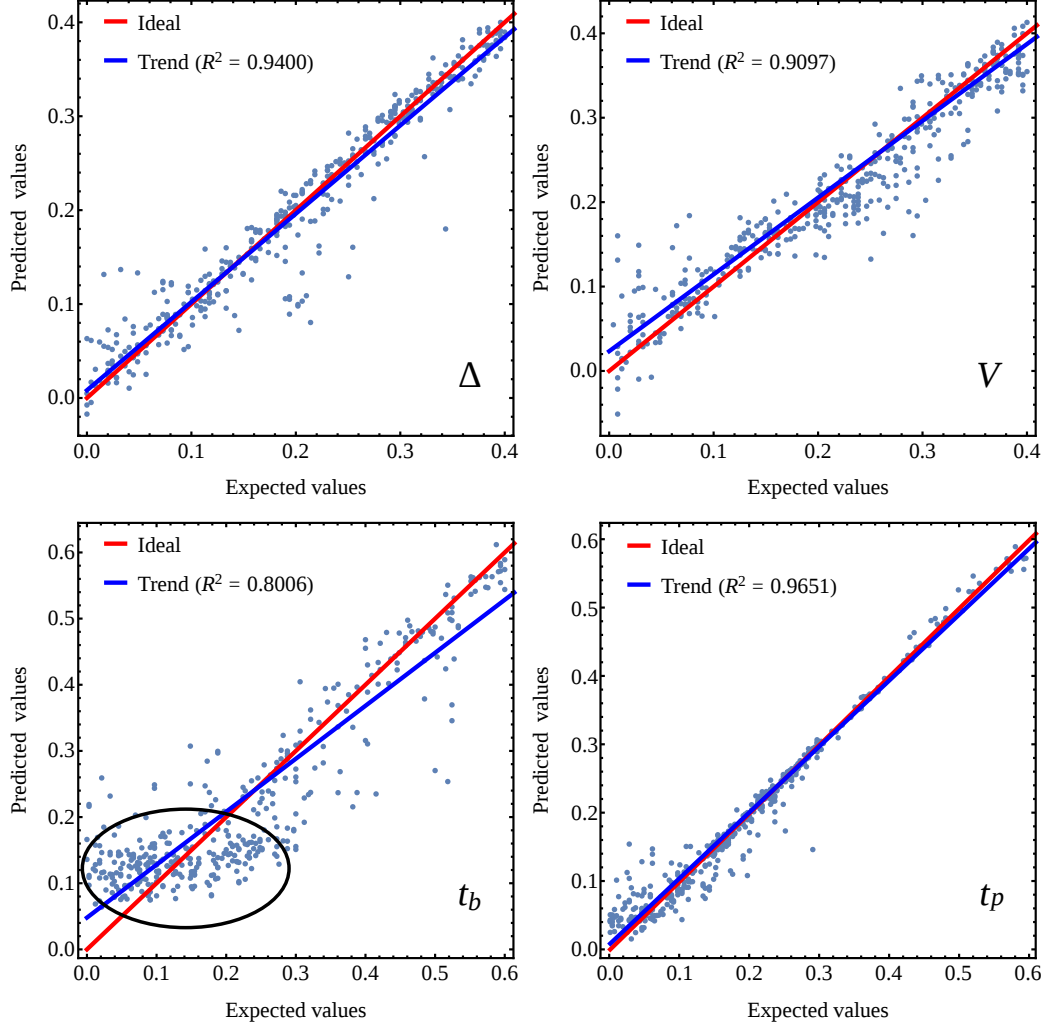


Figure 8: Scatter plots of the predicted and expected parameter values for the heat bath algorithm. For small t_b , the lack of predictive power of the neural network is demonstrated due to the independence of phase diagrams on t_b . However, in other regions, the determination coefficient R^2 demonstrates excellent convergence

Model \ Parameter	Δ		V		t_b		t_p	
MFA	0.016	0.829	0.013	0.937	0.031	0.967	0.014	0.939
Heat bath	0.028	0.940	0.034	0.919	0.033	0.861	0.025	0.965

Table 2: Best U-Net prediction quality metrics (RMSE | R^2) for different model parameters and various methods to obtain phase diagrams

Comparison of results for two approaches

Table 1 shows the prediction quality metrics (RMSE and R^2) for the best neural network predictions based on the U-Net architecture for each of the four model parameters. These results pertain to the largest training datasets and are evaluated for both the MFA approach and the heat bath algorithm.

It is important to note that the calculation was performed exclusively within regions of the parameter space with pronounced correlation (excluding the regions of insensitivity outlighted in black in Figures 6 and 8).

Thus, the machine learning model serves as a universal tool for identifying physically significant parametric regions and demonstrates consistent results for different computational methods. The observed limitations in predictive capability reflect the fundamental properties of the physical system rather than architectural deficiencies of the U-Net model. This confirms the validity of this approach for analyzing the parametric sensitivity of complex multi-parametric systems, including those with quantum effects.

A promising research direction involves investigating transfer learning between different computational methods. Pre-training the model on extensive MFA datasets followed by fine-tuning on limited heat bath algorithm data could significantly enhance learning efficiency and prediction accuracy. This approach would offset the computational cost of more accurate but complex methods by using fast approximate calculations for model initialization. Cross-method transfer learning research is especially relevant for problems where obtaining more accurate data is computationally constrained.

Conclusions

This study demonstrates the effectiveness of machine learning for solving the inverse problem of reconstructing the parameters of a model Hamiltonian for a cuprate superconductor Hamiltonian based on its phase diagram.

We considered a pseudospin model of cuprates, where a pseudospin $S = 1$ describes the three multi-electron states of the $[\text{CuO}_4]^{7-,6-,5-}$ centers. The complex multi-parametric Hamiltonian includes local and non-local density-density correlations for $[\text{CuO}_4]^{7-,5-}$ centers, Heisenberg exchange interaction for magnetic $[\text{CuO}_4]^{6-}$ centers, as well as three types of correlated single-particle transfers and a two-particle transfer. Phase diagrams in the temperature–density variables were obtained using two approaches: mean-field approximation and heat bath Monte Carlo algorithm.

Each parameter of the model Hamiltonian influences the resulting phase diagrams in a complex and highly sensitive manner. This complexity makes the inverse problem of precisely tuning parameters to reproduce experimental results an extremely resource-intensive task.

To address this inverse problem, three deep learning architectures were considered: VGG, ResNet, and U-Net. A comparative analysis revealed a clear advantage of U-Net architecture both in prediction accuracy and computational efficiency. This is attributed to the ability of U-Net to preserve spatial context through its skip-connection mechanism, which is proved crucial for the analysis of phase diagrams.

The use of a U-Net architecture adapted for regression tasks enabled high-accuracy predictions across all considered model parameters. Analysis of regions with low correlations between predicted and expected parameters revealed fundamental limitations in predictive capabilities, which are not related to shortcomings of the machine learning model, but rather to physical properties inherent to the system.

Areas of low prediction accuracy for parameters correspond to regions of parametric insensitivity in the phase diagrams. This indicates that neural networks can serve as effective tools for identifying the significance of model parameters on the system. This was confirmed by parametric analysis for both the mean-field approximation and the heat bath Monte Carlo algorithm. The latter demonstrated a greater insensitivity to the two-particle transport parameter across a wide parameter range. Thus, the U-Net neural network proved itself not merely as a “black box” but as a tool for verifying the physical properties of the system.

The model demonstrated high performance across different computational methods. Despite a significantly smaller training dataset for the heat bath algorithm data, the U-Net maintained high prediction accuracy ($R^2 > 0.9$ for most parameters), which confirms the reliability of the proposed approach and its applicability to real computational data.

Future prospects for this work involve extending the model to incorporate additional parameters, utilizing more complex neural network architectures, investigating transfer learning between different computational methods, and extending the approach to other multi-parametric physical models, including complex quantum systems with strong correlations. The proposed method opens up the possibility of automated selection of theoretical model parameters based on experimental phase diagrams. This is particularly significant in the light of increasing computational complexity of modern physical models.

Acknowledgments

The research was supported by the Russian Science Foundation, grant no. 24-21-20147.

References

- [1] A. S. Moskvina, True charge-transfer gap in parent insulating cuprates, *Physical Review B* 84 (7) (2011) 075116. doi:10.1103/PhysRevB.84.075116.
- [2] A. Moskvina, Perspectives of disproportionation driven superconductivity in strongly correlated 3d compounds, *Journal of Physics: Condensed Matter* 25 (8) (2013) 085601. doi:10.1088/0953-8984/25/8/085601.
- [3] A. Moskvina, Y. Panov, Model of charge triplets for high- T_c cuprates, *Journal of Magnetism and Magnetic Materials* 550 (2022) 169004. doi:10.1016/j.jmmm.2021.169004.
- [4] Y. D. Panov, Critical Temperatures of a Model Cuprate, *Physics of Metals and Metallography* 120 (13) (2019) 1276–1281. doi:10.1134/S0031918X19130222.
- [5] A. S. Moskvina, Y. D. Panov, Nature of the Pseudogap Phase of HTSC Cuprates, *Physics of the Solid State* 62 (9) (2020) 1554–1561. doi:10.1134/S1063783420090206.
- [6] A. Moskvina, Y. Panov, Effective-Field Theory for Model High- T_c Cuprates, *Condensed Matter* 6 (3) (2021) 24. doi:10.3390/condmat6030024.

- [7] Y. D. Panov, A. S. Moskvina, A. A. Chikov, V. A. Ulitko, Monte carlo simulation of a model cuprate, *Journal of Physics: Conference Series* 2043 (1) (2021) 012007. doi:10.1088/1742-6596/2043/1/012007.
- [8] V. Ulitko, Y. D. Panov, A. Moskvina, Classical monte carlo algorithm for simulation of a pseudospin model for cuprates, *arXiv preprint arXiv:2301.11708* (2023). doi:10.48550/arXiv.2301.11708.
- [9] J. Lund, H. Wang, R. Braatz, R. García, Machine learning of phase diagrams, *Materials Advances* 3 (23) (2022) 8485–8497. doi:10.1039/D2MA00524G.
- [10] C. Casert, T. Vieijra, J. Nys, J. Ryckebusch, Interpretable machine learning for inferring the phase boundaries in a nonequilibrium system, *Physical Review E* 99 (2) (2019) 023304. doi:10.1103/PhysRevE.99.023304.
- [11] F. Zipoli, V. Viterbo, O. Schilter, L. Kahle, T. Laino, Prediction of phase diagrams and associated phase structural properties, *Industrial & Engineering Chemistry Research* 61 (24) (2022) 8378–8389. doi:10.1021/acs.iecr.2c00355.
- [12] D. Bayo, B. Çivitcioglu, J. J. Webb, A. Honecker, R. A. Römer, Machine learning of phases and structures for model systems in physics, *Journal of the Physical Society of Japan* 94 (3) (2025) 031002. doi:10.7566/JPSJ.94.031002.
- [13] M. Aghaaminiha, S. A. Ghanadian, E. Ahmadi, A. M. Farnoud, A machine learning approach to estimation of phase diagrams for three-component lipid mixtures, *Biochimica et Biophysica Acta (BBA)-Biomembranes* 1862 (9) (2020) 183350. doi:10.1016/j.bbamem.2020.183350.
- [14] J. Sun, J. Qu, C. Zhao, X. Zhang, X. Liu, J. Wang, C. Wei, X. Liu, M. Wang, P. Zeng, et al., Precise prediction of phase-separation key residues by machine learning, *Nature Communications* 15 (1) (2024) 2662. doi:10.1038/s41467-024-46901-9.
- [15] E. P. Van Nieuwenburg, Y.-H. Liu, S. D. Huber, Learning phase transitions by confusion, *Nature Physics* 13 (5) (2017) 435–439. doi:10.1038/nphys4037.

- [16] W. Zhang, J. Liu, T.-C. Wei, Machine learning of phase transitions in the percolation and xy models, *Physical Review E* 99 (3) (2019) 032142. doi:10.1103/PhysRevE.99.032142.
- [17] A. Canabarro, F. F. Fanchini, A. L. Malvezzi, R. Pereira, R. Chaves, Unveiling phase transitions with machine learning, *Physical Review B* 100 (4) (2019) 045129. doi:10.1103/PhysRevB.100.045129.
- [18] H. Y. Kwon, H. Yoon, C. Lee, G. Chen, K. Liu, A. Schmid, Y. Wu, J. Choi, C. Won, Magnetic hamiltonian parameter estimation using deep learning techniques, *Science advances* 6 (39) (2020) eabb0872. doi:10.1126/sciadv.abb0872.
- [19] F. A. Gómez Albarracín, H. D. Rosales, Machine learning techniques to construct detailed phase diagrams for skyrmion systems, *Physical Review B* 105 (21) (2022) 214423. doi:10.1103/PhysRevB.105.214423.
- [20] Y. Che, C. Gneiting, T. Liu, F. Nori, Topological quantum phase transitions retrieved through unsupervised machine learning, *Physical Review B* 102 (13) (2020) 134213. doi:10.1103/PhysRevB.102.134213.
- [21] V. Stanev, C. Oses, A. G. Kusne, E. Rodriguez, J. Paglione, S. Curtarolo, I. Takeuchi, Machine learning modeling of superconducting critical temperature, *npj Computational Materials* 4 (1) (2018) 29.
- [22] Y. Zhang, X. Xu, Predicting the superconducting transition temperature of high-temperature layered superconductors via machine learning, *Physica C: Superconductivity and its Applications* 595 (2022) 1354031. doi:10.1016/j.physc.2022.1354031.
- [23] S. Xie, Y. Quan, A. Hire, B. Deng, J. DeStefano, I. Salinas, U. Shah, L. Fanfarillo, J. Lim, J. Kim, et al., Machine learning of superconducting critical temperature from eliashberg theory, *npj Computational Materials* 8 (1) (2022) 14. doi:10.1038/s41524-021-00666-7.
- [24] P. J. G. Nieto, E. G. Gonzalo, L. A. M. García, L. Álvarez-de Prado, A. B. Sánchez, Predicting the critical superconducting temperature using the random forest, mlp neural network, m5 model tree and multivariate linear regression, *Alexandria Engineering Journal* 86 (2024) 144–156. doi:10.1016/j.aej.2023.11.034.

- [25] Y. D. Panov, A. S. Moskvina, A. A. Chikov, I. L. Avvakumov, Competition of Spin and Charge Orders in a Model Cuprate, *J. Supercond. Novel Magn.* 29 (4) (2016) 1077–1083. doi:10.1007/s10948-016-3378-5.
- [26] R. Micnas, J. Ranninger, S. Robaszkiewicz, Superconductivity in narrow-band systems with local nonretarded attractive interactions, *Reviews of Modern Physics* 62 (1) (1990) 113. doi:10.1103/RevModPhys.62.113.
- [27] E. Fradkin, S. A. Kivelson, J. M. Tranquada, Colloquium : Theory of intertwined orders in high temperature superconductors, *Reviews of Modern Physics* 87 (2) (2015) 457–482. doi:10.1103/RevModPhys.87.457.
- [28] Y. Miyatake, M. Yamamoto, J. Kim, M. Toyonaga, O. Nagai, On the implementation of the ‘heat bath’ algorithms for monte carlo simulations of classical heisenberg spin systems, *Journal of Physics C: solid state physics* 19 (14) (1986) 2539. doi:10.1088/0022-3719/19/14/020.
- [29] Y. D. Panov, S. V. Nuzhin, V. S. Ryumshin, A. S. Moskvina, The monte carlo method for the orthonickelate model, *Fizika Tverdogo Tela* 66 (7) (2024) 1221–1227.
- [30] Y. D. Panov, V. A. Ulitko, D. N. Yasinskaya, A. S. Moskvina, Modified monte carlo method with the heat bath algorithm for a model cuprate, *arXiv preprint arXiv:2510.08289v1* (2025). doi:10.1088/1742-6596/2043/1/012007.
- [31] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, *arXiv preprint arXiv:1409.1556* (2014). doi:10.48550/arXiv.1409.1556.
- [32] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [33] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: *International Conference on Medical image computing and computer-assisted intervention*, Springer, 2015, pp. 234–241. doi:10.1007/978-3-319-24574-4_28.

- [34] R. Azad, E. K. Aghdam, A. Rauland, Y. Jia, A. H. Avval, A. Bozorgpour, S. Karimijafarbigloo, J. P. Cohen, E. Adeli, D. Merhof, Medical image segmentation review: The success of u-net, *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2024). doi:10.1109/TPAMI.2024.3435571.
- [35] W.-J. Lee, W.-T. Hsieh, B.-H. Fang, K.-H. Kao, N.-Y. Chen, Device simulations with a u-net model predicting physical quantities in two-dimensional landscapes, *Scientific reports* 13 (1) (2023) 731. doi:10.1038/s41598-023-27599-z.
- [36] J. Ye, Y. Huang, K. Liu, U-net based vortex detection in bose–einstein condensates with automatic correction for manually mislabeled data, *Scientific Reports* 13 (1) (2023) 21278. doi:10.1038/s41598-023-48719-9.
- [37] A. Kamali, K. Laksari, Physics-informed unets for discovering hidden elasticity in heterogeneous materials, *Journal of the mechanical behavior of biomedical materials* 150 (2024) 106228. doi:10.1016/j.jmbbm.2023.106228.
- [38] B.-R. Chen, Y.-S. Hsiao, W.-C. Lin, W.-J. Lee, N.-Y. Chen, T.-L. Wu, Using u-net convolutional neural network to model pixel-based electrostatic potential distributions in gan power mis-hemts, *Scientific reports* 14 (1) (2024) 8151. doi:10.1038/s41598-024-58112-9.
- [39] W. Yao, Z. Zeng, C. Lian, H. Tang, Pixel-wise regression using u-net and its application on pansharpening, *Neurocomputing* 312 (2018) 364–371.
- [40] D. Gertsvolf, M. Horvat, D. Aslam, A. Khademi, U. Berardi, A u-net convolutional neural network deep learning model application for identification of energy loss in infrared thermographic images, *Applied Energy* 360 (2024) 122696.