Uncovering Conformal Towers Using Deep Learning

Lior Oppenheim
The Racah Institute of Physics
The Hebrew University of Jerusalem
Jerusalem 9190401, Israel

Maciej Koch-Janusz
Haiqu Inc.
San Francisco 94103, California, USA

Snir Gazit
The Racah Institute of Physics
The Hebrew University of Jerusalem
Jerusalem 9190401, Israel

Zohar Ringel
The Racah Institute of Physics
The Hebrew University of Jerusalem
Jerusalem 9190401, Israel

Abstract

Extracting the operator spectrum (conformal towers) of critical models with space-time dimensionality larger than 2 is a formidable numerical task, closely related to diagonalizing very large element-wise non-negative matrices. Here we demonstrate the ability of a new ML-based numerical tool (extended RSMI-NE) to tackle such problems. We focus on critical properties of the Ising-Higgs gauge theory in $(2+1)D$ along the self-dual line, which has recently been a subject of debate. We determine, for the first time, the low energy operator content of the associated field-theory. Our approach enables us to largely refute a standing conjecture about the universality class of this transition.

Introduction

The hallmark of critical phenomena is the emergence of a universal behavior governing the long wave-length theory. In this limit, dynamics are often controlled by collective degrees of freedom dictated solely by symmetry and dimensionality as in symmetry-breaking transitions, whose critical fluctuations are governed by an order parameter directed along the symmetry-breaking axis [1].

A major challenge in modern condensed matter theory is addressing critical phenomena beyond the above Landau paradigm [2, 3]. This includes spin liquids [4], fractional Hall effect [5], and symmetry protected topological transitions [6, 7]. In such cases, identifying the low-energy theory often becomes a formidable task, due to the absence of clear symmetry-based candidates for the low-lying degrees of freedom [8, 9].

This outstanding problem attracted much recent interest, particularly the case of the Self-Dual Ising-Higgs Gauge theory in $(2+1)D$ ("SD-IHG")[10, 11, 12, 13, 14, 15, 16, 17, 18]. Here the two transitions, meeting at a multicritical point (MCP), are of the $3D$ Ising and Ising* universality classes. Following the general insight that MCPs of two Ising transitions may lead to a continuous $U(1)$ symmetry Bonati *et al.* [17] argued, with supporting Monte Carlo results, that the emergent theory is of the XY^* type [19] — an XY transition exhibiting only gauge invariant operators. However, in the absence of a direct identification of low energy degrees of freedom in terms of the microscopic ones, the validity of such a phenomenological description is unclear.

The ideal way to verify this conjecture is to obtain the operator spectrum of the theory, or at least its leading orders. Indeed, the putative XY^* transition should contain a smoking gun: three degenerate operators with scaling dimension 2, namely the three vector components of the current operator associated with the emergent $U(1)$ symmetry. While for $2D$ critical points, such data is readily

Machine Learning and the Physical Sciences Workshop, NeurIPS 2023.

accessible through transfer matrix diagonalization [20], in $(2+1)D/3D$ it is a challenging numerical problem. Despite recent progress [21, 22, 23], we currently lack a generic tool for this task.

More broadly, extracting the operator content beyond the leading order from microscopic samples, thus constructively connecting the micro- and macroscopic descriptions, is an open challenge in many fields. Recently, methods based on information theory and deep learning have shown promise in this task [24, 25, 26, 27, 28, 29]. One approach is the Real-Space Mutual Information Neural-Estimator (RSMI-NE) algorithm, which was used to identify and extract leading operators in the field theory from microscopic Monte Carlo samples [27, 30, 31]. The possibility of using such techniques to extract sub-leading parts of the operator spectrum remained, however, unexplored.

In this work we address the problem of systematic extraction of sub-leading orders of operators in the spectrum by extending the RSMI-NE algorithm. Applying this technique to the multi-critical point of the SD-IHG theory in $(2+1)D$ we obtain both leading and sub-leading operators, namely the energy operator and its derivatives. Crucially, a current operator required by the conjecture [17] does *n*ot appear in the spectrum. This is in contrast to a model known to exhibit an emergent $U(1)$ symmetry based on coinciding Ising transitions, where we obtain all expected operators, *up to and including the current operator*. We thus rule out the existence of a local current operator for the SD-IHG theory and, with it, the classification of the critical theory as XY^*. More broadly, we showcase a numerically tractable scheme of extracting the leading spectrum of exponentially-large transfer matrices, which is highly relevant to many disciplines in physics.

Models

We investigate two $(2+1)D$ quantum models on a discrete lattice: the SD-IHG model – the principal subject of interest – and the Ashkin-Teller Transverse Field Ising (AT-TFI) model, which is used to compare and contrast the numerical results [21].

The SD-IHG model describes $\mathbb{Z}_2$ gauge- and matter- fields $\sigma_{ij} = \pm 1$ and $\tau_i = \pm 1$, residing, respectively, on the bonds and sites of a cubic $(2+1)D$ lattice (as conventional 2 refers to spatial and 1 to temporal dimensions). The classical action is given by (see Fig. 1):

$$\mathcal{S}_{\text{SD-IHG}} = K \sum_{\square} \prod_{\langle i,j \rangle \in \square} \sigma_{ij} + J \sum_{\langle i,j \rangle} \tau_i \sigma_{ij} \tau_j \tag{1}$$

The gauge-invariant quantities are either a closed loop of gauge fields or two matter fields with a string of gauge fields stretched between them. The model is self-dual along the line $\tanh(K) = \frac{1-\tanh(J)}{1+\tanh(J)}$. The model exhibits two $\mathbb{Z}_2$ transitions, which meet at the self-dual line to form an MCP (see Fig. 1). Though the universality of the MCP is yet unknown, two relevant primary operators in the spectrum can be inferred using the duality symmetry (denoted A and S) [14].

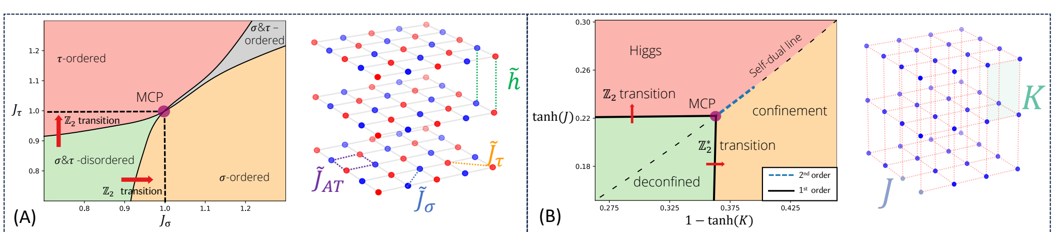


Figure 1: *(A)* Phase diagram and visual representation of the Hamiltonian (Eq. 2) for the AT-TFI model. The four phases correspond to the separate ordering of the two Ising fields. The blue and red sites are the σ and τ fields. The $\tilde{J}_\sigma, \tilde{J}_\tau, \tilde{J}_{AT}, \tilde{h}$ are the coupling constants after the quantum-to-classical mapping. *(B)* Phase diagram and visual representation of the action (Eq. 1) for the SD-IHG model. The two Ising second-order transitions meet at the MCP. A first-order transition also takes place along the self-dual line. The blue sites represent the matter field τ_i and the red bonds represent the gauge field σ_{ij}.

The second system we consider is the AT-TFI model in $(2+1)D$, which serves as a benchmark of our approach, as it contains a fully understood, yet non-trivial, critical point of the XY^* universality class, described by a similar field theory to the one proposed in [17]. The model is phrased in terms of quantum spins $\hat{\sigma}_i, \hat{\tau}_i$ residing on interlaced sublattices of a $2D$ square lattice. The 2D quantum Hamiltonian (equivalent by quantum-to-classical mapping to a model on a $3D$ classical lattice, see Fig. 1) is given by [21]:

$$\hat{\mathcal{H}}_{\text{AT-TFI}} = -\sum_{\langle i,j \rangle} \left[J_\sigma \hat{\sigma}_i^z \hat{\sigma}_j^z + J_\tau \hat{\tau}_i^z \hat{\tau}_j^z - J_{AT} \hat{\sigma}_i^z \hat{\sigma}_j^z \hat{\tau}_i^z \hat{\tau}_j^z \right] - h \sum_i \hat{\sigma}_i^x + \hat{\tau}_i^x, \tag{2}$$

At the MCP, the σ and τ fields merge into a continuous complex field $\psi = \sigma + i\tau$ which forms a $|\psi|^4$ theory [21]. The relevant part of the spectrum is fully characterized and consists of three primary operators: charge-0 ($\sigma^2 + \tau^2$), charge-2 ($\sigma^2 - \tau^2, \sigma\tau$), and a Noether current ($\sigma\partial\tau - \tau\partial\sigma$).

Methods

Recently, a correspondence between the solutions to a certain information-theoretic variational problem and the leading operators in the transfer matrix spectrum was shown [32]. This result explains how leading eigenvectors and eigenvalues of the transfer matrix can be learned using the framework of the information-bottleneck (IB) compression theory [33]. It further suggests that IB can also be used to extract sub-leading operators of the spectrum.

In parallel, motivated by progress in mutual information estimation algorithms [34, 35, 36] and using a different theoretical MI-based goal (similar to, yet distinct from IB), a deep-learning scheme, RSMI-NE, was devised to detect the leading local operator based solely on a corpus of Monte Carlo samples [27, 30]. It takes a pair of random variables $(\mathscr{V}, \mathscr{E})$ where $\mathscr{V}$ is a spatial block of the system and $\mathscr{E}$ is a distant environment of $\mathscr{V}$, spatially separated from it by a buffer $\mathscr{B}$, and output a compressed encoding $P(\mathscr{H}|\mathscr{V})$ (where $|H|$ has a fixed cardinality) that maximizes the mutual information $I(\mathscr{H};\mathscr{E})$. Effectively, $P(\mathscr{H}|\mathscr{V})$ tracks the leading operator that can be extracted from $\mathscr{V}$. The method has already proven successful in $2D$ systems, including interacting spin and dimer models, on regular and aperiodic lattices [24, 30, 27, 31].

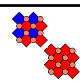
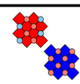
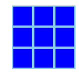
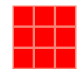
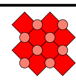
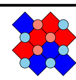
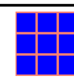
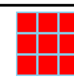
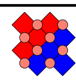
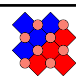
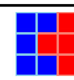
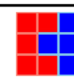
In this work, we refine and extend the RSMI-NE in two important ways: in contrast to the previous works we make essential use of non-linear operators parameterized by neural networks (as opposed to linear maps in [27, 30]), and we extend the algorithm to allow to systematically uncover not only the leading but also sub-leading operators in the spectrum, which is typically very difficult.

As a key step in learning non-linear operators we introduce a batch normalization layer controlling the variance of $\mathscr{H}$, and thus limiting the amount of information being learned. This adds noise to the encoding and effectively introduces a compression penalty term similar to IB. This ensures that [32] the non-linear encoder is pressured by the optimization process to learn the pristine operator.

In order to extract sub-leading operators, we devised two methods. In the first method, the encoding in $\mathscr{H}$ is learned in a consecutive manner, where in each step a new encoder is learned based on all of the previous ones $P(\mathscr{H}_N|\mathscr{H}_1\mathscr{H}_2 \dots \mathscr{H}_{N-1}\mathscr{V})$ ($|\mathscr{H}| = 2$), which are held constant during the training process of the N'th encoder. The RSMI-NE neural network is then forced to learn in each step only the encoder that yields the biggest change in the overall mutual information, given all the previous encoders. The second method makes use of symmetries in order to focus on operators lying in a particular symmetry sector. Given a symmetry group G, we can partially symmetrize the dataset by acting on samples of $\mathscr{V}$ (but not of $\mathscr{E}$) with random elements in G, which is a form of data augmentation. This symmetrization washes out the information gained from non-symmetric operators in the spectrum. By employing this method, one can directly target the leading operator in the G-invariant symmetry sector. This procedure can also be used in order to learn the leading operator which is *not* invariant under the symmetry.

The output of the extended algorithm is an ordered set of encoders parameterized by neural networks ("neural operators"). As such, this tool can provide a complete signature of the underlying universality, and unlike exact diagonalization approaches, *e.g.* the critical torus energy spectrum (CTES) [37], which scale exponentially with system size, we expect it to scale polynomially.

Table 1: AT-TFI and SD-IHG leading operators content

AT-TFI				SD-IHG			
RSMI-NE Scaling Dimension {Expected[38]}	Analytic Operator {Deg.}	Neural Operator Projection		RSMI-NE Scaling Dimension {Expected[14]}	Analytic Operator {Deg.}	Neural Operator Projection	
		Maximum	Minimum			Maximum	Minimum
$1.24(1)$ $1.22(1)$ $\{1.23629\}$	$\langle\sigma\rangle^2 - \langle\tau\rangle^2$ $\langle\sigma\rangle\langle\tau\rangle$ $\{2\}$			$1.24(1)$ $\{1.222\}$	$\langle A\rangle$ $\{1\}$		
$1.49(2)$ $\{1.51136\}$	$\langle\sigma\rangle^2 + \langle\tau\rangle^2$ $\{1\}$			$1.54(2)$ $\{1.502\}$	$\langle S\rangle$ $\{1\}$		
$2.02(3)$ $\{2.0\}$	$\langle\sigma\rangle\langle\partial\tau\rangle - \langle\tau\rangle\langle\partial\sigma\rangle$ $\{3\}$			$2.20(6)$ $\{2.222\}$	$\langle\partial A\rangle$ $\{3\}$		

AT-TFI and SD-IHG three leading operators at their MCPs identified using the RSMI-NE method, including their scaling dimensions and degeneracies. Expected scaling dimensions are taken from [38] and [14] respectively. Errors are purely statistical. $2D$ slices of the $(2+1)D$ local configurations that maximize/minimize the neural operators are shown (blue and red colors denote the spin states). In the AT-TFI case, the τ sub-lattice appears as plaquettes for clarity. In the SD-IHG case, we draw the values of the gauge invariant plaquettes and bonds.

Results

By applying the extended RSMI-NE method to the well-understood case of the AT-TFI model, we extracted the first three leading primary operators (and their degeneracies), including the current operator, *i.e.* the operator of scaling dimension 2. Higher operators can also be constructed by iterating the procedure but are less relevant in the context of the field-theoretic problem studied.

The left column of Table 1 presents the computed scaling dimensions of the neural operators extracted (using Widom scaling [39]), which are in agreement with the theoretically expected values of the $U(1)$ theory [38]. Slight discrepancies in scaling dimensions are attributed to a mixture between operators within the same symmetry sector. Further, by computing operator-operator correlations between a neural operator and the analytically known operators, we could unambiguously identify the neural operator either as a superposition of known operators, or as a yet unknown one.

Moreover, the scaling operators parameterized by neural networks can be accessed directly, rather than just through their scaling exponents. Owing to their non-linearity the analysis and visualization are, however, more complicated than in *e.g.* [27]. To understand their action on the local degrees of freedom, we calculate configurations that extremize the values for the numerical operator. Example $2D$ spatial slices through these $(2+1)D$ configurations are shown in Table 1.

Having validated the method and its ability to detect non-linear relevant operators in the CFT spectrum, particularly the emergent currents, we apply it to the intriguing case of the SD-IHG model. We detect the three leading operators in the CFT spectrum, namely the S and A primary operators and the descendant ∂A. The operators generally act non-linearly on the gauge-invariant constituents, *i.e.* plaquettes and bonds.

The right column of Table 1 shows a very good correspondence between the scaling dimensions of the *two* leading neural operators and the values of the theoretically expected operators A and S. Strikingly, however, the next three neural operators are inconsistent with a conjectured current operator (and have a higher scaling dimension). No operator with the characteristics of a current (namely a vector operator with a scaling dimension of 2.0) has been found. Its absence in the RSMI-NE results (in contrast to the AT-TFI case) is a strong indication that no such operator exists, and, therefore, the self-dual MCP of the SD-IHG theory does not belong to the XY^* universality class.

Conclusions and Outlook

We demonstrated that recently developed numerical RG methods based on information theory and ML can be brought to a level where they shed light on the current open questions in field theory. In particular, we provide strong evidence against the hypothesis that the multi-critical point of the $(2+1)D$ self-dual Ising-Higgs gauge theory belongs to the XY^* universality class, showcasing the ability of our extension of the RSMI-NE algorithm to extract conformal data for high-dimensional systems, including sub-leading and descendant operators.

We expect that our extension of RSMI-NE will be a valuable addition to the arsenal of numerical tools in statistical physics and field theory, especially for critical phenomena beyond the Landau paradigm. Apart from extracting the conformal tower including operators that are difficult to resolve by symmetry, it can guide the construction of field-theoretical description by providing microscopic interpretations for the most relevant degrees of freedom.

References

[1] L. Landau and E. Lifshitz, *Statistical Physics: Volume 5 (3rd Edition)*. Butterworth-Heinemann, 1980.

[2] S. Sachdev and X. Yin, "Quantum phase transitions beyond the Landau–Ginzburg paradigm and supersymmetry," *Annals of Physics*, vol. 325, no. 1, pp. 2–15, 2010.

[3] C. Xu, "Unconventional quantum critical points," *International Journal of Modern Physics B*, vol. 26, no. 18, p. 1230007, 2012.

[4] W. X. Gang, *Quantum field theory of many-body systems: from the origin of sound to an origin of light and electrons*, pp. 354–440. Oxford: Oxford University Press, 2007.

[5] S. M. Girvin, *Introduction to the Fractional Quantum Hall Effect*, pp. 133–162. Basel: Birkhauser Basel, 2005.

[6] M. Cheng, M. Zaletel, M. Barkeshli, A. Vishwanath, and P. Bonderson, "Translational Symmetry and Microscopic Constraints on Symmetry-Enriched Topological Phases: A View from the Surface," *Phys. Rev. X*, vol. 6, p. 041068, 2016.

[7] M. Barkeshli, P. Bonderson, M. Cheng, and Z. Wang, "Symmetry fractionalization, defects, and gauging of topological phases," *Phys. Rev. B*, vol. 100, p. 115147, 2019.

[8] T. Senthil, L. Balents, S. Sachdev, A. Vishwanath, and M. P. A. Fisher, "Quantum criticality beyond the Landau-Ginzburg-Wilson paradigm," *Physical Review B*, vol. 70, no. 14, 2004.

[9] S. Sachdev, "Topological order, emergent gauge fields, and Fermi surface reconstruction," *Reports on Progress in Physics*, vol. 82, no. 1, p. 014001, 2018.

[10] A. Kitaev, "Fault-tolerant quantum computation by anyons," *Annals of Physics*, vol. 303, no. 1, pp. 2–30, 2003.

[11] J. Vidal, S. Dusuel, and K. P. Schmidt, "Low-energy effective theory of the toric code model in a parallel magnetic field," *Physical Review B*, vol. 79, no. 3, 2009.

[12] I. S. Tupitsyn, A. Kitaev, N. V. Prokof'ev, and P. C. E. Stamp, "Topological multicritical point in the phase diagram of the toric code model and three-dimensional lattice gauge Higgs model," *Physical Review B*, vol. 82, no. 8, 2010.

[13] S. Gazit, F. F. Assaad, S. Sachdev, A. Vishwanath, and C. Wang, "Confinement transition of $\mathbb{Z}_2$ gauge theories coupled to massless fermions: Emergent quantum chromodynamics and $SO(5)$ symmetry," *Proceedings of the National Academy of Sciences*, vol. 115, no. 30, 2018.

[14] A. M. Somoza, P. Serna, and A. Nahum, "Self-Dual Criticality in Three-Dimensional $\mathbb{Z}_2$ Gauge Theory with Matter," *Phys. Rev. X*, vol. 11, p. 041008, 2021.

[15] N. Iqbal and J. McGreevy, "Mean string field theory: Landau-Ginzburg theory for 1-form symmetries," *SciPost Phys.*, vol. 13, p. 114, 2022.

[16] C. Bonati, A. Pelissetto, and E. Vicari, "Scalar gauge-Higgs models with discrete Abelian symmetry groups," *Phys. Rev. E*, vol. 105, p. 054132, 2022.

[17] C. Bonati, A. Pelissetto, and E. Vicari, "Multicritical point of the three-dimensional gauge Higgs model," *Physical Review B*, vol. 105, no. 16, 2022.

[18] N. Manoj and V. B. Shenoy, "Arboreal topological and fracton phases," *Phys. Rev. B*, vol. 107, p. 165136, 2023.

[19] S. V. Isakov, R. G. Melko, and M. B. Hastings, "Universal Signatures of Fractionalized Quantum Critical Points," *Science*, vol. 335, no. 6065, pp. 193–195, 2012.

[20] J. L. Cardy, "Operator content of two-dimensional conformally invariant theories," *Nuclear Physics B*, vol. 270, pp. 186–204, 1986.

[21] M. Schuler, L.-P. Henry, Y.-M. Lu, and A. M. Läuchli, "Emergent XY* transition driven by symmetry fractionalization and anyon condensation," *SciPost Phys.*, vol. 14, p. 001, 2023.

[22] Z. Zhou, L. Hu, W. Zhu, and Y.-C. He, "The $\mathrm{SO}(5)$ Deconfined Phase Transition under the Fuzzy Sphere Microscope: Approximate Conformal Symmetry, Pseudo-Criticality, and Operator Spectrum," 2023.

[23] W. Zhu, C. Han, E. Huffman, J. S. Hofmann, and Y.-C. He, "Uncovering Conformal Symmetry in the 3D Ising Transition: State-Operator Correspondence from a Quantum Fuzzy Sphere Regularization," *Physical Review X*, vol. 13, no. 2, 2023.

[24] M. Koch-Janusz and Z. Ringel, "Mutual information, neural networks and the renormalization group," *Nature Physics*, vol. 14, no. 6, pp. 578–582, 2018.

[25] S.-H. Li and L. Wang, "Neural Network Renormalization Group," *Phys. Rev. Lett.*, vol. 121, p. 260601, 2018.

[26] D. Di Sante, M. Medvidović, A. Toschi, G. Sangiovanni, C. Franchini, A. M. Sengupta, and A. J. Millis, "Deep Learning the Functional Renormalization Group," *Phys. Rev. Lett.*, vol. 129, p. 136402, 2022.

[27] D. E. Gökmen, Z. Ringel, S. D. Huber, and M. Koch-Janusz, "Statistical Physics through the Lens of Real-Space Mutual Information," *Phys. Rev. Lett.*, vol. 127, p. 240603, 2021.

[28] G. Margalit, O. Lesser, T. Pereg-Barnea, and Y. Oreg, "Renormalization-group-inspired neural networks for computing topological invariants," *Phys. Rev. B*, vol. 105, p. 205139, 2022.

[29] Z. Zhang and Y.-Z. You, "Observing schrodinger's cat with artificial intelligence: Emergent classicality from information bottleneck," *arXiv preprint arXiv:2306.14838*, 2023.

[30] D. E. Gökmen, Z. Ringel, S. D. Huber, and M. Koch-Janusz, "Symmetries and phase diagrams with real-space mutual information neural estimation," *Phys. Rev. E*, vol. 104, p. 064106, 2021.

[31] D. E. Gökmen, S. Biswas, S. D. Huber, Z. Ringel, F. Flicker, and M. Koch-Janusz, "Compression theory for inhomogeneous systems," 2023.

[32] A. Gordon, A. Banerjee, M. Koch-Janusz, and Z. Ringel, "Relevance in the Renormalization Group and in Information Theory," *Physical Review Letters*, vol. 126, no. 24, 2021.

[33] N. Tishby, F. C. Pereira, and W. Bialek, "The Information Bottleneck method," in *Proc. of the 37-th Annual Allerton Conference on Communication, Control and Computing*, pp. 368–377, 1999.

[34] M. I. Belghazi, A. Baratin, S. Rajeshwar, S. Ozair, Y. Bengio, A. Courville, and D. Hjelm, "Mutual Information Neural Estimation," in *Proceedings of the 35th International Conference on Machine Learning* (J. Dy and A. Krause, eds.), vol. 80 of *Proceedings of Machine Learning Research*, pp. 531–540, PMLR, 10–15 Jul 2018.

[35] A. A. Alemi, B. Poole, I. Fischer, J. V. Dillon, R. A. Saurous, and K. Murphy, "Fixing a broken elbo," 2018.

[36] A. van den Oord, Y. Li, and O. Vinyals, "Representation Learning with Contrastive Predictive Coding," 2019.

[37] M. Schuler, S. Whitsitt, L.-P. Henry, S. Sachdev, and A. M. Läuchli, "Universal signatures of quantum critical points from finite-size torus spectra: A window into the operator content of higher-dimensional conformal field theories," *Physical Review Letters*, vol. 117, nov 2016.

[38] S. M. Chester, W. Landry, J. Liu, D. Poland, D. Simmons-Duffin, N. Su, and A. Vichi, "Carving out OPE space and precise $O(2)$ model critical exponents," *JHEP*, vol. 06, p. 142, 2020.

[39] A. W. Sandvik and B. Zhao, "Consistent scaling exponents at the deconfined quantum-critical point," *Chinese Physics Letters*, vol. 37, no. 5, p. 057502, 2020.