

A Basis-Dependent Kernel for Learning Quantum Phases of Matter

Clemens Lindner,^{1,*} Vinícius Mohr,² Juan Carrasquilla,² and Adrián Pérez-Salinas²

¹*Faculty of Information Technology, University of Jyväskylä, Finland*

²*Institute for Theoretical Physics, ETH Zürich, Switzerland*

The existence and unraveling of different phases of matter for various configurations of the same system constitutes a central task in quantum many body physics. Recent developments addressed this problem from data, by studying the differences between states, e. g. through fidelity measures, and inferring phase transitions from them. We introduce the IPR-kernel, a kernel method built for learning quantum phases of matter from inverse participation ratios (IPRs) in physically motivated bases. Unlike standard fidelity approaches, the IPR-kernel can identify states with physically-motivated similarity even when they are orthogonal. We demonstrate the method on ground states of the transverse-field Ising (TFIM) and Heisenberg XXZ models, and show for TFIM Gibbs states that the IPR-kernel retains phase-relevant structure at nonzero temperature, while fidelity-based approaches are insensitive to such structure.

1 Motivation

Learning the structure of quantum many-body systems, commonly divided in phases, is a natural setting for quantum machine learning. Computation from first-principles is usually challenging, and recent research has developed data-driven methods to find these phases [1, 2]. Quantum kernel methods [3–8] have already been applied to phase classification [9, 10]. In these approaches, the comparison between states is typically done through a Hilbert-Schmidt (HS) fidelity measure, which follows standard quantum feature-space approaches [11, 12]. One could however argue that two perfectly orthogonal states, e.g. with imperfections, could also belong to the same phase of matter.

In this work, we propose an alternative kernel based on basis-dependent effective sparsity. The guiding intuition comes from multiple-basis representations [13]: a state can be highly localized in one physically meaningful basis and completely delocalized in another. Across a phase transition, the basis in which the state is localized can change. Thus, we compare how states localize across a small set of phase-relevant bases. As a result, the IPR-kernel can remain large for two orthogonal states when those states share the same basis-dependent localization structure. This is especially useful beyond ideal ground states, where mixed states, including noisy or finite-temperature ones, may retain phase information even after direct state overlap has become small.

2 IPR-kernel

For a pure n -qubit state $|\psi\rangle$ and a basis specified by a unitary U_a , the inverse participation ratio is defined as

$$I^{(a)}(\psi) = \sum_{x \in \{0,1\}^n} \left| \langle x | U_a^\dagger | \psi \rangle \right|^4 = \sum_{x \in \{0,1\}^n} (p_x^{(a)})^2. \quad (1)$$

The IPR is $\mathcal{O}(1)$ when $|\psi\rangle$ is concentrated on few basis states and $\exp(-\mathcal{O}(n))$ in the other extreme. IPRs and related Rényi entropies are widely used to quantify Hilbert-space localization [14–17], and quantum algorithms for estimating them have recently been proposed [18], provided that U_a is easy to implement. We collect the IPR values for M bases in a feature vector and define

$$K_{\text{IPR}}(\psi, \phi) = \langle \mathbf{I}(\psi), \mathbf{I}(\phi) \rangle = \sum_{a=1}^M I^{(a)}(\psi) I^{(a)}(\phi), \quad (2)$$

*Email: clindner@jyu.fi

which is positive semi-definite by definition of the IPR. In the numerical classification experiments we use the normalized kernel $\tilde{K}_{\text{IPR}}(\psi, \phi)$, which removes the overall scale of the IPR feature vector and emphasizes its direction. To select the bases U_a , we follow a Hamiltonian depending heuristics. If the system of interest is defined by a Hamiltonian of the form

$$\hat{H} = \sum_a \hat{H}_a, \quad [H_a, H_{a'}] \neq 0, \quad (3)$$

then $\{U_a\}$ are chosen as the eigenbases of $\{H_a\}$.

For mixed states we define the basis-dependent IPR by keeping only the diagonal probabilities in the chosen basis,

$$I^{(a)}(\rho) = \sum_{x \in \{0,1\}^n} \left(\langle x | U_a^\dagger \rho U_a | x \rangle \right)^2. \quad (4)$$

This extension is essential for finite-temperature states. For comparison, we use the normalized HS-kernel $\tilde{K}_{\text{HS}}(\rho, \sigma) \propto \text{Tr}(\rho\sigma)$, which compares density matrices through their normalized Hilbert-Schmidt inner product, reducing to the usual state-overlap kernel for pure states. For large system sizes, $\tilde{K}_{\text{HS}}$ becomes a diagonal Gram matrix due to exponential concentration [19, 20], a phenomenon which doesn't occur for $K_{\text{IPR}}(\psi, \phi)$. The IPR-kernel still suffers from sampling based precision when estimating each IPR value.

3 Ground-state and thermal TFIM

Our learning task uses support vector classification (SVC), i.e. given a dataset with states obtained from a Hamiltonian $H(\lambda_i)$ and phase labels y_i

$$(|\psi(\lambda_i)\rangle, y_i), \quad y_i \in \{-1, +1\}, \quad (5)$$

predict the label of a new state. We test the kernel on the one-dimensional transverse-field Ising model

$$H_{\text{TFIM}}(h) = -J \sum_{i=1}^L Z_i Z_{i+1} - h \sum_{i=1}^L X_i, \quad (6)$$

with $J = 1$ and critical point $h_c = 1$. The two limiting regimes are naturally described by the computational Z basis and the Hadamard X basis. We therefore use

$$\mathbf{I}(\psi_{GS}(h)) = \left(I^{(Z)}(\psi_{GS}(h)), I^{(X)}(\psi_{GS}(h)) \right). \quad (7)$$

Ground states are computed as matrix product states using DMRG for L up to 100, and IPRs are estimated by Monte Carlo sampling from the matrix product states (MPS) Born distribution. The resulting IPR curves crossing drifts towards the critical $h_c = 1$ point with increasing system size, see fig. 1. The normalized IPR-kernel built from these features gives near-perfect ($> 99\%$ balanced accuracy) global and local phase classification, where global testing uses the full $h \in [0, 4]$ range, while local testing restricts the test states to a window $\delta = 0.25$ around h_c .

At finite temperature we consider Gibbs states

$$\rho(\beta, h) = Z^{-1}(\beta, h) e^{-\beta H_{\text{TFIM}}(h)}, \quad Z(\beta, h) = \text{Tr}(e^{-\beta H_{\text{TFIM}}(h)}). \quad (8)$$

The Gibbs states are represented as matrix product operators (MPO), the IPR features are estimated from the corresponding distribution and the HS-kernel is evaluated by MPO contraction.

The distinction between global and local testing is important. Global classification includes states deep inside both phases and is therefore easier. Local classification uses test points near h_c , where the two phases are hardest to distinguish. In this regime, the IPR-kernel shows better local performance (fig. 1) because its decision functions are sharper around the transition, indicating that it captures the phase transition more faithfully.

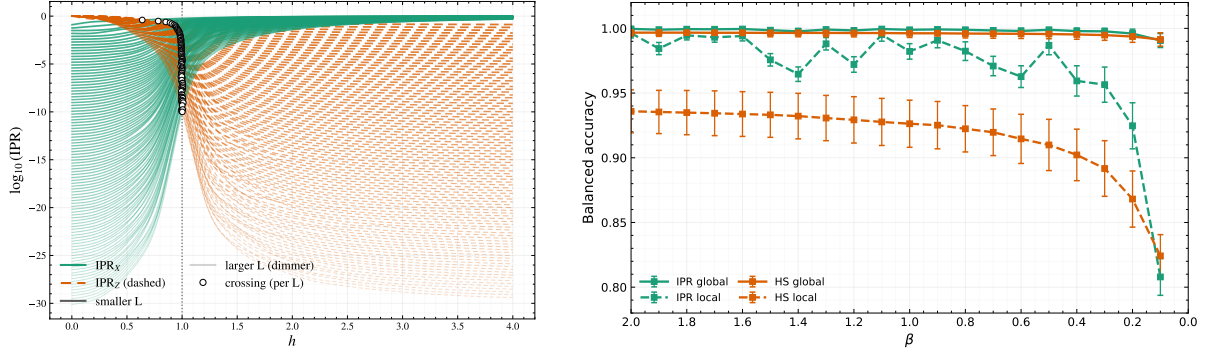


Figure 1: Zero- and finite-temperature TFIM results. Left: Ground-state IPRs in the X and Z bases for $L = 3, \dots, 100$, showing an IPR crossover moving towards $h_c = 1$. Right: Global and local SVM balanced accuracy for thermal TFIM classification at $L = 100$ using the IPR- and HS-kernel, including standard deviation over 100 train-test splits. The local task is more sensitive to whether the kernel resolves the phase boundary.

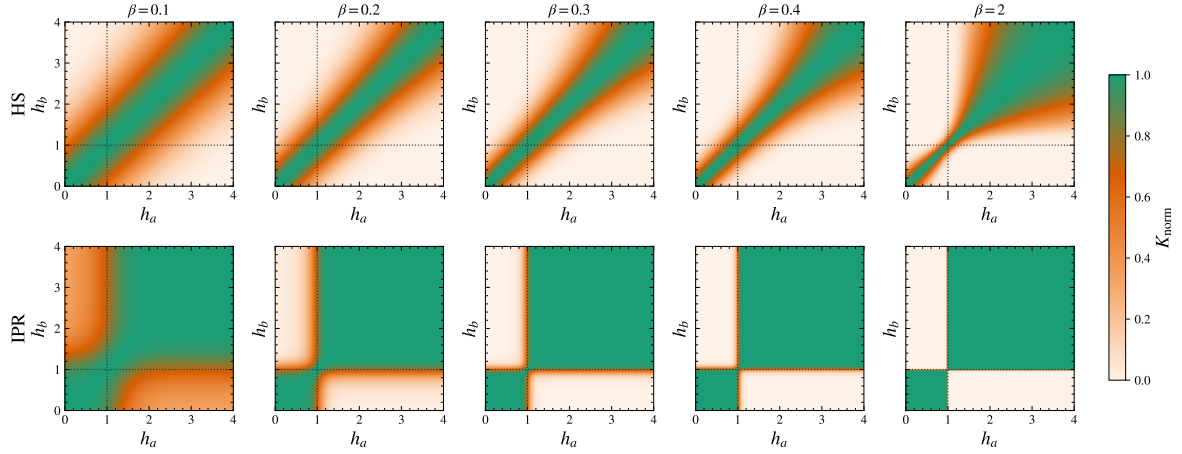


Figure 2: Thermal TFIM Gram matrices at $L = 100$ for varying inverse temperatures. Top row: normalized HS-kernel. Bottom row: normalized IPR-kernel.

The key distinction between the two kernels is visible in fig. 2. At low temperature ($\beta \approx 2$) both kernels recover a phase-structured, block-diagonal Gram matrix, which then gets lost faster for the HS-kernel with increasing temperatures. Our extended results also include further analytical analysis on the thermal decay and the Heisenberg model,

$$H_{\text{XXZ}} = \sum_{j=1}^L (X_j X_{j+1} + Y_j Y_{j+1} + \Delta Z_j Z_{j+1}) := H_a + \Delta H_b, \quad (9)$$

where the H_a and H_b eigenbasis IPR curves are calculated up to $L = 24$, leading to a finite size crossover away from the expected critical point for ground states $\Delta_c = 1$ at $\Delta \approx 2.1$.

4 Conclusion

We introduced the IPR-kernel, a basis-dependent kernel for learning quantum phases of matter. This similarity measure compares IPR profiles in selected physical bases rather than direct state overlap. For the TFIM, the X and Z bases provide a clean testbed: their IPR crossover tracks the critical point and the resulting kernel retains phase-relevant structure even into the finite temperature regime. More broadly, our results support a physics-informed view of quantum kernels: useful feature maps should be tailored to the structure of the problem at hand, rather than made maximally expressive by default. Future work could apply IPR-based features directly to quantum circuits and variational algorithms, in line with recent work on phase detection with variational states, localization-based trainability, and problem-informed Ansätze [21–24]. In this sense, the IPR-kernel points towards problem-aware QML and highlights the potential of quantum-computer-assisted learning for problems rooted in quantum many-body physics and beyond.

References

- [1] Juan Carrasquilla and Roger G. Melko. Machine learning phases of matter. 13(5):431–434. ISSN 1745-2481. doi: 10.1038/nphys4035.
- [2] Anna Dawid, Julian Arnold, Borja Requena, Alexander Gresch, Marcin Płodzień, Kaelan Donatella, Kim A. Nicoli, Paolo Stornati, Rouven Koch, Miriam Büttner, Robert Okuła, Gorka Muñoz-Gil, Rodrigo A. Vargas-Hernández, Alba Cervera-Lierta, Juan Carrasquilla, Vedran Dunjko, Marylou Gabrié, Patrick Huembeli, Evert van Nieuwenburg, Filippo Vicentini, Lei Wang, Sebastian J. Wetzel, Giuseppe Carleo, Eliška Greplová, Roman Krems, Florian Marquardt, Michał Tomza, Maciej Lewenstein, and Alexandre Dauphin. *Modern Applications of Machine Learning in Quantum Sciences*. doi: 10.1017/9781009504942. URL <http://arxiv.org/abs/2204.04198>.
- [3] Hsin-Yuan Huang, Michael Broughton, Masoud Mohseni, Ryan Babbush, Sergio Boixo, Hartmut Neven, and Jarrod R. McClean. Power of data in quantum machine learning. *Nature Communications*, 12:2631, 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-22539-9. URL <https://www.nature.com/articles/s41467-021-22539-9>.
- [4] Yunchao Liu, Srinivasan Arunachalam, and Kristan Temme. A rigorous and robust quantum speed-up in supervised machine learning. *Nature Physics*, 17:1013–1017, 2021. ISSN 1745-2481. doi: 10.1038/s41567-021-01287-z. URL <https://www.nature.com/articles/s41567-021-01287-z>.
- [5] Jonas M. Kübler, Simon Buchholz, and B. Scholkopf. The Inductive Bias of Quantum Kernels. In *Neural Information Processing Systems*, 2021. URL <https://www.semanticscholar.org/paper/The-Inductive-Bias-of-Quantum-Kernels-K%C3%BCbler-Buchholz/0197d510416b1bf8ea88561724b4c3487dd0e8c9>.
- [6] Thomas Hubregtsen, David Wierichs, Elies Gil-Fuster, Peter-Jan H. S. Derks, Paul K. Faehrmann, and Johannes Jakob Meyer. Training quantum embedding kernels on near-term quantum computers. *Physical Review A*, 106:042431, 2022. doi: 10.1103/PhysRevA.106.042431. URL <https://link.aps.org/doi/10.1103/PhysRevA.106.042431>.
- [7] Abdulkadir Canatar, Evan Peters, Cengiz Pehlevan, Stefan M. Wild, and Ruslan Shaydulin. Bandwidth Enables Generalization in Quantum Kernel Models, 2023. URL <http://arxiv.org/abs/2206.06686>.
- [8] Elies Gil-Fuster, Jens Eisert, and Vedran Dunjko. On the expressivity of embedding quantum kernels. *Machine Learning: Science and Technology*, 5:025003, 2024. ISSN 2632-2153. doi: 10.1088/2632-2153/ad2f51. URL <https://doi.org/10.1088/2632-2153/ad2f51>.
- [9] Teresa Sancho-Lorente, Juan Román-Roche, and David Zueco. Quantum kernels to learn the phases of quantum matter. *Physical Review A*, 105:042432, 2022. ISSN 2469-9926, 2469-9934. doi: 10.1103/PhysRevA.105.042432. URL <https://link.aps.org/doi/10.1103/PhysRevA.105.042432>.
- [10] Yusen Wu, Bujiao Wu, Jingbo Wang, and Xiao Yuan. Quantum Phase Recognition via Quantum Kernel Methods. *Quantum*, 7:981, 2023. ISSN 2521-327X. doi: 10.22331/q-2023-04-17-981. URL <http://arxiv.org/abs/2111.07553>.
- [11] Vojtech Havlicek, Antonio D. Córcoles, Kristan Temme, Aram W. Harrow, Abhinav Kandala, Jerry M. Chow, and Jay M. Gambetta. Supervised learning with quantum enhanced feature spaces. *Nature*, 567:209–212, 2019. ISSN 0028-0836, 1476-4687. doi: 10.1038/s41586-019-0980-2. URL <http://arxiv.org/abs/1804.11326>.
- [12] Maria Schuld. Supervised quantum machine learning models are kernel methods, 2021. URL <http://arxiv.org/abs/2101.11020>.
- [13] Adrián Pérez-Salinas, Patrick Emonts, Jordi Tura, and Vedran Dunjko. Multiple-basis representation of quantum states, 2025. URL <http://arxiv.org/abs/2411.03110>.
- [14] Ignacio Baena, Pedro Pérez-Fernández, Manuela Rodríguez-Gallardo, and José Miguel Arias. Entropies and IPR as Markers for a Phase Transition in a Two-Level Model for Atom-Diatomic Molecule Coexistence. *Entropy*, 24:113, 2022. ISSN 1099-4300. doi:

- 10.3390/e24010113. URL <http://arxiv.org/abs/2201.04571>.
- [15] Philipp Frey, David Mikhail, Stephan Rachel, and Lucas Hackl. Probing Hilbert space fragmentation and the block inverse participation ratio. *Physical Review B*, 109:064302, 2024. doi: 10.1103/PhysRevB.109.064302. URL <https://link.aps.org/doi/10.1103/PhysRevB.109.064302>.
 - [16] Wen-Han Kao and Natalia B. Perkins. Disorder upon disorder: Localization effects in the Kitaev spin liquid. *Annals of Physics*, 435:168506, 2021. ISSN 00034916. doi: 10.1016/j.aop.2021.168506. URL <http://arxiv.org/abs/2102.05689>.
 - [17] M. Rossignolo and L. Dell’Anna. Localization transitions and mobility edges in coupled Aubry-André chains. *Physical Review B*, 99:054211, 2019. doi: 10.1103/PhysRevB.99.054211. URL <https://link.aps.org/doi/10.1103/PhysRevB.99.054211>.
 - [18] Yingjian Liu, Piotr Sierant, Paolo Stornati, Maciej Lewenstein, and Marcin Płodzień. Quantum Algorithms for Inverse Participation Ratio Estimation in multi-qubit and multi-qudit systems. *Physical Review A*, 111:052614, 2025. ISSN 2469-9926, 2469-9934. doi: 10.1103/PhysRevA.111.052614. URL <http://arxiv.org/abs/2405.03338>.
 - [19] Supanut Thanasilp, Samson Wang, M. Cerezo, and Zoë Holmes. Exponential concentration in quantum kernel methods. *Nature Communications*, 15:5200, 2024. ISSN 2041-1723. doi: 10.1038/s41467-024-49287-w. URL <https://www.nature.com/articles/s41467-024-49287-w>.
 - [20] Pranav Kairon, Jonas Jäger, and Roman V. Krems. Equivalence between exponential concentration in quantum machine learning kernels and barren plateaus in variational algorithms, 2025. URL <http://arxiv.org/abs/2501.07433>.
 - [21] Chenfeng Cao, Filippo Maria Gambetta, Ashley Montanaro, and Raul A. Santos. Unveiling quantum phase transitions from traps in variational quantum algorithms. *npj Quantum Information*, 11:93, 2025. ISSN 2056-6387. doi: 10.1038/s41534-025-01038-5. URL <https://www.nature.com/articles/s41534-025-01038-5>.
 - [22] Chenfeng Cao, Yeqing Zhou, Swamit Tannu, Nic Shannon, and Robert Joynt. Exploiting many-body localization for scalable variational quantum simulation. *Quantum*, 9:1942, 2025. ISSN 2521-327X. doi: 10.22331/q-2025-12-12-1942. URL <http://arxiv.org/abs/2404.17560>.
 - [23] Kasidit Srimahajariyapong, Supanut Thanasilp, and Thiparat Chotibut. Connecting phases of matter to the flatness of the loss landscape in analog variational quantum algorithms, 2025. URL <http://arxiv.org/abs/2506.13865>.
 - [24] Chae-Yeun Park and Nathan Killoran. Hamiltonian variational ansatz without barren plateaus. *Quantum*, 8:1239, 2024. doi: 10.22331/q-2024-02-01-1239. URL <https://quantum-journal.org/papers/q-2024-02-01-1239/>.