Deep Learning the Functional Renormalization Group

Domenico Di Sante,[1,2,*] Matija Medvidović,[2,3] Alessandro Toschi,[4] Giorgio Sangiovanni,[5] Cesare Franchini,[1,6] Anirvan M. Sengupta,[2,7,8] and Andrew J. Millis[2,3]

[1]*Department of Physics and Astronomy, University of Bologna, 40127 Bologna, Italy*
[2]*Center for Computational Quantum Physics, Flatiron Institute, 162 5th Avenue, New York, New York 10010, USA*
[3]*Department of Physics, Columbia University, New York, New York 10027, USA*
[4]*Institute of Solid State Physics, TU Wien, A-1040 Vienna, Austria*
[5]*Institut für Theoretische Physik und Astrophysik and Würzburg-Dresden Cluster of Excellence ct.qmat, Universität Würzburg, 97074 Würzburg, Germany*
[6]*University of Vienna, Faculty of Physics and Center for Computational Materials Science, A-1090 Vienna, Austria*
[7]*Center for Computational Mathematics, Flatiron Institute, 162 5th Avenue, New York, New York 10010, USA*
[8]*Department of Physics and Astronomy, Rutgers University, Piscataway, New Jersey 08854, USA*

(Received 18 March 2022; revised 18 July 2022; accepted 26 August 2022; published 21 September 2022)

We perform a data-driven dimensionality reduction of the scale-dependent four-point vertex function characterizing the functional renormalization group (FRG) flow for the widely studied two-dimensional t-t' Hubbard model on the square lattice. We demonstrate that a deep learning architecture based on a neural ordinary differential equation solver in a low-dimensional latent space efficiently learns the FRG dynamics that delineates the various magnetic and d-wave superconducting regimes of the Hubbard model. We further present a dynamic mode decomposition analysis that confirms that a small number of modes are indeed sufficient to capture the FRG dynamics. Our Letter demonstrates the possibility of using artificial intelligence to extract compact representations of the four-point vertex functions for correlated electrons, a goal of utmost importance for the success of cutting-edge quantum field theoretical methods for tackling the many-electron problem.

DOI: 10.1103/PhysRevLett.129.136402

Introduction.—Interacting electron systems exhibit a rich variety of distinct phenomena at different energy and temperature scales. Upon lowering these scales, new effective degrees of freedom and collective behaviors emerge, typically including competing spin, charge, and pairing fluctuations. The difficulties inherent in treating these competing, scale-dependent phenomena on an equal footing represent one of the major obstacles to the numerical solution of theoretical models.

The renormalization group (RG) provides a powerful approach to these problems [1–5]. The characteristic of the RG of keeping only the relevant information, as a scale parameter is reduced, makes it a valuable tool to deal with interacting fermions. In its *exact* or *functional* ("FRG") form, the RG is formulated as an exact functional flow equation which, as a function of a continuously decreasing energy scale, provides an effective action description of a microscopic model [6–9].

In quantum condensed matter physics the FRG has been used to study model systems such as the two-dimensional (2D) Hubbard model and its extensions [10–18]. Applications of the FRG to real materials have thus far remained sporadic, requiring considerable numerical effort to incorporate realistic band dispersions, multiorbital characters, and realistic interactions. Successful applications, including the study of superconductivity (SC) and competing phases in iron-based compounds [19–22], cobaltates [23], doped and twisted-bilayer graphene [24–26], buckled Dirac semimetals, doped topological insulators [27,28], and quantum spin liquid phases in frustrated antiferromagnets [29,30], show the potential of this approach, but the underlying computational complexity suggests a simplification of the FRG approach would be desirable.

In the standard RG procedure, say, for a $(4-\epsilon)$-dimensional ϕ^4 field theory with ϵ expansion [1], the effective action of the theory takes a simple form parametrized by a small number of coupling constants and the functional flow equations collapse to a small set of coupled differential equations describing the flow of these coupling constants. In contrast, the common formulation of the fermionic FRG keeps track of the entire frequency and momentum dependence of the interaction vertexes during the flow [9,13]. Thus, the apparent dimensionality space of "couplings" is high, although schemes for finding compressed representations have been investigated [16,17,31–35]. In this Letter we present results from a data-driven approach indicating that this apparent high dimensionality of the vertex function can be, in some cases, illusory.

In the context of high-dimensional data, the advent of machine learning (ML) techniques and data-driven approaches applied to many body quantum physics has triggered enormous interest [36]. ML ideas have been

applied so far to several categories of methods for interacting electron systems, including density functional theory (DFT) [37–40], the Anderson impurity model [41], quantum-embedding and dynamical mean field theory (DMFT) [42–44], and the numerical renormalization group (NRG) [45,46]. Interacting spins models have also been studied [47]. In fact, the ability of neural networks to approximate a very large class of functions promoted the use of deep net architecture as a new numerical tool for solving the quantum many-body problem.

The essential object in the FRG is the vertex function $\mathrm{V}(\mathrm{k}_1, \mathrm{k}_2, \mathrm{k}_3)$, whose description, in principle, requires the computation and storage of a function of three continuous momentum variables. By studying a particular theoretical model, the two-dimensional t-t'-U Hubbard model, believed relevant for cuprates and wide classes of organic conductors, we show that a lower dimensional representation can capture the FRG flow of the apparently high-dimensional vertex functions. To learn this simpler representation, we use a neural network architecture known as neural ordinary differential equations (NODEs) [48]. The latent variable space of the NODEs provides us with insight into the low-dimensional structure of the FRG flow. In order to further investigate the reason for this simplicity, we apply dynamic mode decomposition (DMD) [49,50], a complementary dimensionality reduction scheme that is specifically tailored for dynamical systems. We observe that a small number of modes are able to approximately capture the FRG dynamics of this model. In other words, the key to the success of the method is that the momentum dependence of the vertex can be approximated by combining a small number of patterns characteristic of the different competing many-body phases.

Thus our ML approach to the FRG achieves what reduced order models [51] wish to accomplish. However, we do not apply predetermined ansätze or make simplifying approximations, potentially discarding some relevant information in the vertex function. Instead, we let data guide the choice of the lower-dimensional representation.

The FRG ground states of the Hubbard model.—The microscopic Hamiltonian we consider is

$$H = -t\sum_{\mathrm{NN},s} c^{\dagger}_{i,s}c_{j,s} - t'\sum_{\mathrm{NNN},s} c^{\dagger}_{i,s}c_{j,s} + U\sum_{i} n_{i,\uparrow}n_{i,\downarrow} \quad (1)$$

with hopping amplitudes t and t' between nearest neighbors (NN) and next-nearest neighbors (NNN) on the 2D square lattice, and onsite Coulomb repulsion U. The two-particle properties of this model are investigated through the temperature-flow one-loop FRG scheme [14,15], where the RG flow of $\mathrm{V}^{\Lambda}(\mathrm{k}_1, \mathrm{k}_2, \mathrm{k}_3)$ is

$$\frac{d\mathrm{V}^{\Lambda}}{d\Lambda} = \mathrm{V}^{\Lambda}\circ\mathrm{L}^{\Lambda}\circ\mathrm{V}^{\Lambda} \quad (2)$$

[see Fig. 1(a)], with the RG scale Λ given by the temperature T, and $\circ$ defining integration over the internal

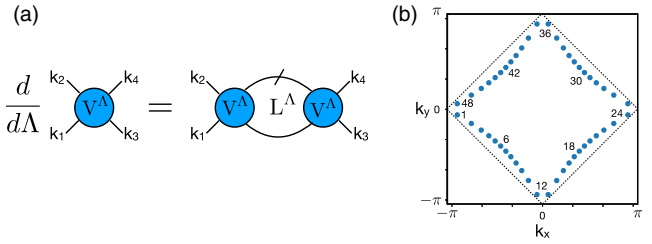

FIG. 1. (a) Diagrammatic representation of the one-loop FRG flow equation for the two-particle vertex function V^{Λ}. (b) Fermi surface (FS) of the $t-t'$ tight-binding model for $t' = -0.25t$. The blue points indicate the 48 momenta used to patch the FS. The black dashed lines are the umklapp surface of perfect nesting at $t' = 0$.

degrees of freedom. Comprehensive reviews of the FRG scheme applied to fermionic problems are found in Refs. [52,53]; Ref. [54] gives an excellent pedagogical introduction. In essence, for spin-rotation invariant systems, the right-hand side of Eq. (2) splits into the sum of three contributions, which describe the particle-particle, direct particle-hole, and crossed particle-hole channels [13] necessary to account on similar footing for the SC and density-wave instabilities. L^{Λ} is a scale-dependent loop kernel that contains information on the single-particle properties of the microscopic model [55].

Neglecting the frequency dependences of the vertex couplings, which have a negative scaling dimension (irrelevant couplings) under the RG flow [56], but keeping a full momentum description in terms of a discrete set of N_k wave vectors on the Fermi surface (FS), Eq. (2) is recasted into a set of N_k^3 coupled ordinary differential equations (ODEs). The solution to this problem, with initial conditions $\mathrm{V}^{\Lambda_0}(\mathrm{k}_1, \mathrm{k}_2, \mathrm{k}_3) = U$ when $\Lambda_0 = 8t$ is the bandwidth, yields the gradual evolution of the two-particle vertex function $\mathrm{V}^{\Lambda}(\mathrm{k}_1, \mathrm{k}_2, \mathrm{k}_3)$ as $\Lambda \to 0$ approaches the FS. For a typical $N_k = 48$ FS discretization, as depicted in Fig. 1(b), the complexity of Eq. (2) already amounts to more than 10^5 coupled ODEs.

When Eq. (2) is numerically solved at varying t' for chemical potential fixed at the van Hove filling ($\mu = 4t'$) and at weak-coupling $U = 3t$, the Hubbard model in Eq. (1) experiences three different regimes [14]: (i) The first regime is close to half-filling, with $t' > -0.2t$ and dominant antiferromagnetic (AFM) scattering processes between FS regions connected by wave vectors $\sim(\pi, \pi)$. These can be seen in Fig. 2(b) (left) as bright features corresponding to repulsive couplings on the line $\mathrm{k}_2 - \mathrm{k}_3 \sim (\pi, \pi)$. (ii) Further decreasing t', the d-wave SC takes over, with the dominant $d_{x^2-y^2}$ symmetry of the pairing scattering, as seen from the sign profile of the diagonal features of Fig. 2(b) (central). (iii) After a quantum critical point for $t' \sim -0.33t$, scattering processes with small momentum transfer $\mathrm{k}_2 - \mathrm{k}_3 \sim (0, 0)$ dominate, see bright features in Fig. 2(b) (right), leading to a change

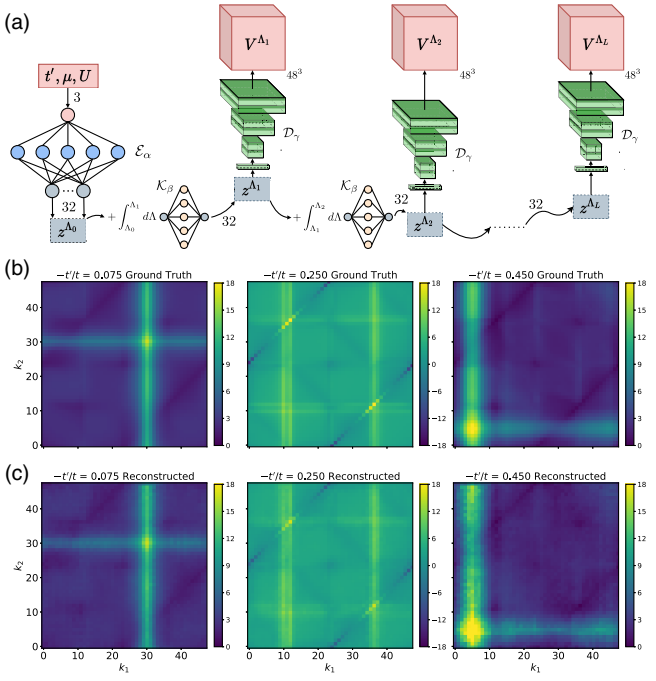

FIG. 2. (a) The deep learning architecture defined by Eqs. (3) and detailed in Ref. [57]. (b) False-color representation of the amplitude of the vertex function $\mathrm{V}^{\Lambda}(\mathrm{k}_1, \mathrm{k}_2, \mathrm{k}_3)$ at a late stage of the renormalization process, for different representative initial conditions. To represent a function of three momenta we fix the first outgoing wave vector k_3 and present the vertex as a function of the two independent incoming vectors. Following Ref. [15], (left) V for $t' = -0.075t$ with k_3 fixed at patch 6; (central) V for $t' = -0.25t$, and k_3 is fixed at patch 25 to emphasize the diagonal features important for the Cooper channel; (right) V for $t' = -0.45t$ and k_3 fixed at patch 6. These choices correspond to AFM, d-wave SC, and FM instabilities, respectively [14]. (c) Same as (b) but for the predicted data $\hat{\mathrm{V}}^{\Lambda}(\mathrm{k}_1, \mathrm{k}_2, \mathrm{k}_3)$. We highlight that these data belong to the test set.

of ground state from d-wave singlet SC to ferromagnetism (FM).

The deep learning FRG: Results and interpretation.—By inspecting the $\mathcal{O}(10^5)$ couplings of the two-particle vertex functions of Fig. 2(b) just before the FRG flow runs to strong coupling and the one-loop approximation breaks down, we recognize that many of them either have remained nearly constant or have become vanishingly small under the RG flow. Only *few* of them have grown positively or negatively (bright features) under the RG evolution [56]. However, as mentioned before, contrary to the standard RG procedure for traditional critical phenomena [3], the fermionic FRG does not discard any coupling in the vertex $\mathrm{V}^{\Lambda}(\mathrm{k}_1, \mathrm{k}_2, \mathrm{k}_3)$ during the flow. Our approach is to find a simpler representation in a data-driven manner, using the power of neural nets to find useful features.

In recent years, there have been many developments in utilizing neural networks for predicting sequence data. These range from the conventional recurrent neural network (RNN), gated RNNs, like long short-term memory (LSTM), and those using the gated recurrent unit (GRU) [63], all the way to an encoder-decoder with attention [64]. Since we are interested in finding latent variables whose dynamics itself is governed by an ODE, the natural candidate is a flexible dimensionality reduction scheme based on the parameterized NODE architecture [48,65]. The method, sketched in Fig. 2(a), focuses on three deep neural networks—the encoder $\mathcal{E}$, the NODE $\mathcal{K}$, and the decoder $\mathcal{D}$. The complete action of our model is defined by

$$\mathbf{z}^{\Lambda_0} = \mathcal{E}_\alpha(t', U, \mu); \quad \frac{d\mathbf{z}^{\Lambda}}{d\Lambda} = \mathcal{K}_\beta(\mathbf{z}^{\Lambda}); \quad \hat{\mathrm{V}}^{\Lambda} = \mathcal{D}_\gamma(\mathbf{z}^{\Lambda}), \quad (3)$$

where α, β, γ are parameter sets corresponding to each neural network (details are found in Ref. [57], and our PyTorch implementation, NeuralFRG, is at Ref. [66]). The ground truth data are generated by solving the FRG problem in Eq. (2) for 35 values of t' in the range $0 \leq -t'/t < 0.5$ (and $U = 3t$), and storing for each t' the whole vertex dynamics, for a total of ∼7000 collected vertices, each with $\mathcal{O}(N_k^3 \sim 10^5)$ elements.

The encoder maps the Hubbard model parameters and the FRG initial condition to a low-dimensional latent representation $\mathbf{z}^{\Lambda_0}$ of drastically smaller dimension than N_k^3. All the results here are obtained with a latent space dimension of 32, but are robust against the use of either smaller or larger values, as we show in Ref. [57]. The NODE then defines a differential equation propagation rule for latent variables in Λ. Finally, at each step of the flow, a decoder network is employed to map the latent representation $\mathbf{z}^{\Lambda}$ to a reconstructed four-point vertex function $\hat{\mathrm{V}}^{\Lambda}(\mathrm{k}_1, \mathrm{k}_2, \mathrm{k}_3)$. We use a modified version of the mean squared error (MSE, Ref. [57]) between $\hat{\mathrm{V}}^{\Lambda}$ and V^{Λ}, in conjunction with a gradient-based optimizer [67]. All three networks, $\mathcal{E}_\alpha$, $\mathcal{K}_\beta$, and $\mathcal{D}_\gamma$, are optimized simultaneously.

We find that the learned dynamics successfully captures the final instability for the entire range of next-nearest neighbor couplings t'. Satisfactory prediction of the qualitative features of the vertex data is achieved in the limit $\Lambda \to 0$, as presented in Fig. 2(c). More interestingly, Fig. 3 shows that, during the FRG dynamics in the latent space, three highly statistically correlated latent space representations $\mathbf{z}$ emerge as a learned feature of the NODE neural network. At $\Lambda = \Lambda_0$, a first classification task is performed by the encoder $\mathcal{E}_\alpha$, which produces highly correlated latent variables according to the value of t' [Fig. 3(a)]. The NODE $\mathcal{K}_\beta$ takes it over to finite RG time $\ln \Lambda_0/\Lambda$, and during the final stages of the FRG evolution in Λ, three markedly correlated areas appear, as shown in Fig. 3(b).

The boundaries of these three regions roughly coincide with the values of t' at which the FRG predicts a change in the leading two-particle instability [14]. It is also interesting to notice that while the AFM and d-wave SC areas show similar normalized $\bar{\mathbf{z}}$ and are thus well aligned in the

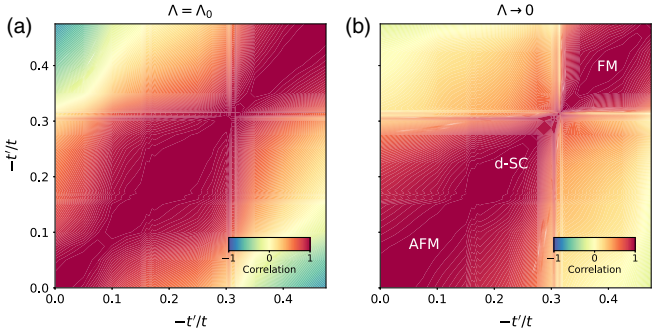

FIG. 3. Correlation matrix of the latent vectors $\mathbf{z}^{\Lambda}$ at different t' values for $\Lambda = \Lambda_0$ (a) and $\Lambda \to 0$ (b), respectively. The correlation matrix is defined as the scalar product kernel $[\bar{\mathbf{z}}(t'_1) \cdot \bar{\mathbf{z}}(t'_2)]$ between normalized latent variables $\bar{\mathbf{z}}$, where t'_1 and t'_2 are any of the $0 \leq -t'/t < 0.5$. Red (blue) features correspond to a high (low) degree of statistical correlation.

latent space (the scalar product kernel $[\bar{\mathbf{z}}(t'_1) \cdot \bar{\mathbf{z}}(t'_2)] \sim 1$), reflecting their common origin in the dominant spin fluctuations, the FM region stands on its own, separated from the other two phases by either a quantum critical point (in the FRG framework [14]) or a first order transition (in the Hartree-Fock treatment [68]). The neural network distinguishes between these three many-body regimes by learning specific low-dimensional hidden representations. This is accomplished by activating three different groups of neurons in $\mathcal{K}_\beta$ as a function of t', as shown in Fig. 4(a). Each instability ground state corresponds indeed to a specific pattern of active neurons. This is manifestly evident when the neuron activation pattern of Fig. 4(a) is contrasted to Fig. 4(b), where we show the dependence on t' of the most negative eigenvalues $w_0^{\mathrm{ch}}(\Lambda)$ of the FRG channel couplings $W^{\Lambda,\mathrm{ch}}(\mathrm{k}_1, \mathrm{k}_2) = \sum_i w_i^{\mathrm{ch}}(\Lambda) f_i^{\mathrm{ch}}(\mathrm{k}_1)^* f_i^{\mathrm{ch}}(\mathrm{k}_2)$ [53,57], with channels ch = AFM, SC, FM, and $f_i^{\mathrm{ch}}(\mathrm{k})$ lattice harmonics transforming as an irreducible representation of the symmetry group of $W^{\Lambda,\mathrm{ch}}(\mathrm{k}_1, \mathrm{k}_2)$. These leading eigenvalues are the ones

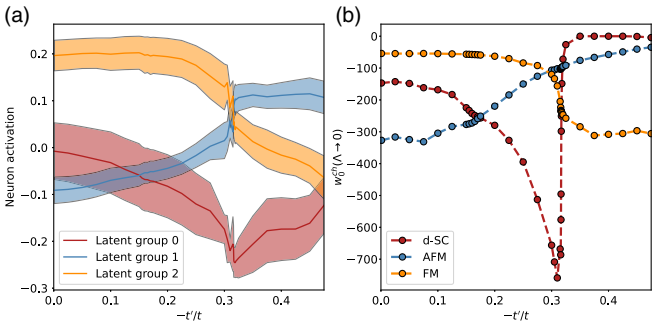

FIG. 4. (a) Three-group K-means clustering [69] of neuron activation in $\mathcal{K}_\beta$ as a function of t' for $\Lambda \to 0$. (b) Evolution of the three channel leading eigenvalues $w_0^{\mathrm{ch}}(\Lambda)$ as a function of t' before the breaking of the one-loop approximation.

associated with the highest ordering temperature T_c for their specific channel [53].

The dynamic mode decomposition (DMD).—The success of the NODE-based neural network in finding a relatively low-dimensional hidden representation prompts us to uncover possible simplicity hidden in the collection of the dynamic trajectories of the vertex functions themselves. To explore this possibility, we apply conventional DMD [49], without any specialized kernel [70]. Given a time series of data, DMD computes a set of modes each of which is associated with an eigenvalue on the complex plane, approximating eigenvalues and eigenvectors of the Koopman operator [50]. Although the mathematical procedure for identifying the DMD modes and eigenvalues is purely linear, the dynamic itself can be nonlinear.

For each individual FRG flow for $0 \leq -t'/t < 0.5$, we collected the four-point vertex functions in the form of a snapshot sequence $\mathcal{V}_0^N = \{\mathrm{V}^{\Lambda_0}, \mathrm{V}^{\Lambda_1}, \ldots, \mathrm{V}^{\Lambda_N}\}$. These snapshots are assumed to be related via a linear mapping that defines a linear dynamical system in the Hilbert space of functions that best approximates the nonlinear dynamics of the data [71]

$$\mathrm{V}^{\Lambda_{i+1}} = \mathcal{A}\mathrm{V}^{\Lambda_i}, \tag{4}$$

with the Koopman operator $\mathcal{A}$ having the prohibitive dimension $N_k^3 \times N_k^3$. Eigenvalues and eigenvectors of $\mathcal{A}$ are referred to as the DMD eigenvalues and DMD modes, respectively. Since a direct eigendecomposition of $\mathcal{A}$ is unfeasible, the problem can be made tractable with the help of a singular value decomposition (SVD) of $\mathcal{V}_0^{N-1}$ [49,57].

One can select a restricted number N^{SVD} of largest singular values of $\mathcal{V}_0^{N-1}$ performing a low-rank truncation, and optimally reconstruct the flow dynamics $\hat{\mathrm{V}}^{\Lambda}(\mathrm{k}_1, \mathrm{k}_2, \mathrm{k}_3)$ only using the leading N^{SVD} DMD modes (see Ref. [72] for details about the optimal reconstruction procedure used here). In Fig. 5(a), for exemplary cases of the three instability regimes, we show the MSE between the reconstructed data $\hat{\mathrm{V}}^{\Lambda}$ and the ground truth data V^{Λ}, which rapidly decays with the number of included DMD modes.

Interestingly, one can also use several different FRG flows in the same snapshot sequence and perform the DMD analysis. This way, the DMD modes inherit features from all included four-point vertex functions. We choose $-t'/t = 0$ (AFM), -0.315 (d-wave SC), -0.475 (FM) in order to span all three instability regimes, and reconstruct the FRG flow even on unseen data. Figure 5(b) shows, for instance, the reconstructed data for $-t'/t = 0.17$ and $N^{\mathrm{SVD}} = 8$ (right panel) as compared to the FRG ground truth (left panel). These results suggest also that the four-point vertex function (in principle a complicated function of three variables) can indeed be very compactly parametrized by a small number of basis functions in each of the relevant two particle regimes. Whether this finding applies more

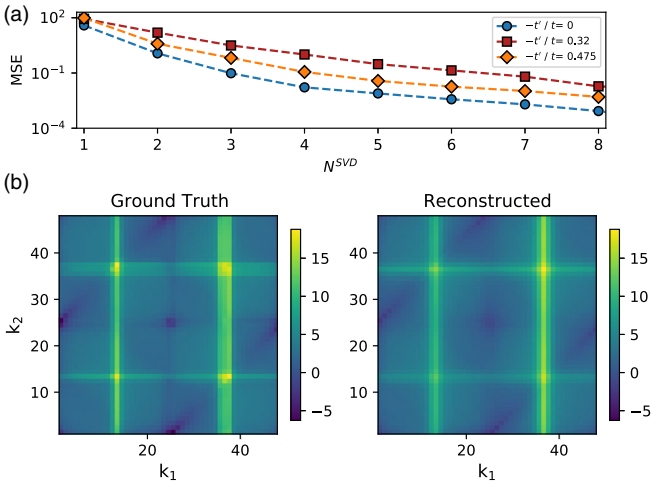

FIG. 5. (a) MSEs from the DMD reconstructed FRG dynamics for $-t'/t = 0$ (AFM), $-t'/t = 0.32$ d-wave SC, and $-t'/t = 0.475$ FM regimes as a function of the number of singular values N^{SVD} included in the SVD scheme. (b) Comparison between unseen ground truth $\mathrm{V}^{\Lambda}(\mathrm{k}_1, \mathrm{k}_2, \mathrm{k}_3)$ and DMD reconstructed data $\hat{\mathrm{V}}^{\Lambda}(\mathrm{k}_1, \mathrm{k}_2, \mathrm{k}_3)$ for $-t'/t = 0.17$ and first outgoing wave vector k_3 fixed at patch 1.

generally to the vertex functions is an important open question.

Outlook.—Our Letter presents an application of artificial intelligence to the FRG, which successfully unveils a dimensionality reduced dynamics for the Hubbard model on a square lattice at specific sets of electron filling. Nonetheless, the relevance of the procedure outlined in our Letter goes beyond the test bed cases considered. In particular, the identification of how to extract and manipulate relevant information encoded in the four-point (or two-particle) vertex functions of many-electron problems, separating it from the nonrelevant ones, represents a goal of utmost importance for the success of several cutting-edge quantum field theoretical methods for quantum materials. The four-point vertex is, in fact, the building block of advanced approaches based on the solution of FRG flow equations [52] or resummation of parquet diagrams [73], including their most recent developments such as the multiloop extension [74,75] of the FRG, the merger of the FRG and DMFT [76], and the diagrammatic extensions [77] of DMFT itself, specifically designed for computing the most challenging nonperturbative regimes.

In addition, the promising outcome of our deep learning-based FRG procedure naturally suggests its extension beyond the van Hove filling illustrated here as well as to more realistic models including nonlocal electronic interactions. It will be also important to explore whether transfer learning [78] could mitigate the burden of training deep nets for similar Hamiltonians on other lattice geometries, and whether a rich set of learned latent variables allows for extracting equations from the data (symbolic regression) [79] and for a resolution increase of the vertex.

The authors are grateful to Ronny Thomale for providing the FORTRAN N-patch FRG code used to generate the ground truth data, and to Daniel Springer and Tobias Müller for stimulating discussions. The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 897276 (BITMAP). This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through Project-ID 258499086-SFB 1170 and through the Würzburg-Dresden Cluster of Excellence on Complexity and Topology in Quantum Matter-ct.qmat Project-ID 390858490-EXC 2147 as well as the Austrian Science Fund (FWF) through the project I 2794-N35. M. M. acknowledges support from the CCQ graduate fellowship in computational quantum physics. The Flatiron Institute is a division of the Simons Foundation.

*domenico.disante@unibo.it

[1] K. G. Wilson and J. G. Kogut, Phys. Rep. **12**, 75 (1974).
[2] K. G. Wilson, Rev. Mod. Phys. **47**, 773 (1975).
[3] M. E. Fisher, *Scaling, Universality and Renormalization Group Theory* (Springer, Berlin, 1983).
[4] R. Shankar, Rev. Mod. Phys. **66**, 129 (1994).
[5] J. Polchinski, arXiv:hep-th/9210046.
[6] C. Wetterich, Phys. Lett. B **301**, 90 (1993).
[7] T. R. Morris, Int. J. Mod. Phys. A **09**, 2411 (1994).
[8] M. Salmhofer and C. Honerkamp, Prog. Theor. Phys. **105**, 1 (2001).
[9] P. Kopietz and T. Busche, Phys. Rev. B **64**, 155101 (2001).
[10] D. Zanchi and H. J. Schulz, Phys. Rev. B **61**, 13609 (2000).
[11] C. J. Halboth and W. Metzner, Phys. Rev. B **61**, 7364 (2000).
[12] C. J. Halboth and W. Metzner, Phys. Rev. Lett. **85**, 5162 (2000).
[13] C. Honerkamp, M. Salmhofer, N. Furukawa, and T. M. Rice, Phys. Rev. B **63**, 035109 (2001).
[14] C. Honerkamp and M. Salmhofer, Phys. Rev. Lett. **87**, 187004 (2001).
[15] C. Honerkamp and M. Salmhofer, Phys. Rev. B **64**, 184516 (2001).
[16] C. Hille, F. B. Kugler, C. J. Eckhardt, Y.-Y. He, A. Kauch, C. Honerkamp, A. Toschi, and S. Andergassen, Phys. Rev. Research **2**, 033372 (2020).
[17] C. Hille, D. Rohe, C. Honerkamp, and S. Andergassen, Phys. Rev. Research **2**, 033068 (2020).
[18] D. Vilardi, P. M. Bonetti, and W. Metzner, Phys. Rev. B **102**, 245128 (2020).
[19] F. Wang, H. Zhai, Y. Ran, A. Vishwanath, and D.-H. Lee, Phys. Rev. Lett. **102**, 047005 (2009).
[20] R. Thomale, C. Platt, J. Hu, C. Honerkamp, and B. A. Bernevig, Phys. Rev. B **80**, 180505(R) (2009).
[21] R. Thomale, C. Platt, W. Hanke, and B. A. Bernevig, Phys. Rev. Lett. **106**, 187003 (2011).
[22] R. Thomale, C. Platt, W. Hanke, J. Hu, and B. A. Bernevig, Phys. Rev. Lett. **107**, 117001 (2011).

[23] M. L. Kiesel, C. Platt, W. Hanke, and R. Thomale, Phys. Rev. Lett. **111**, 097001 (2013).
[24] C. Honerkamp, Phys. Rev. Lett. **100**, 146404 (2008).
[25] M. L. Kiesel, C. Platt, W. Hanke, D. A. Abanin, and R. Thomale, Phys. Rev. B **86**, 020507(R) (2012).
[26] D. M. Kennes, J. Lischner, and C. Karrasch, Phys. Rev. B **98**, 241407(R) (2018).
[27] D. Di Sante, X. Wu, M. Fink, W. Hanke, and R. Thomale, Phys. Rev. B **99**, 201106(R) (2019).
[28] X. Wu, M. Fink, W. Hanke, R. Thomale, and D. Di Sante, Phys. Rev. B **100**, 041117(R) (2019).
[29] S. Chillal, Y. Iqbal, H. O. Jeschke, J. A. Rodriguez-Rivera, R. Bewley, P. Manuel, D. Khalyavin, P. Steffens, R. Thomale, A. T. M. N. Islam *et al.*, Nat. Commun. **11**, 2348 (2020).
[30] F. L. Buessen, M. Hering, J. Reuther, and S. Trebst, Phys. Rev. Lett. **120**, 057201 (2018).
[31] N. Wentzell, G. Li, A. Tagliavini, C. Taranto, G. Rohringer, K. Held, A. Toschi, and S. Andergassen, Phys. Rev. B **102**, 085106 (2020).
[32] J. Kaye, K. Chen, and O. Parcollet, Phys. Rev. B **105**, 235115 (2022).
[33] C. Husemann and M. Salmhofer, Phys. Rev. B **79**, 195125 (2009).
[34] W.-S. Wang, Y.-Y. Xiang, Q.-H. Wang, F. Wang, F. Yang, and D.-H. Lee, Phys. Rev. B **85**, 035414 (2012).
[35] J. Lichtenstein, D. Sánchez de la Peña, D. Rohe, E. Di Napoli, C. Honerkamp, and S. Maier, Comput. Phys. Commun. **213**, 100 (2017).
[36] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Rev. Mod. Phys. **91**, 045002 (2019).
[37] J. C. Snyder, M. Rupp, K. Hansen, K.-R. Müller, and K. Burke, Phys. Rev. Lett. **108**, 253002 (2012).
[38] F. Brockherde, L. Vogt, L. Li, M. E. Tuckerman, K. Burke, and K.-R. Müller, Nat. Commun. **8**, 872 (2017).
[39] J. R. Moreno, G. Carleo, and A. Georges, Phys. Rev. Lett. **125**, 076402 (2020).
[40] J. Kirkpatrick, B. McMorrow, D. H. P. Turban, A. L. Gaunt, J. S. Spencer, A. G. D. G. Matthews, A. Obika, L. Thiry, M. Fortunato, D. Pfau *et al.*, Science **374**, 1385 (2021).
[41] L. F. Arsenault, A. Lopez-Bezanilla, O. A. von Lilienfeld, and A. J. Millis, Phys. Rev. B **90**, 155136 (2014).
[42] J. Rogers, T.-H. Lee, S. Pakdel, W. Xu, V. Dobrosavljević, Y.-X. Yao, O. Christiansen, and N. Lanatà, Phys. Rev. Research **3**, 013101 (2021).
[43] T. Song and H. Lee, Phys. Rev. B **100**, 045153 (2019).
[44] E. Sheridan, C. Rhodes, F. Jamet, I. Rungger, and C. Weber, Phys. Rev. B **104**, 205120 (2021).
[45] S.-H. Li and L. Wang, Phys. Rev. Lett. **121**, 260601 (2018).
[46] J. B. Rigo and A. K. Mitchell, Phys. Rev. Research **4**, 013227 (2022).
[47] G. Carleo and M. Troyer, Science **355**, 602 (2017).
[48] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud, *Neural Ordinary Differential Equations*, Advances in Neural Information Processing Systems (Curran Associates, Inc., 2018), Vol. 31.
[49] P. J. Schmid, J. Fluid Mech. **656**, 5 (2010).
[50] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, J. Fluid Mech. **641**, 115 (2009).
[51] A. Quarteroni and G. Rozza, eds., *Reduced Order Methods for Modeling and Computational Reduction* (Springer, Cham, 2014).
[52] W. Metzner, M. Salmhofer, C. Honerkamp, V. Meden, and K. Schönhammer, Rev. Mod. Phys. **84**, 299 (2012).
[53] C. Platt, W. Hanke, and R. Thomale, Adv. Phys. **62**, 453 (2013).
[54] P. Kopietz, L. Bartosch, and F. Schütz, *Introduction to the Functional Renormalization Group* (Springer-Verlag, Berlin Heidelberg, 2010).
[55] Commonly, the feedback of the 2-point vertex function $\Sigma^{\Lambda}(k)$ in L^{Λ} is neglected [52,53].
[56] The reader interested in how the contributions to the effective action are classified as relevant, marginal and irrelevant, within the FRG framework, is referred to Refs. [9,52,54], where rescaled versions of the FRG flow equations are derived and couplings are classified.
[57] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.129.136402 for further details and results. It also includes Refs. [58–62].
[58] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga *et al.*, in *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, Vancouver, BC, Canada* (2019), Vol. 32, ISSN 10495258, arXiv:1912.01703.
[59] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith *et al.*, Nature (London) **585**, 357 (2020).
[60] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright *et al.*, Nat. Methods **17**, 261 (2020).
[61] J. D. Hunter, Comput. Sci. Eng. **9**, 90 (2007).
[62] F. Pedregosa, G. Varoquaux, A. GramfortV. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, J. Mach. Learn. Res. **12**, 2825 (2011), http://jmlr.org/papers/v12/pedregosa11a.html.
[63] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, Cambridge, MA, 2016).
[64] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, *Attention is All you Need*, Advances in Neural Information Processing Systems (Curran Associates, Inc., 2017), Vol. 30.
[65] K. Lee and E. J. Parish, Proc. R. Soc. A **477**, 20210162 (2021).
[66] https://github.com/BITMAPdds/NeuralFRG.
[67] D. P. Kingma and J. L. Ba, in *3rd Int. Conf. Learn. Represent. ICLR 2015—Conf. Track Proc.* (International Conference on Learning Representations, ICLR, 2015), arXiv:1412.6980.
[68] H. Q. Lin and J. E. Hirsch, Phys. Rev. B **35**, 3359 (1987).
[69] S. Lloyd, IEEE Trans. Inf. Theory **28**, 129 (1982).
[70] M. O. Williams, C. W. Rowley, and I. G. Kevrekidis, arXiv:1411.2260.
[71] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems* (Society for Industrial and Applied Mathematics, Philadelphia, 2016).
[72] M. R. Jovanović, P. J. Schmid, and J. W. Nichols, Phys. Fluids **26**, 024103 (2014).

[73] N. E. Bickers, in *Theoretical Methods for Strongly Correlated Electrons*, edited by D. Sénéchal, A. Tremblay, and C. Bourbonnais (Springer, New York, 2004), Chap. 6.
[74] F. B. Kugler and J. von Delft, Phys. Rev. Lett. **120**, 057403 (2018).
[75] D. Kiese, T. Müller, Y. Iqbal, R. Thomale, and S. Trebst, Phys. Rev. Research **4**, 023185 (2022).
[76] C. Taranto, S. Andergassen, J. Bauer, K. Held, A. Katanin, W. Metzner, G. Rohringer, and A. Toschi, Phys. Rev. Lett. **112**, 196402 (2014).
[77] G. Rohringer, H. Hafermann, A. Toschi, A. A. Katanin, A. E. Antipov, M. I. Katsnelson, A. I. Lichtenstein, A. N. Rubtsov, and K. Held, Rev. Mod. Phys. **90**, 025003 (2018).
[78] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, in *International Conference on Artificial Neural Networks* (Springer, New York, 2018), pp. 270–279.
[79] S. L. Brunton, J. L. Proctor, and J. N. Kutz, Proc. Natl. Acad. Sci. U.S.A. **113**, 3932 (2016).