Machine learning effective models for quantum systems

Jonas B. Rigo and Andrew K. Mitchell
School of Physics, University College Dublin, Belfield, Dublin 4, Ireland

The construction of good effective models is an essential part of understanding and simulating complex systems in many areas of science. It is a particular challenge for correlated many-body quantum systems displaying emergent physics. Using information theoretic techniques, we propose a model machine learning approach that optimizes an effective model based on an estimation of its partition function. The success of the method is exemplified by application to the single impurity Anderson model and double quantum dots, with new non-perturbative results obtained for the old problem of mapping to effective Kondo models. We also show that the correct effective model is not in general obtained by attempting to match observables to those of its parent Hamiltonian.

The approach to understanding and simulating complex quantum systems can be divided into two groups: ab initio studies in which one tries to account for all microscopic details, or studies of simplified effective models that still capture the essential physical phenomena of interest. A prerequisite for the latter is to construct a good effective model. The question of how to do this systematically, starting from a more complex microscopic system, is an important one for many areas of physics.

Effective models are often defined in a reduced Hilbert space involving only those degrees of freedom relevant to describe the low-temperature physics of a complex microscopic model. They can be derived by perturbatively eliminating degrees of freedom, coarse-graining, or by using renormalization group (RG) methods [1–4]: at low energies, microscopic details only enter through effective interactions and renormalized coupling constants. For the purposes of simulation, and to make realistic contact with experiment, not only the structure of the effective model but also its parameters and emergent energy scales must be determined. By contrast to this bottom-up approach, one can also target specific states or experimental observables of interest, and try to engineer effective models that yield the desired properties [5].

In this Letter, we use information theory and machine learning (ML) methods to find good effective models. We describe two different approaches for quantum many-body systems. The first approach is based on comparing the low-energy eigenspectrum of the effective and microscopic models (see Fig. 1), which gives a deceptively simple and intuitive optimization condition on their partition functions (‘model ML’). The second approach compares the distribution of diagrams in a perturbative expansion, yielding an optimization condition on specific local observables. Perhaps surprisingly, these two approaches yield different results for the optimized effective model, with implications for the Gibbs-Bogoliubov-Feynman (GBF) [6] inequality. By studying specific example systems, we show explicitly that the correct low-energy physics (with the correct emergent energy scales) is obtained by optimization based on the partition function. In general, observables of the correct effective model and the parent Hamiltonian need not agree, due to information monotonicity along RG flow [7]. ML optimization schemes employing cost functions involving physical observables [8] cannot be relied upon to give the correct low-energy physics. Furthermore, different effective models are required to match different observables.

ML stands on the foundations of statistical inference and information theory. The goal is to approximate an unknown probability distribution by optimizing an auxiliary probabilistic model [9, 10]. ML is generally treated as a ‘black box’ method, since the probabilistic models are of high complexity and abstraction with no physical content [11]. However, mapping the input to the desired output in deep learning by processing through successive layers of ever increasing abstraction is conceptually similar to RG [12–17]. Applications of ML in physics include the search for eigenstates of Hamiltonians [18–23], the inverse problem of finding the parent Hamiltonian of a given state [24–26], predicting properties of materials [27, 28], and identifying phases of matter [29–32]. In our model ML approach, the optimized probabilistic model is the desired effective model, and has physical meaning.

![Figure 1. Schematic comparison of bare [left] and effective [right] models: (a) Finite size spectra, which agree up to some high-energy cutoff, $E_{\text{cut}}$. (b) Density of states (Boltzmann weighting $e^{-\beta E}$ as dotted line). (c) Spectrum of the thermal density matrix. (d) Thermodynamics, such as entropy, match at low temperatures $T \ll E_{\text{cut}}$ when $Z_{\text{eff}} = Z_{\text{bare}}$.](image-url)
Partition function condition on effective models.– As illustrated in Fig. 1, the goal is to find an effective model with the same low-energy eigenspectrum, or density of states $\rho(\omega)$, as the bare model. Since the effective model lives in a restricted Hilbert space, its high-energy spectrum is typically more sparse than the bare model. The regime of applicability of the effective model is therefore restricted below some cutoff $E_{\text{cut}}$. At low temperatures $T \ll E_{\text{cut}}$, the thermally-weighted density of states (density matrix spectrum) $q(\omega) = \exp(-\beta \omega)\rho(\omega)$ should therefore agree. This guarantees that the bare and effective models have the same low-temperature thermodynamics, including the same emergent energy scales. Note that if $q_{\text{bare}}(\omega) \sim q_{\text{eff}}(\omega)$ at a given temperature, then the partition functions $Z = \int d\omega q(\omega)$ necessarily match. However, two models with different $q(\omega)$ and therefore different low-energy physics may also have the same $Z$.

In principle, optimizing an effective model could be achieved by minimizing the difference between the bare and effective probability distributions $P(\omega) = q(\omega)/Z$ by minimizing their Kullback-Leibler (KL) divergence [33],

$$D_{\text{KL}} = \int d\omega P_{\text{eff}}(\omega) \log[P_{\text{eff}}(\omega)/P_{\text{bare}}(\omega)]. \quad (1)$$

ML algorithms based in this way on optimizing with respect to the density matrix are referred to as ‘quantum Boltzmann machines’ [34]. The problem is that this rigorous prescription only applies in the eigenbasis of the models, and the gradient descent update required to find the optimal effective model involves taking derivatives of Eq. 1 with respect to tuning parameters. In most cases this is not practicable, and the ML algorithm itself would need to be run on a quantum computer.

Our central result is that this can be avoided if we restrict our attention to effective models that can in principle be derived by a continuous RG transformation from the bare/microscopic model. In particular, the low-energy spectrum of the effective model should remain in one-to-one correspondence with the bare model, with the same quantum numbers, and the symmetries of the bare model are preserved (although the effective model may have larger symmetries). We exclude, for example, a large class of effective models involving a non-interacting quantum gas fine-tuned to trivially reproduce the desired eigenspectrum, or other unphysical models. While an RG-derivable effective model $\hat{H}_{\text{eff}} = \sum_i \theta_i \hat{h}_i$ may have high physical complexity, its parametric complexity $\{\theta_i\}$ is typically modest. A given effective model has correspondingly modest expressibility in terms of describing different physical systems; the structure of an effective model must be appropriate to the physics being described. This is unlike the standard philosophy for Boltzmann machines that employ an unphysical auxiliary energy-based model to represent $P_{\text{bare}}(\omega)$, with high expressibility but also high parametric complexity [9].

Since the RG process can be regarded as a ‘quantum channel’ [7] (a completely positive, trace preserving linear map [35]), the partition function is invariant under RG. An RG-derivable effective model therefore satisfies the condition, $Z_{\text{eff}} = Z_{\text{bare}}$. Optimization can therefore be done directly on the level of the partition functions.

Our model ML does not perform RG: given a suitable structure for the effective model, the method efficiently finds the optimized model parameters by matching partition functions. With loss function $L_J = |\log(Z_{\text{eff}}) - \log(Z_{\text{bare}})|^2$, the gradient descent update for tuning a parameter $\theta_i$ of the effective model is

$$\partial L_J / \partial \theta_i \sim |\log(Z_{\text{eff}}) - \log(Z_{\text{bare}})| \cdot \langle \hat{h}_i \rangle_{\text{eff}}. \quad (2)$$

If the effective model cannot describe the physics of the bare model, such that $Z_{\text{eff}} \neq Z_{\text{bare}}$, the optimization procedure yields coupling constant $\theta_i = 0$.

Finally, note that the partition functions themselves can be estimated by any suitable method at any temperature $T < E_{\text{cut}}$ (the optimization is robust and found to be insensitive to the particular temperature chosen).

Model machine learning for the Anderson model.– As a simple but non-trivial demonstration, we apply the model ML scheme to the Anderson impurity model (AIM) [36],

$$\hat{H}_{\text{A}} = \hat{H}_{\text{bath}} + \sum_{\sigma} \bar{n}_{d\sigma} U \bar{n}_{d\sigma} + V \sum_{\sigma} \left( \bar{c}_{\sigma0\sigma}^\dagger c_{\sigma\sigma}^\dagger + \bar{c}_{\sigma\sigma}^\dagger c_{\sigma0\sigma}^\dagger \right)$$

where $\hat{H}_{\text{bath}} = \sum_{k,\sigma} \epsilon_k c_{k\sigma}^\dagger c_{k\sigma}$, $\bar{n}_{d\sigma} = d_{\sigma}^\dagger d_{\sigma}$, and $V c_{\sigma\sigma} = \sum_k V_k c_{k\sigma}$. For simplicity we consider particle-hole symmetry $\epsilon = -U/2$, and a flat conduction electron density of states in a band of half-width $D = 1$. The well-known low-energy effective model is the Kondo Hamiltonian [36], describing impurity-mediated spin-flip scattering,

$$\hat{H}_{\text{K}} = \hat{H}_{\text{bath}} + J \hat{S}_{d} \cdot \hat{\bar{S}}_{0}, \quad (4)$$

where $\hat{S}_{d}$ is a spin-$1/2$ operator for the impurity, and $\hat{\bar{S}}_{0} = 1/2 \sum_{\sigma,\sigma'} \bar{n}_{\sigma\sigma'} \bar{c}_{\sigma\sigma'}^\dagger c_{\sigma0\sigma'}$ is the spin density of conduction electrons at the impurity. Traditionally, the Kondo model is derived from the Anderson model by means of the Schrieffer-Wolff (SW) transformation [3], which perturbatively eliminates excitations out of the singly-occupied spin manifold of impurity states. SW yields $J_{\text{SW}} = 8V^2 / U$ to second order in the impurity-bath hybridization. However, even SW to infinite-order neglects renormalization effects (retardation) [37, 38].

We use the numerical renormalization group (NRG) method [4, 39–41] to determine the partition functions of the Anderson and Kondo models ($Z_{\text{A}}$ and $Z_{\text{K}}$ respectively) at temperature $T$. Optimization of the Kondo coupling $J$ by minimizing $L_J$ using a bisection method converges exponentially rapidly. NRG results are presented in Fig. 2, comparing the bare AIM (circle points) with effective Kondo models: $J$ determined by model ML (red lines), the SW result (blue lines), and $J$ obtained by observable matching (green lines, discussed shortly).
Figure 2. Model ML derivation of the Kondo model from the Anderson impurity model. (a) Impurity contribution to entropy, \( S(T) \); (b) \( T = 0 \) spectrum of the t-matrix, \( t(\omega) \). AIM results for \( U = 0.5 \) and \( 8V^2/U = 0.25 \) shown as circle points. Lines are for Kondo models with \( J = J_{\text{ML}} \) optimized by model ML (red lines), \( J_{\text{obs}} \) obtained by observable matching (green lines), and \( J_{\text{SW}} \) for Schrieffer-Wolff (blue lines). Insets show the same but for reference AIM with \( 8V^2/U = 0.4 \). (c) Ratio of partition functions \( Z_K/Z_A \) for the model ML optimized Kondo model and the reference AIM; (d) corresponding ratio of spin-spin correlators.

Panel (a) shows the impurity contribution to entropy \( S(T) \), for \( U = 0.5 \) and \( J_{\text{SW}} = 0.25 \) (inset for \( U = 0.5 \) and \( J_{\text{SW}} = 0.4 \)), while panel (b) shows the scattering t-matrix spectrum \( t(\omega) \), at \( T = 0 \) for the same parameters. Panels (c,d) show the ML optimization procedure.

Figs 2(a,b) demonstrate that the model ML method perfectly determines the true coupling \( J \) of the effective Kondo model — even in the case where an incipient local moment is never fully developed (insets). By contrast, the SW result substantially over-estimates the coupling, leading to the wrong low-energy physics and Kondo scales. In more complex models, errors in determining the couplings may lead to incorrect ground state assignment. Deviations between the AIM and Kondo model with \( J = J_{\text{ML}} \) set in only at high temperature scales \( T \sim U \) (impurity charge fluctuations in the AIM cannot be described by the Kondo model [36]).

In the AIM and Kondo models, the density of states \( \rho(\omega) = -\frac{1}{\pi} \sum_{kk'} \text{Im} G_{kk'}(\omega) \), is related to the t-matrix via \( G_{kk'}(\omega) = G_{kk'}^0(\omega) + G_{kk'}^0(\omega) t_{kk'}(\omega) G_{kk'}^{\alpha}(\omega) \), where \( G_{kk'}^0 \) and \( G_{kk'}^{\alpha} \) are the full and free electron Green’s functions [36]. All non-trivial correlations are encoded in the t-matrix spectrum \( t(\omega) \) [41] plotted in Fig. 2(b). Matching the low-energy density of states as per Fig. 1(b) is therefore equivalent to matching the low-energy t-matrix. However, we have shown that this is achieved automatically by satisfying the simpler condition \( Z_K = Z_A \).

Fig. 3(a) shows the evolution of the Kondo coupling \( J \) obtained by model ML for a reference AIM with fixed \( J_{\text{SW}} = 0.3 \), but varying \( U \) (red line). Panel (b) shows the corresponding Kondo temperature \( T_K \) of the ML-optimized Kondo model, compared with the true \( T_K \) of the AIM (circle points). The breakdown of SW is clearly seen: only at very large \( U \) such greater than the conduction bandwidth \( D = 1 \), does SW apply. In particular, note that modest differences in \( J \) translate to dramatic deviations in \( T_K \). We find that, with \( J = J_{\text{ML}} \) the Kondo scale is given very accurately by,

\[
T_K \approx \alpha \frac{\sqrt{\rho_0 J_{\text{ML}}}}{\sqrt{J_{\text{ML}} + \gamma \rho_0 J_{\text{ML}}}} \exp\left(-\frac{1}{\rho_0 J_{\text{ML}} + \gamma \rho_0 J_{\text{ML}}}\right), \tag{5}
\]

where \( \rho_0 \) is the Fermi level free density of states, \( \gamma = \pi^2/4 \), and \( \alpha = \mathcal{O}(1) \). Eq. 5 is consistent with exact results for a pure Kondo model [41, 42]. Eq. 5 applies quite generally to the AIM, provided the correct coupling \( J \) is identified, and generalizes previous results for the AIM valid at large-\( U \) [36]. As a simple rule of thumb, we find:

\[
T_K(J_{\text{ML}}) \approx T_K(J_{\text{SW}}) \times \frac{U}{U + 15\sqrt{\rho_0 J_{\text{SW}}}}. \tag{6}
\]

Inverting Eq. 6 provides an accurate estimate of the true coupling \( J \) of the AIM in terms of the SW result.

Fig. 3(c) shows how the results of model ML depend on temperature. We perform optimization of the Kondo coupling \( J \) by matching partition functions at temperature \( T \) for three reference AIM with the same \( T_K \) but different \( U,V \). We find that \( J_{\text{ML}} \) is robust and essentially constant for all \( T \ll U \), where one expects the Kondo model to apply (in practice, \( J \) is obtained with less than 3% error for \( U/T > 100 \)). This has the important implication that model ML can be performed using estimates of the partition functions at relatively high temperatures, making it amenable to treatment with e.g. quantum Monte Carlo methods [43–45]. Note that for \( T \gtrsim U \), the Kondo model is not a good effective model, and the resulting \( J_{\text{ML}} \) vanishes as per Eq. 2.

A further application of model ML to double quantum dots is given in the Supplementary Material [41].

Optimization using observables. – ML employing heuristic cost functions based on physical observables might seem appealing if the goal is to reproduce observables of the bare model within the simpler description of an effective model. However, as shown below, this is not always possible. Indeed, different effective models are in general required to reproduce different observables.
The correct low-energy effective model (in the sense discussed in this paper) cannot reproduce all observables of the bare model, even at low temperatures, due to information monotonicity along RG flow. The information content of the effective model is necessarily lower than the bare model. Under RG the value of observables may vary with respect to the coupling constants of the effective model, even at low temperatures, due to information monotonicity along RG flow. The information content of the effective model using the observable-matching criterion follows only when the impurity observables match. GBF implies that, when optimizing the effective model with respect to \( \theta \), the corresponding observable \( \langle \hat{h}_i \rangle \) in the effective model is merely bounded by its value in the bare model, not necessarily equal to it.

In the case of mapping AIM to Kondo, we find that the proper effective model (\( J \) determined by model ML) yields \( \langle \hat{S}_d \cdot \hat{S}_0 \rangle_K \leq \langle \hat{S}_d \cdot \hat{S}_0 \rangle_\Lambda \), with the GBF bound satisfied only in the SW limit \( U \rightarrow \infty \), see inset to Fig. 3(a).

To compare with model ML, we implement optimization of the effective Kondo model using the observable-based cost function \( L_J = (\langle \hat{S}_d \cdot \hat{S}_0 \rangle_K - \langle \hat{S}_d \cdot \hat{S}_0 \rangle_\Lambda)^2 \). The green lines in Fig. 2 show the result of minimizing \( L_J \).

The Kondo model with \( J = J_{\text{obs}} \) has the same impurity-bath spin correlation as the reference AIM, but does not yield the correct low-energy physics or Kondo scale (panels a,b). Panels (c,d) show that \( \langle \hat{S}_d \cdot \hat{S}_0 \rangle_K = \langle \hat{S}_d \cdot \hat{S}_0 \rangle_\Lambda \) and \( Z_K = Z_\Lambda \) cannot be simultaneously satisfied. Fig. 3(a) shows how \( J_{\text{obs}} \) varies with \( U \) for fixed \( J_{SW} \). Only for \( U \rightarrow \infty \) in the reference AIM does \( J_{\text{obs}} \rightarrow J_{ML} \) (in which case \( J_{SW} \) is the correct result). For \( U < 1 \), \( J_{\text{obs}} \) is a poor approximation to the true \( J \), high-lited by the differences in Kondo scales in panel (b).

**ML based on diagrammatic expansion.** To gain further insight, we formulate an alternative ML approach using a classical probability distribution extracted from a diagrammatic expansion of the partition function, serving as a proxy for the Hamiltonian. We focus on generalized quantum-impurity problems \( \hat{H} = \hat{H}_{\text{imp}} + \hat{H}_{\text{bath}} + \hat{H}_{\text{hyb}} \), of which AIM and Kondo are the simplest examples. The partition function may be written in terms of the continuous-time hybridization expansion [46–48], viz:

\[
Z = Z_{\text{bath}} \int dx \, w(x), \tag{7a}
\]

\[
w(x) = \det(\Delta(x)) \Lambda(x) \exp[-\beta(\hat{H}_{\text{imp}}) x], \tag{7b}
\]

where \( w(x) \) is the weight of a distinct Feynman diagram labelled by the tuple \( x \), \( \Lambda(x) \) is the antiperiodic hybridization matrix, \( \Lambda(x) \) is a product of Dyson orbitals, and \( (\hat{H}_{\text{imp}}) x \) is the expectation value of the impurity Hamiltonian on diagram \( x \).

Observables are evaluated as \( \langle \mathcal{O} \rangle = \int dx \, w(x) \langle \mathcal{O} \rangle_x \). Further details and discussion are provided in the Supplementary Material [41].

For \( w(x) > 0 \), one may define a classical probability distribution of diagrams, \( P(x) = (Z_{\text{bath}}/Z) w(x) \). As with the classical Boltzmann machine, \( P(x) \) is in the form of an energy-based model, with the weights here distributed according to the impurity Hamiltonian. The distribution of diagrams for bare and effective models can be compared using the KL divergence, Eq. 1. The parameters \( \theta_i \) of the effective impurity Hamiltonian can be optimized by minimizing the KL divergence. The ML gradient descent update is \( \nabla_\theta D_{KL}[P_{\text{eff}} : P_{\text{bare}}] = \beta \langle \nabla_\theta \hat{H}_{\text{eff}} \rangle_{\text{bare}} - \beta \langle \nabla_\theta \hat{H}_{\text{eff}} \rangle_{\text{eff}}, \) provided \( \nabla_\theta \Lambda(x) = 0 \). This is a strictly convex optimization problem, with the minimum found when the impurity observables match.

This rigorous derivation highlights several important features. The observable-matching criterion follows only if there is no ‘sign problem’, \( w(x) > 0 \), such that \( P(x) \) can be regarded as a proper probability distribution, and when the impurity eigenbasis does not depend on the parameters being tuned, \( \nabla_\theta \Lambda(x) = 0 \). Even with these conditions fulfilled, observable matching merely minimizes the KL divergence – it does not imply that \( P_{\text{eff}}(x) = P_{\text{bare}}(x) \) or that \( Z_{\text{eff}} = Z_{\text{bare}} \).
Conclusion and applications.– Finding a good effective low-energy model for a given bare Hamiltonian is a challenging and subtle problem. In this Letter we approach the problem using ML techniques, showing that optimization on the level of the partition function yields the correct low-energy physics for RG-derivable effective models. The success and efficiency of our model ML method is demonstrated for quantum impurity problems. We show that local observables of the correct effective model do not necessarily match those of the bare model.

The model ML framework we introduce is general; applications include deriving effective models for complex molecular junctions [49], and solving inverse problems for rational design. Model ML may also be adapted to find the effective equilibrium problem for non-equilibrium systems [50], or to find simplified/coarse-grained effective descriptions within multi-orbital/cluster dynamical systems [53, 54], or to find simplified/coarse-grained effective models. The success and efficiency of our model ML approach the problem using ML techniques, showing that

Acknowledgments.– We thank Sudeshna Sen for useful discussions, and acknowledge funding from the Irish Research Council Laureate Awards 2017/2018 through grant IRCLA/2017/169.

[1] J. L. Cardy, Scaling and renormalization in statistical physics (1996).
[2] K. G. Wilson, Phys. Rev. B 4, 3174 (1971).
[3] J. R. Schrieffer and P. A. Wolff, Phys. Rev. 149, 491 (1966).
[4] K. G. Wilson, Rev. Mod. Phys. 47, 773 (1975); R. Bulla, T. A. Costi, and T. Pruschke, Rev. Mod. Phys. 80, 395 (2008).
[5] H. Fujita, Y. O. Nakagawa, S. Sugiuira, and M. Oshikawa, Phys. Rev. B 97, 075114 (2018).
[6] R. Feynman, Statistical Mechanics: A Set Of Lectures, Advanced Books Classics (Avalon Publishing, 1998).
[7] C. Bény and T. J. Osborne, Phys. Rev. A 92, 022330 (2015).
[8] K. Pakrouski, arXiv:1907.05898 (2019).
[9] I. Goodfellow, Y. Bengio, and A. Courville, Deep learning (MIT Press, 2017).
[10] C. M. Bishop, Pattern Recognition and Machine Learning (Information Science and Statistics) (Springer-Verlag, Berlin, Heidelberg, 2006).
[11] F. Doshi-Velez and B. Kim, arXiv:1702.08608 (2017).
[12] E. de Mello Koch, R. de Mello Koch, and L. Cheng, arXiv:1906.05212 (2017).
[13] P. Mehta and D. J. Schwab, arXiv:1410.3831 (2014).
[14] S.-H. Li and L. Wang, Phys. Rev. Lett. 121, 260601 (2018).
[15] M. Koch-Janusz and Z. Ringel, Nature Physics 14, 578 (2018).
[16] C. Bény and T. J. Osborne, New Journal of Physics 17, 083005 (2015).
[17] S. Apenko, Physica A: Statistical Mechanics and its Applications 391, 62 (2012).
[18] G. Carleo and M. Troyer, Science 355, 602 (2017).
[19] G. Carleo, K. Choo, D. Hofmann, J. E. T. Smith, T. Westerhout, F. Alet, E. J. Davis, S. Efthymiou, I. Glasser, S.-H. Lin, M. Mauri, G. Mazzola, C. B. Mendl, E. van Nieuwenburg, O. O’Reilly, H. Theseinuaat, G. Torlai, and A. Wietek, arXiv:1904.00031 (2019).
[20] I. Glasser, N. Pancotti, M. August, I. D. Rodriguez, and J. I. Cirac, Phys. Rev. X 8, 011006 (2018).
[21] K. Choo, G. Carleo, N. Regnault, and T. Neupert, arXiv:1807.03325 (2018).
[22] Z. Cai and J. Liu, arXiv:1704.05148 (2017).
[23] G. Torlai and R. G. Melko, arXiv:1905.04312 (2019).
[24] A. Valenti, E. van Nieuwenburg, S. Huber, and E. Grempel, arXiv:1907.02540 (2017).
[25] H. Zheng, K. T. Changlani, Hitesh J.and Williams, W. Busemeyer, and B. L. K., arXiv:1712.00477 (2017).
[26] E. Bairey, I. Arad, and N. H. Linde, Phys. Rev. Lett. 122, 020504 (2019).
[27] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, Nature 559, 547 (2018).
[28] R. Lot, F. Pellegrini, Y. Shaidu, and E. Kucukbenli, arXiv:1906.08534 (2019).
[29] J. Carrasquilla and R. G. Melko, Nature Physics 13, 431 EP (2017).
[30] Y.-T. Hsu, X. Li, D.-L. Deng, and S. Das Sarma, Phys. Rev. Lett. 121, 245701 (2018).
[31] Y. Zhang, A. Mesaros, K. Fujita, S. D. Edkins, M. H. Hamidian, K. Ch'ung, H. Eisaki, S. Uchida, J. C. S. Davis, E. Khatami, and E.-A. Kim, Nature 570, 484 (2019).
[32] W. Rzadkowski, N. Defenu, S. Chiaichiera, A. Trombettoni, and G. Bighin, arXiv:1907.05417 (2019).
[33] S. Kullback and R. A. Leibler, Ann. Math. Statist. 22, 1951 (1951).
[34] M. H. Amin, E. Andriyash, J. Rolfe, B. Kulchitskyy, and R. Melko, Phys. Rev. X 8, 021050 (2018).
[35] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Phys. Rev. Lett. 78, 2275 (1997).
[36] A. C. Hewson, The Kondo Problem to Heavy Fermions, Cambridge Studies in Magnetism (Cambridge University Press, 1993).
[37] S. K. Kehrein and A. Mielke, Annals of Physics 252, 1 (1996).
[38] R. Chan and M. Gulacsi, Philosophical Magazine 84, 1265 (2004).
[39] A. Weichselbaum and J. von Delft, Phys. Rev. Lett. 99, 076402 (2007).
[40] F. B. Anders and A. Schiller, Phys. Rev. Lett. 95, 196801 (2005).
[41] See Supplementary Material, which contains references [53, 54].
[42] N. Andrei, K. Furuya, and J. H. Lowenstein, Rev. Mod. Phys. 55, 331 (1983).
[43] C. H. Bennett, Journal of Computational Physics 22, 245 (1976).
[44] W. von der Linden, Physics Reports 220, 53 (1992).
[45] A. N. Rubtsov, V. V. Savkin, and A. I. Lichtenstein, Phys. Rev. B 72, 035122 (2005).
[46] F. D. M. Haldane, Journal of Physics C: Solid State Physics 11, 5015 (1978).
[47] E. Guill, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner, Rev. Mod. Phys. 83, 349 (2011).
[48] K. Haule, Phys. Rev. B 75, 155113 (2007).
[49] A. K. Mitchell, K. G. L. Pedersen, P. Hedegård, and J. Paaske, Nature Communications 8, 15210 EP (2017).
article.
[50] A. Schiller and S. Hershfield, Phys. Rev. B 62, R16271 (2000).
[51] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti, Rev. Mod. Phys. 78, 865 (2006).
[52] T. Maier, M. Jarrell, T. Pruschke, and M. H. Hettler, Rev. Mod. Phys. 77, 1027 (2005).
[53] R. Žitko and J. Bonča, Phys. Rev. B 74, 045312 (2006).
[54] A. K. Mitchell, M. R. Galpin, S. Wilson-Fletcher, D. E. Logan, and R. Bulla, Phys. Rev. B 89, 121105 (2014); K. M. Stadler, A. K. Mitchell, J. von Delft, and A. Weichselbaum, Phys. Rev. B 93, 235101 (2016).
DOUBLE QUANTUM DOT

As a further demonstration of the model ML method, we now consider the parallel double quantum dot (DQD) model illustrated in Fig. S1. The DQD Hamiltonian reads,

\[ H_{\text{DQD}} = H_{\text{bath}} + \sum_{\alpha=1,2} \left( H_{\text{imp}}^\alpha + H_{\text{hyb}}^\alpha \right), \]  

\[ H_{\text{imp}}^\alpha = \sum_\sigma c_{\alpha\sigma} \hat{n}_\alpha + U \hat{n}_\alpha \hat{n}_\alpha + \hat{\sigma}_\alpha \]  

\[ H_{\text{hyb}}^\alpha = V \sum_\sigma (d_{\alpha\sigma}^\dagger c_{0\sigma} + c_{0\sigma} d_{\alpha\sigma}), \]  

where \( \hat{n}_\alpha = d_{\alpha\sigma}^\dagger d_{\alpha\sigma} \) is the number operator for dot \( \alpha = 1, 2 \). The physics of the DQD is much richer than that of the single dot case, due to the interplay between Kondo physics and an emergent RKKY interaction between the dots, mediated by the conduction electrons [1]. No simple perturbative approach to this problem exists, making the DQD a more stringent test of the model ML methodology.

On the scale \( T \sim U \), charge fluctuations on the dots are frozen out, and the DQD can be understood as two free spins-\( \frac{1}{2} \). On a lower scale \( T \sim J_{\text{RKKY}} \), the emergent ferromagnetic RKKY interaction between the dots favours a low-energy spin-1 DQD configuration. Finally, below the Kondo scale \( T_K \), the collective DQD spin-1 state is Kondo underscreened to leave a residual spin-\( \frac{1}{2} \) local moment and a singular Fermi liquid. The low-energy physics is therefore described by an effective spin-1 Kondo model [1] (Eq. 4 but with a spin-1 impurity operator). Applying model ML, we determine the effective \( J \) of this spin-1 Kondo model, which is seen in Fig. S1 to perfectly reproduce the correct low-temperature behavior of the entropy, including the correct Kondo scale \( T_K \). The effective spin-1 Kondo model applies for \( T \ll J_{\text{RKKY}} \sim 10^{-3} \), an emergent scale which can be read off from Fig. S1 as the temperature at which the entropy crosses over from \( S = \ln(4) \) (two free spins-\( \frac{1}{2} \)) to \( S = \ln(3) \) (free spin-1).

The inset shows the evolution of \( J_{\text{ML}} \) as a function of interaction strength \( U \) for fixed \( J_{\text{SW}} = 8V^2/U = 0.2 \). We find that Eqs. 5 and 6 again hold (but with \( \gamma = \pi^2/2 \)).

Finally, we note that a DQD in the serial configuration, rather than in parallel as shown above, has a singlet rather than doublet ground state. Since the spin-1 Kondo model is no longer a good effective model, the model ML method yields \( J_{\text{ML}} = 0 \), signaling that a different effective model should be chosen.

Figure S1. Impurity contribution to the entropy, \( S(T) \), as a function of temperature, \( T \), for the parallel double quantum dot model (Eq. S1) with \( U = 0.5 \) and \( 8V^2/U = 0.2 \) (black line) compared with an effective spin-1 Kondo model with \( J \) determined by model ML (red line). Inset shows the evolution of \( J_{\text{ML}} \) with \( U \) for the same fixed \( 8V^2/U = 0.2 \).

DETAILS OF ML BASED ON DIAGRAMMATIC EXPANSION

For a general Hamiltonian partitioned as \( \hat{H} = \hat{H}_0 + \hat{H}_1 \), one can compute the partition function as an expansion in powers of \( \hat{H}_1 \) by dividing into infinitesimal imaginary time slices [2],

\[ Z = \sum_{n=0}^\infty \frac{(-1)^n}{n!} T^\beta T \int_0^\beta d\tau_1 ... \int_0^\beta d\tau_n \times \text{tr} \left[ e^{-\beta \hat{H}_0} \hat{H}_1(\tau_n) ... \hat{H}_1(\tau_1) \right]. \]  

This expansion is discussed extensively in the context of continuous time quantum Monte Carlo (CT-QMC) [3–5]. In the case of hybridization expansion CT-QMC [5] \( \hat{H}_1 = \hat{H}_{\text{hyb}} \) describes the hybridization between a non-interacting bath \( \hat{H}_{\text{bath}} \) and an interacting quantum impurity \( \hat{H}_{\text{imp}} \). For the following discussion, we therefore consider Hamiltonians of the form

\[ \hat{H}_0 = \hat{H}_{\text{bath}} + \hat{H}_{\text{imp}} \]  

\[ \hat{H}_1 = \hat{H}_{\text{hyb}} = \sum_k \sum_\sigma V_k^\sigma d_{k\sigma}^\dagger c_{k\sigma} + \text{H.c.}, \]  

where we assume that the hybridization tensor is diagonal in the spin quantum number \( \sigma \) [5]. Eq. S2 can be interpreted as sum over all possible diagrams obtained by allowing electrons to hop between the impurity and the bath. Following the approach of Ref. [6], we bring
the impurity operators into the eigenbasis \((E_\alpha, |\alpha\rangle)\) of \(\hat{H}_{\text{imp}}\) and instead of occupation diagrams we obtain diagrams involving impurity eigenstates. This leaves us with a partition function in the form of Eq. 7,

\[
Z = Z_{\text{bath}} \int dx \ w(x) ,
\]

\[
w(x) = \det(\Delta(x)) \Lambda(x) \exp[-\beta\langle \hat{H}_{\text{imp}} \rangle x] .
\]

where \(x = (n, \{k_i, k'_i, \sigma_i, \tau_i, \tau'_{i+1}\})\) denotes a diagram in the eigenbasis [2]. Going from Eq. S2 to Eq. 7 we have integrated out the non-interacting bath and applied Wick’s theorem to obtain the antiperiodic hybridization function

\[
\det \left[ V_{k_i k'_j} V^{\sigma_i \sigma'_j}_{k_i k'_j} \right] = Z_{\text{bath}} \det \left[ V_{k_i k'_j} V^{\sigma_i \sigma'_j}_{k_i k'_j} \exp \left( T \hat{c}_{k_i \sigma_i} \hat{c}^\dagger_{k'_i \sigma'_j} \right) \right] = Z_{\text{bath}} \det(\Delta(x)) .
\]

Bringing Eq. S2 into the eigenbasis has the effect that \(e^{-\beta \hat{H}_{\text{imp}}}\) can be trivially evaluated, with

\[
\beta \langle \hat{H}_{\text{imp}} \rangle x = \sum_{i=1}^n \left[ E_{\alpha_i}(\tau_i - \tau'_i) + E_{\alpha_i}'(\tau'_i - \tau_{i+1}) \right] ,
\]

where periodic boundary conditions \(k+1 \rightarrow 1\) are implied. The transformation requires a projection of impurity operators onto the eigenbasis, viz

\[
\Lambda(x) = \prod_{i=1}^n \langle \alpha_i | d_i^\dagger \sigma_i' \rangle \langle \alpha_i' | d_i \sigma'_i | \alpha_{i+1} \rangle .
\]

The partition function now takes the form of Eq. 7.

From Eq. 7 we extract a classical probability distribution \(P(x) = (Z_{\text{bath}}/Z) w(x)\), provided \(w(x) > 0\). This distribution acts as a proxy for the impurity Hamiltonian; to emphasize this dependency, we introduce the notation \(P(\hat{H}_{\text{bare}}; x)\). This proxy can be used to compare two Hamiltonians. We write \(P(\hat{H}_{\text{bare}}; x) = (Z_{\text{bath}}/Z_{\text{bare}}) w_{\text{bare}}(x)\) for the bare model, and \(P(\hat{H}_{\text{eff}}; x) = (Z_{\text{bath}}/Z_{\text{eff}}) w_{\text{eff}}(x)\) for the effective model. The ‘distance’ between the bare and effective Hamiltonians can then be measured using the Kullback-Leibler divergence [7] of \(P(\hat{H}_{\text{bare}}; x)\) and \(P(\hat{H}_{\text{eff}}; x)\). From Eq. 1 we find,

\[
D_{\text{KL}}[\hat{H}_{\text{bare}} : \hat{H}_{\text{eff}}] = -\log(Z_{\text{bare}}) + \log(Z_{\text{eff}}) + \int dx \ P(\hat{H}_{\text{bare}}; x) \left[ \log(\Lambda_{\text{bare}}(x)) - \log(\Lambda_{\text{eff}}(x)) \right] - \beta \langle \hat{H}_{\text{bare}} \rangle + \beta \langle \hat{H}_{\text{eff}} \rangle ,
\]

where we have used \(\langle \hat{O}_{\text{bare}} \rangle = \int dx \ w_{\text{bare}}(x) \langle \hat{O} \rangle_{\text{bare}}\) and \(\langle \hat{O} \rangle_{\text{eff}} = \int dx \ w_{\text{eff}}(x) \langle \hat{O} \rangle_{\text{eff}}\). In order to minimize this distance we use a gradient descent (GD) method, for which we require the gradient of Eq. 8,

\[
\nabla D_{\text{KL}}[\hat{H}_{\text{bare}} : \hat{H}_{\text{eff}}] = \beta \langle \hat{H}_{\text{bare}} \rangle + \beta \langle \hat{H}_{\text{eff}} \rangle - \beta \langle \hat{H}_{\text{bare}} \rangle + \beta \langle \hat{H}_{\text{eff}} \rangle - \beta \langle \hat{H}_{\text{bare}} \rangle + \beta \langle \hat{H}_{\text{eff}} \rangle .
\]

In particular, if the eigenbasis of \(\hat{H}_{\text{imp}}\) does not depend on the coupling constants \(\{\beta\}\), then \(\nabla_\beta \log(\Lambda_{\text{eff}}(x)) = 0\), and so

\[
\nabla_\beta D_{\text{KL}}[\hat{H}_{\text{bare}} : \hat{H}_{\text{eff}}] = \beta \langle \hat{O} \rangle_{\text{bare}} - \beta \langle \hat{O} \rangle_{\text{eff}} .
\]

Eq. S10 is suitable for GD optimization methods. The optimized effective model is found when the gradient vanishes, \(\nabla_\beta D_{\text{KL}}[\hat{H}_{\text{bare}} : \hat{H}_{\text{eff}}] = 0\), which arises when the observables evaluated in bare and effective models match. This constitutes a rigorous derivation for quantum impurity type problems of an ML optimization scheme for effective models based on observables. However, as we have shown, this prescription does not necessarily yield an effective model with the correct low-energy physics. On the other hand, if an effective model is sought that reproduces the value of a particular observable in the bare model, this ML scheme is effective and efficient due to the guaranteed convexity of the optimization landscape.

**DETAILS OF NRG CALCULATIONS**

In this work, we used Wilson’s NRG method [8], formulated in the complete Anders-Schiller basis [9], to calculate the partition functions of Anderson and Kondo models at finite temperature \(T\). Numerical results for the impurity contribution to entropy were obtained by standard thermodynamic NRG [8], while the dynamical t-matrix and local spin-spin correlator observables were obtained using the full density matrix approach [10]. The spectrum of the t-matrix is defined \(t(\omega) = -\pi \rho_0 \text{Im} \sum_{kk'} t_{kk'}(\omega)\). All parameters are given in units of the conduction electron bandwidth, here set to \(D = 1\). Calculations were performed with NRG discretization parameter \(\Lambda = 2.5\), retaining \(N\) = 3500 states at each iteration, and exploiting \(Q\) and \(S\) quantum numbers. Interleaved NRG (iNRG) and ‘z-trick’ averaging were not required, but could be used with model ML in principle [11]. We note that for the purposes of model ML, the partition functions need not be calculated exactly: within some given level of approximation applied to both bare and effective models, only the difference \(Z_{\text{bare}} - Z_{\text{eff}}\) needs to be estimated to perform the optimization [12]. Although NRG is used here, the ML methods we describe do not depend on the techniques used to estimate the partition functions or
observables.

*Note on definition of Kondo scales.*– All physical observables are universal functions of $T/T_K$ in the scaling limit of the AIM or Kondo models, and so the definition of the Kondo temperature $T_K$ can be chosen somewhat arbitrarily up to an overall constant. In this work, we choose $\alpha = 1.4$ in Eq. 5, so that the conductance half-width-at-half-maximum through an Andersonian dot is precisely at $T = T_K$. As such, this is a natural experimentally-relevant measure. However, for the Anderson / Kondo models, this definition is equivalent to choosing $S_{\text{imp}}(T = T_K) \simeq 0.5$, $T_K \chi_{\text{imp}}(T = T_K) \simeq 0.1$, or $t(\omega = T_K, T = 0) \simeq 0.63$.

[1] R. Žitko and J. Bonča, Phys. Rev. B 74, 045312 (2006).
[2] F. D. M. Haldane, Journal of Physics C: Solid State Physics 11, 5015 (1978).
[3] A. N. Rubtsov, V. V. Savkin, and A. I. Lichtenstein, Phys. Rev. B 72, 035122 (2005).
[4] E. Gull, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner, Rev. Mod. Phys. 83, 349 (2011).
[5] P. Werner and A. J. Millis, Phys. Rev. B 74, 155107 (2006).
[6] K. Haule, Phys. Rev. B 75, 155113 (2007).
[7] S. Kullback and R. A. Leibler, Ann. Math. Statist. 22, 79 (1951).
[8] K. G. Wilson, Rev. Mod. Phys. 47, 773 (1975); R. Bulla, T. A. Costi, and T. Pruschke, Rev. Mod. Phys. 80, 395 (2008).
[9] F. B. Anders and A. Schiller, Phys. Rev. Lett. 95, 196801 (2005).
[10] A. Weichselbaum and J. von Delft, Phys. Rev. Lett. 99, 076402 (2007).
[11] A. K. Mitchell, M. R. Galpin, S. Wilson-Fletcher, D. E. Logan, and R. Bulla, Phys. Rev. B 89, 121105 (2014); K. M. Stadler, A. K. Mitchell, J. von Delft, and A. Weichselbaum, Phys. Rev. B 93, 235101 (2016).
[12] C. H. Bennett, Journal of Computational Physics 22, 245 (1976).