Unsupervised Model Learning for Quantum Impurity Systems

Jonas B. Rigo[1, 2, 3, *] and Andrew K. Mitchell[2, 3, †]

[1] *Forschungszentrum Jülich GmbH, Peter Grünberg Institute, Quantum Control, 52425 Jülich, Germany*
[2] *School of Physics, University College Dublin, Belfield, Dublin 4, Ireland*
[3] *Centre for Quantum Engineering, Science, and Technology, University College Dublin, Dublin 4, Ireland*

Generalized quantum impurity models – which feature a few localized and strongly-correlated degrees of freedom coupled to itinerant conduction electrons – describe diverse physical systems, from magnetic moments in metals to nanoelectronics quantum devices such as quantum dots or single-molecule transistors. Correlated materials can also be understood as self-consistent impurity models through dynamical mean field theory. Accurate simulation of such models is challenging, especially at low temperatures, due to many-body effects from electronic interactions, resulting in strong renormalization. In particular, the interplay between local impurity complexity and Kondo physics is highly nontrivial. A common approach, which we further develop in this work, is to consider instead a simpler effective impurity model that still captures the low-energy physics of interest. The mapping from bare to effective model is typically done perturbatively, but even this can be difficult for complex systems, and the resulting effective model parameters can anyway be quite inaccurate. Here we develop a non-perturbative, unsupervised machine learning approach to systematically obtain low-energy effective impurity-type models, based on the renormalization group framework. The method is shown to be general and flexible, as well as accurate and systematically improvable. We benchmark the method against exact results for the Anderson impurity model, and provide an outlook for more complex models beyond reach of existing methods.

In the field of condensed matter physics, the microscopic Hamiltonian describing a quantum many-body system (the *bare model*) is in many cases well known. However, for interacting systems, the complexity of these models grows quickly with the number of quantum degrees of freedom (for example orbitals or spins), such that a brute force solution becomes analytically and/or numerically intractable for many realistic scenarios of interest. The challenge in many-body theory is therefore not in writing down the bare model, but rather in solving it. However, in many situations it is not necessary to consider the entire configuration space of the system, because only a (relatively) small *active* subspace controls the phenomena of interest [1]. Thus, *effective models* can be devised, that faithfully capture the phenomena of interest, while only keeping the active part of the configuration space. Such effective models have a reduced complexity and increased expressiveness. The challenge in many-body theory can therefore be restated as one of finding the best *solvable* model that describes approximately but accurately the physics of interest.

As an example, it is often more convenient to analyse an effective lattice model in second quantized form than the *ab initio* treatment involving an all-orbital description of the constituent atoms [2, 3]. Methods such as the constrained random-phase approximation [4, 5], coupled cluster downfolding [6], density matrix downfolding [7] or the Pariser-Parr-Pople model for molecules [8] were devised to systematically eliminate inactive degrees of freedom and account for them with renormalized parameters of a reduced model.

These methods have in common that they require some prior knowledge about the bare model to derive the effective model. This situation shares some similarity with that of the *inverse problem* [9] encountered in statistical inference or machine learning [10], that seeks to infer a probabilistic model from observed data. Recently machine learning techniques have been explored to construct effective Hamiltonians from simulated data [11–14] or from experimentally measured data [15–17]. Another field where the inference of Hamiltonians has gained considerable relevance, is in the analogue simulation of quantum Hamiltonians on quantum hardware [18–20]. Indeed, machine learning inspired process- and Hamiltonian-tomography methods have been developed to infer the precise Hamiltonian that NISQ devices actually simulate, given hardware imperfections [21–26].

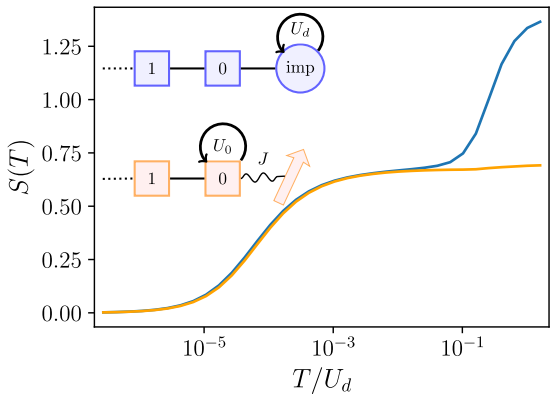

FIG. 1. Schematic of the simplest bare (blue) and effective (orange) impurity models, together with an illustrative comparison of their impurity entropies after optimization.

* j.rigo@fz-juelich.de
† andrew.mitchell@ucd.ie

To derive a good effective Hamiltonian, it is not enough to infer a model with the desired properties – it must also be meaningfully simpler than the bare model. The canonical method to systematically eliminate degrees of freedom and obtain effective models is the renormalization group (RG) [27–29]. On the level of the thermal density matrix a single RG transformation acts as the partial trace over the high energy degrees of freedom [30]. In the context of quantum physics, a central concept from information theory is that thermal states encode the corresponding Hamiltonian [31], such that comparing the ensemble of thermal states from two systems is equivalent to comparing the two Hamiltonians. This provides a way to measure the 'difference' between a Hamiltonian before and after renormalization. By minimizing the distance in Hamiltonian space, one can therefore in principle optimize a simplified effective model to best approximate some more complicated bare Hamiltonian after renormalization. We rigorously derive such an approach in this paper for a specific class of systems known as 'quantum impurity' models [32]. These models involve localized interacting quantum degrees of freedom coupled to one or more continuum baths of noninteracting conduction electrons. Generalized quantum impurity models comprise the microscopic description of semiconductor quantum dot devices [33] and complex single-molecule junctions [34]. They also underpin the theoretical understanding of correlated materials through dynamical mean field theory [35]. Developing a strategy for systematically and accurately deriving simple effective models to better understand and simulate complex systems is the ultimate application of this work.

In particular, a prerequisite for the design and use of complex quantum nanoelectronics devices with advanced functionality beyond the classical paradigm, is a fundamental understanding of their low-temperature correlated electron physics and quantum transport properties. However, this is a notoriously difficult theoretical challenge because of the subtle interplay between the orbital and spin complexity of the nanostructure, determined by its structure and chemistry; strong electron interactions due to quantum confinement; and the coupling to $\sim 10^{23}$ conduction electrons in the external circuit. This results in nontrivial quantum phenomena such as Coulomb blockade [36], various forms of Kondo effect [34, 36, 37], and quantum interference [38, 39] – all of which strongly affect low-temperature electronic conductance through the device, and hence its functionality. As with coupled quantum dot devices [40–42], entangled spin and charge degrees of freedom can give rise to new physics in single molecule junctions. It is therefore a formidable task to derive simplified effective models that can still describe this range of physics in these kinds of quantum device.

A perturbative approach to this problem in the context of molecular electronics [34] maps the microscopic model of a single-molecule junction to an effective two-channel Kondo model. This method captures simultaneously the effect on quantum transport from quantum interference and Kondo physics – but is inevitably approximate. Since thermodynamic observables flow under RG and the bare and effective model are defined at different energy scales, it is also not *a priori* clear whether such minimal Kondo models are sufficiently general to reproduce local observables of interest in the bare model [43]. In this paper we show how an RG analysis of effective interaction terms can be used to determine thermodynamic observables that are comparable across different energy scales. This allows us to introduce a novel machine learning (ML) methodology to derive accurate effective models for complex quantum impurity problems, that works by optimizing generalized (minimally constrained) models, and ensuring that local observables are correctly reproduced. We show by way of explicit examples that the low-energy Kondo physics is simultaneously captured. The parameters of the effective model are optimized by minimizing the Kullback-Leibler divergence (KLD) [44] that compares its ensemble of thermal states to that of the bare model. Information on the target is extracted from a numerical simulation of the system of interest. However, the full solution of the bare model is *not* required: minimization of the KLD requires only an estimation of thermal expectation values corresponding to specific local operators at relatively high temperatures. This can be achieved with any suitable quantum impurity solver [45–48]. Furthermore, the KLD can be shown to be convex under reasonable assumptions (see Appendix A) and its gradient is known analytically in closed form [49], which makes it an ideal optimization problem. We refer to this method as *unsupervised model learning* (UML). We demonstrate the efficacy of the UML method by application to the Anderson impurity model (AIM) [32] and obtain some new non-perturbative results for this old problem. Finally we give an outlook to the application of this framework to more complex problems, where the development of tractable effective models is essential for the study of nontrivial correlated electron physics at low temperatures.

I. METHOD

Unsupervised learning is a type of ML with the goal to recreate the probability distribution of some target data. Examples such as the Boltzmann machine [9] achieve this by minimizing the distance between the probabilistic ansatz, at the core of the machine, and the heuristic estimation of the target distribution given by some sample data [10]. The distinguishability between probability distributions can be computed using the aforementioned KLD [44]. The KLD can be generalized for quantum density matrices $\hat{\rho}$ in form of the von Neumann relative entropy [50, 51],

$$D_{KL}\left[\hat{\rho}_2 : \hat{\rho}_1\right] = \mathrm{tr}\left[\hat{\rho}_1 \log(\hat{\rho}_1)\right] - \mathrm{tr}\left[\hat{\rho}_1 \log(\hat{\rho}_2)\right] \ . \quad (1)$$

The generalized KLD quantifies how distinguishable $\hat{\rho}_1$ is from $\hat{\rho}_2$. The thermal density matrix $\hat{\rho} = \frac{1}{\mathcal{Z}} e^{-\beta \hat{H}}$

is fully defined by the system Hamiltonian $\hat{H}$ (and inverse temperature β). Thus, the thermal density matrix can be seen as a proxy for its defining Hamiltonian, and the KLD as a measure of distinguishability between two Hamiltonians. To emphasize this we denote the KLD for two thermal density matrices $\hat{\rho}_1 = \frac{1}{\mathcal{Z}_1} e^{-\beta \hat{H}_1}$ and $\hat{\rho}_2 = \frac{1}{\mathcal{Z}_2} e^{-\beta \hat{H}_2}$ as $D_{KL}[\hat{H}_2 : \hat{H}_1]$.

Given the target Hamiltonian $\hat{H}_{\text{bare}}$, we seek to optimize the simpler effective model

$$\hat{H}_{\text{eff}}(\boldsymbol{\theta}) = \sum_i \theta_i \hat{h}_i \ , \tag{2}$$

by minimizing $D_{KL}[\hat{H}_{\text{bare}} : \hat{H}_{\text{eff}}(\boldsymbol{\theta})]$ with respect to $\boldsymbol{\theta}$ for a set of operators $\{\hat{h}_i\}$. The minimization yields the optimal couplings $\boldsymbol{\theta}^*$ to represent $\hat{H}_{\text{bare}}$ with $\hat{H}_{\text{eff}}(\boldsymbol{\theta}^*)$. To achieve this we represent the impurity thermal density matrix $\hat{\rho}_{\text{imp}} = \text{tr}_{\text{bath}} \left[e^{-\beta \hat{H}} / \mathcal{Z} \right]$ as a classical probability distribution which can be used to evaluate the KLD. Using the hybridization perturbation expansion one can write the partition function of $\hat{H}_{\text{bare}}$ as a summation of the weights of all impurity occupation diagrams x, which is subdivided into a summation over all corresponding diagrams in terms of impurity eigenstates $\{\alpha\}_x$ [52]

$$\begin{aligned} \mathcal{Z} &= \mathcal{Z}_{\text{bath}} \sum_x \sum_{\{\alpha\}_x} w(\{\alpha\}_x) \ , \\ w(\{\alpha\}_x) &= e^{-\langle \hat{H}^{\text{imp}} \rangle_{\{\alpha\}_x}} \Lambda_{\{\alpha\}_x} \det(\Delta_x) \ , \end{aligned} \tag{3}$$

where $w(\{\alpha\}_x)$ is the weight of a distinct Feynman diagram labelled by the eigenstate diagram $\{\alpha\}_x$, Δ_x is the antiperiodic hybridization matrix, $\Lambda_{\{\alpha\}_x}$ the sequence of impurity operators comprising the diagram $\{\alpha\}_x$ while being projected onto the eigenbasis of $\hat{H}^{\text{imp}}$ and $\langle \hat{H}^{\text{imp}} \rangle_{\{\alpha\}_x}$ is the average value of the impurity Hamiltonian over the diagram $\{\alpha\}_x$.

From Eq. 3 we extract the distribution

$$P(\{\alpha\}_x) = (\mathcal{Z}_{\text{bath}} / \mathcal{Z}) w(\{\alpha\}_x) \ ,$$

which can be interpreted as a classical probability distribution provided that $w(\{\alpha\}_x) > 0$. This distribution acts as a proxy for the impurity density matrix and hence also for the impurity Hamiltonian. As with the classical Boltzmann machine, the probability distribution P is in the form of an energy-based model, with the weights $w(\{\alpha\}_x)$ here distributed according to the impurity Hamiltonian, $\hat{H}^{\text{imp}}$. We can therefore evaluate the KLD

$$\begin{aligned} &D_{KL}[\hat{H}_{\text{bare}} : \hat{H}_{\text{eff}}(\boldsymbol{\theta})] = \\ &\sum_x \sum_{\{\alpha\}_x:\ \text{ad}} P_{\text{bare}}(\{\alpha\}_x) \log \left[\frac{P_{\text{bare}}(\{\alpha\}_x)}{P_{\text{eff}}(\{\alpha\}_x)} \right] \ , \end{aligned} \tag{4}$$

where we have used the term *admissible* (ad) to denote diagrams that involve eigenstates of $\hat{H}^{\text{imp}}_{\text{bare}}$ for which $\hat{H}^{\text{imp}}_{\text{eff}}$ has analogues eigenstates. To minimize the distinguishability between $\hat{H}_{\text{bare}}$ and $\hat{H}_{\text{eff}}(\boldsymbol{\theta})$ we use gradient descent (GD) methods [10]. For this, we have the analytic gradient of the KLD in closed form

$$\begin{aligned} &\nabla_\theta D_{\text{KL}}[\hat{H}_{\text{bare}} : \hat{H}_{\text{eff}}(\boldsymbol{\theta})] = \\ &\beta \langle \hat{\Omega}^\dagger_{\text{ad}} \nabla_\theta \hat{H}_{\text{eff}} \hat{\Omega}_{\text{ad}} \rangle_{\text{bare}} - \beta \langle \nabla_\theta \hat{H}_{\text{eff}} \rangle_{\text{eff}} \ , \end{aligned} \tag{5}$$

where the *admissibility operator* $\hat{\Omega}_{\text{ad}}$ connects the effective and bare Fock space, by mapping the effective eigenstates to analogous bare eigenstates (see Sec. II for a detailed derivation) and thus, eliminates all non-admissible states from the bare thermal average, while on the effective Hilbert space it holds that

$$\langle \hat{\Omega}_{\text{ad}} \rangle_{\text{eff}} = \langle \hat{\mathbb{1}} \rangle_{\text{eff}} = 1 \ . \tag{6}$$

From this form of the gradient (Eq. 5) it follows immediately that the minimum is found when the impurity observables match

$$\langle \hat{\Omega}^\dagger_{\text{ad}} \hat{h}_i \hat{\Omega}_{\text{ad}} \rangle_{\text{bare}} = \langle \hat{h}_i \rangle_{\text{eff}} \tag{7}$$

for all effective impurity operators $\hat{h}_i$ of the effective model (from now on we simplify the notation $\langle \hat{\Omega}^\dagger_{\text{ad}} \hat{h}_i \hat{\Omega}_{\text{ad}} \rangle_{\text{bare}} \to \langle \hat{h}_i \rangle_{\text{bare}}$). Thus follows that according to Eq. 4 the optimal low-energy effective model $\hat{H}_{\text{eff}}$, for a given bare model $\hat{H}_{\text{bare}}$, matches the thermodynamic expectation values of all effective impurity operators $\hat{h}_i$ in the bare model.

The second order derivative of the KLD corresponds to the second order derivative of the effective free energy $\mathcal{F}_{\text{eff}}(\boldsymbol{\theta}) = -1/\beta \log \text{tr} \left[\exp(-\beta \hat{H}_{\text{eff}}(\boldsymbol{\theta})) \right]$ which, provided that all operators $\{\hat{h}_i\}$ mutually commute, becomes convex (as shown in Appendix A). It is important to note that the gradient obtained from the KLD Eq. 5 is equivalent to the gradient of the thermodynamic entropy, which is the loss-function in terms of the *maximum entropy principle*. The application of the maximum entropy principle to Hamiltonian tomography was recently proposed by Anshu et al. [21] in similar spirit as the here presented method with a great focus on convex problems, while here the focus lies on the RG analysis of the terms constituting $\hat{H}_{\text{eff}}$.

Given the convexity of Eq. 4 we can make some notes about the complexity of the optimization of Eq. 4. The complexity of the optimization depends on the method used to compute the observables comprising the gradient Eq. 5. The most expensive step is to calculate the observables $\langle \hat{h}_i \rangle_{\text{bare}}$ in the bare model. However, this needs to be done just once (e.g. with QMC methods [45, 53]). On the other hand, multiple calculations of $\langle \hat{h}_i \rangle_{\text{eff}}$ are required during the optimization. But these are performed on the simplified effective model, and are therefore inexpensive (they can be done with e.g. the NRG [46–48] (NRG) method). For the results presented in the following, we have used CT-QMC and NRG to calculate the observables and the Adagrad or Adam GD method [54].

II. EFFECTIVE HAMILTONIANS

For the UML method, it is not necessary to assume a certain type of impurity Hamiltonian, be it in *ab-initio* or *tight-binding* form. However, for illustrative purposes and without loss of generality we assume that the molecular impurity Hamiltonian is akin to a *Pariser-Parr-Pople* [55] Hamiltonian

$$\hat{H}^{\text{imp}} = \sum_{i,j,\sigma} \lambda_{ij} \hat{d}^\dagger_{i\sigma} \hat{d}_{j\sigma} + \sum_i U_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} + \sum_{i>j} U'_{ij} \hat{n}_i \hat{n}_j \ , \quad (8)$$

where $\hat{d}^\dagger_{i\sigma}$ ($\hat{d}_{i\sigma}$) creates (annihilates) a spin-σ electron in orbital i of the molecule, while $\hat{n}_{i\sigma} = \hat{d}^\dagger_{i\sigma}\hat{d}_{i\sigma}$ and $\hat{n}_i = \sum_\sigma \hat{n}_{i\sigma}$ are number operators. Properties of the molecule, like the thermodynamic average charge $\langle \hat{Q} \rangle \equiv Q$ can be controlled externally by means of a gate voltage $\hat{H}^{\text{gate}} = V_{\text{g}} \sum_i \hat{n}_i$ at a temperature T. We denote the collection of all bare Hamiltonian couplings with

$$\{\lambda_{11}, \lambda_{12}, .., U_1, U_2, .., U'_{11}, U'_{12}, .., V_{11}, V_{12}, .., V_g\} \equiv \mathbf{C} \ .$$

In general, the molecule can be coupled to any number of non-interacting fermionic baths $\hat{H}^{\text{bath}} = \sum_{k,\sigma} \epsilon_k \hat{c}^\dagger_{k\sigma} \hat{c}_{k\sigma}$. However, for simplicity we consider here only one such bath, to which the molecule couples as $\hat{H}^{\text{hyb}} = \sum_{i,k,\sigma} V_i (\hat{d}^\dagger_{i\sigma} \hat{c}_{\sigma k} + h.\,c.\,)$. Although, instead of the momentum space representation, we represent the bath a semi-infinite tight-binding chain using Wilson chain mapping [46]

$$\begin{aligned} \hat{H}^{\text{bath}} \to \hat{H}^{\text{bath}}_{\text{disc}} &= \sum_\sigma \sum_{i=0}^{\infty} t_i (\hat{c}^\dagger_{\sigma i+1} \hat{c}_{\sigma i} + \hat{c}^\dagger_{\sigma i} \hat{c}_{\sigma i+1}) \\ \hat{H}^{\text{hyb}} &\to \sum_{i,\sigma} V_i (\hat{d}^\dagger_{i\sigma} \hat{c}^\dagger_{\sigma 0} + h.\,c.\,) \ , \end{aligned} \quad (9)$$

where we assume a particle hole symmetric, metallic bath of *half-bandwidth* D. For that choice the hopping behaves like $t_n \propto D\Lambda^{-n/2}$ for large n. Thus, the complete *bare* Hamiltonian reads

$$\hat{H}_{\text{bare}} = \hat{H}^{\text{imp}} + \hat{H}^{\text{gate}} + \hat{H}^{\text{bath}}_{\text{disc}} + \hat{H}^{\text{hyb}} \ .$$

By means of V_g the molecule can be tuned in a specific charge sector, which can host a degenerate GS (of energy E_{GS}) manifold of half-integer or integer spins (we exclude the trivial case $S = 0$). Assuming $E_{\text{ex},1}$ is the energy of the first excited state then, when temperature is less than the energy gap $T < E_{\text{ex},1} - E_{GS}$ the only active part of the molecule Fock space $\mathcal{H}^{\text{imp}}_{\text{bare}}$ is the GS manifold. Thus, the entire molecule can be represented by a spin $\hat{\mathbf{S}}$ and we map

$$\mathcal{H}^{\text{imp}}_{\text{bare}} \to \mathcal{H}^{\text{imp}}_{\text{eff}} \ , \quad (10)$$

which allows for an immense reduction in complexity of the bare molecule Fock space, by replacing it with the effective Fock space $\mathcal{H}^{\text{imp}}_{\text{eff}}$. The spin $\hat{\mathbf{S}}$ replaces a degenerate GS manifold, which represents degrees of freedom to which the bath can couple. Thus, the effective Hamiltonian is structured as

$$\hat{H}_{\text{eff}}(\boldsymbol{\theta}) = \hat{H}^{\text{bath}} + \hat{H}^{ne-\text{imp}}_{\text{eff}}(\boldsymbol{\theta}) \ .$$

The effective impurity Hamiltonian $\hat{H}^{ne-\text{imp}}_{\text{eff}}(\boldsymbol{\theta})$ consists of effective interactions

$$\hat{H}^{ne-\text{imp}}_{\text{eff}}(\boldsymbol{\theta}) = \sum_i \theta_i \hat{h}_i [\{\hat{c}_{\sigma i}\}^n_{i=0}, \{\hat{c}^\dagger_{\sigma i}\}^n_{i=0}, \{\hat{S}\}] \ , \quad (11)$$

that involve the spin $\hat{\mathbf{S}}$ and a number of bath operators $\{\hat{c}^{(\dagger)}_{\sigma i}\}^n_{i=0}$. This structure of the Hamiltonian corresponds to a Kondo-type interaction, allowing for spin-flip assisted scattering of the impurity spin with the bath electrons.

A crucial aspect of the GD optimization of Eq. 4 is to compare bare and effective physical observables. The point of an effective model is to reproduce the bare quantities in the most faithful manner, thus bare and effective quantities need to be compared in a meaningful way. In the case of dynamic quantities, like GFs, this can be achieved by rescaling [56]. But for static quantities, like thermodynamic expectation values, a map between bare and effective eigenstates is required. That map is the admissibility operator $\hat{\Omega}_{\text{ad}}$. Effective model operators are only defined on the effective model Fock space thus, to apply them to a state in the bare Fock space the admissibility operator is required to relate a given bare state to the corresponding effective state. We can construct the admissibility operator by labelling the bare and effective eigenstates with QNs of symmetries $\{\text{S}_1, \text{S}_2, \dots\}$ that are common to both models

$$\begin{aligned} \hat{H}^{\text{imp}} |n\rangle = E_n |n\rangle &\mapsto \hat{H}^{\text{imp}} |\mathbf{Q}, \mathbf{q}; m_{\mathbf{Q},\mathbf{q}}\rangle \\ &= E_{\mathbf{Q},\mathbf{q};m_{\mathbf{Q},\mathbf{q}}} |\mathbf{Q}, \mathbf{q}; m_{\mathbf{Q},\mathbf{q}}\rangle \ , \end{aligned}$$

where $\mathbf{Q} = (Q_1, Q_2, \dots)$ denotes a vector of non-abelian quantum numbers and $\mathbf{q} = (q_1, q_2, \dots)$ is a vector of abelian quantum numbers. By using QNs we can define which bare states are *admissible* ("ad" for short) or not: Any state that is labelled with a QN combination that exists in the effective model Fock space is an admissible state. The label $m_{\mathbf{Q},\mathbf{q}} \in \{1, 2, \dots, M_{\mathbf{Q},\mathbf{q}}\}$ denotes an index which distinguishes multiplets with the same set of QNs. $M_{\mathbf{Q},\mathbf{q}}$ is the number of such multiplets in a given sector, which is in general smaller in the effective model than in the bare model $M_{\text{eff}} \leq M_{\text{bare}}$. This is required in order to have a meaningful reduction in complexity. For simplicity we assume that $M_{\text{eff}} = 1$, which allows us to define the admissibility operator

$$\hat{\Omega}_{\text{ad}} = \sum_{\mathbf{Q},\mathbf{q}:\text{ad}} \left[|\mathbf{Q}, \mathbf{q}\rangle_{\text{eff}} \times \sum_{m_{\mathbf{Q},\mathbf{q}}=1}^{M_{\text{bare}}} \langle \mathbf{Q}, \mathbf{q}; m_{\mathbf{Q},\mathbf{q}}|_{\text{bare}} \right] , \quad (12)$$

where the notation $\mathbf{Q}, \mathbf{q}$: ad indicates that the sum runs only over the quantum numbers that label the eigenstates of $\hat{H}^{\text{imp}}_{\text{eff}}$. Thus, operators like $\hat{h}_i$ can be meaningfully computed in the bare Fock space, using $\hat{\Omega}_{\text{ad}}$

$$\chi_i = \langle \hat{\Omega}^\dagger_{\text{ad}} \hat{h}_i \hat{\Omega}_{\text{ad}} \rangle_{\text{bare}} \, . \tag{13}$$

We provide specific examples for this in action shortly. Note that the charge of the impurity Q behaves differently compared to the other quantum numbers and must be adjusted to be comparable between bare and effective model. In the case of the effective model being spin-1/2 based, the impurity charge is always $Q = 1$, which is adequate for comparison with an sAIM (a special case of Eq. 8) [57] where the $S = 1/2$ subspace also has $Q = 1$. More complicated moieties, like for instance a six membered ring (akin to a simplified benzene molecule) host a spin-1/2 at several different charges, $Q = 7$ just to list one. To account for that we look at a fixed impurity charge sector $\tilde{Q}$ that hosts at least one spin-1/2 representation. This value $\tilde{Q}$ is then the reference value and the admissible states are then defined with respect to $\tilde{Q}$. When the bath sites are attached, the quantum number $\tilde{Q}$ is the reference charge relative to which the new charge quantum numbers in the bare extended impurity are measured to, thus we can simply set

$$\tilde{Q} \to Q = 1 \, .$$

One can take all charge sectors $\tilde{Q}$ that host a spin-1/2 into account by treating them as multiplets and summing over them. Even within a single charge sector of the bare model there can be several representations of spin-1/2, however those can be treated as multiplets and simply summed over. Since in our first application we are only dealing with a simple sAIM, we do not have to sum over different charge sectors or spin-representations.

The UML approach yields the best couplings $\boldsymbol{\theta}^*$ for the chosen Hamiltonian operators $\{\hat{h}_i\}$ in the effective model, Eq. 11, given the target Hamiltonian $\hat{H}_{\text{bare}}$, Eq. 8. Thus, the $\{\hat{h}_i\}$ have to be chosen in the most general way possible, however we will see in the following discussion that some physical intuition or prior knowledge is required. In general and for impurity models in particular the RG procedure provides a systematic way to derive *minimal* effective models. The Wilsonian RG map [27] R_l for a Hamiltonian $\hat{H}(\mathbf{C}) = \sum C_i \hat{X}_i$, with couplings $\mathbf{C} = (C_1, C_2, \dots)$ and associated interaction terms $\{\hat{X}_i\}$ allows to eliminate correlations between states above and below an energy scale $\exp(-l)D < D$ (with D being the half bandwidth) which is iteratively becoming smaller as the transformation R_l is repeated. The RG map produces a new Hamiltonian that is characterized by *renormalized* couplings

$$R_l \left[\hat{H}(\mathbf{C}) \right] = \hat{H}(\mathbf{C}') \, , \tag{14}$$

while leaving the free energy invariant $\mathcal{F}(\mathbf{C}) = \mathcal{F}(\mathbf{C}')$. Within the Wilsonian RG approach all terms $\hat{X}_i$ consistent with the symmetries and Fock space of the system are allowed to appear in $\hat{H}(\mathbf{C})$. For simplicity we make the approximation that the set of interactions $\{\hat{X}\}$ is finite and does not change during the RG procedure. The continuos flow of the couplings $\mathbf{C}$ is described by the infinitesimal RG-transformation

$$\frac{d\mathbf{C}}{dl} = \lim_{\delta l \to 0} \frac{1}{\delta l} [\mathbf{R}_{l+\delta l}(\mathbf{C}) - \mathbf{C}] \equiv \beta_l(\mathbf{C}) \, , \tag{15}$$

with β_l being the *Gell-Mann–Low β-function* [29].

An important property of RG is the existence of fixed points (FP), which are characterized by a set of couplings $\{K^*\}$ that are invariant under further RG transformations

$$R_l \left[\hat{H}(\mathbf{C}^*) \right] = \hat{H}(\mathbf{C}^*) \equiv \hat{H}^* \, , \tag{16}$$

we call $\hat{H}^*$ the FP Hamiltonian. The RG transformation can be linearised at the FP Hamiltonian

$$R_l[\hat{H}^* + \delta\hat{H}] = \hat{H}^* + \mathcal{R}^*_l \cdot \delta\hat{H} + \mathcal{O}(\delta\hat{H}^2) \, , \tag{17}$$

where $\delta\hat{H}$ is a perturbation to the fixed point. The FP can be further analysed by finding the eigensystem of the linearized RG transformation $\mathcal{R}^*_l \cdot \hat{O}^*_i = \lambda^*_i \hat{O}^*_i$. The eigenvalues λ^*_i indicated whether the operator $\hat{O}^*_i$ drives the system away from the FP if $|\lambda^*_i| > 1$ or does not impede the flow to the FP if $|\lambda^*_i| < 1$. Perturbations to the fixed point can be decomposed in to the eigensystem of $\mathcal{R}^*_l$ as $\delta\hat{H} = \sum_i \alpha_i \hat{O}^*_i$. If there are contributions α_i from operators with $|\lambda^*_i| > 1$ the perturbation is said to be *relevant*, conversely if there are only contributions of the type $|\lambda^*_i| < 1$ the perturbation is *irrelevant* (this is a strongly simplified discussion, we refer to [29] for more details).

A typical situation for non-critical FPs is that either $C^*_i = 0$ or $C^*_i = \infty$ [29], which are trivially invariant under rescaling. This makes the application of perturbation theory straightforward and strongly simplified effective FP Hamiltonians can be identified. We refer to these models as *minimal* models, as they only include RG-relevant terms and thus the bare minimum interactions required. An important example is found in the RG-flow of the sAIM at particle-hole symmetry. The symmetric sAIM is given by the bare Hamiltonian,

$$\hat{H}^A_{\text{bare}} = \tfrac{1}{2} U_d (\hat{n}^d - 1)^2 \tag{18}$$

$$+ \hat{H}^{\text{bath}}_{\text{disc}} + V \sum_\sigma (\hat{d}^\dagger_\sigma \hat{c}_{0\sigma} + \hat{c}^\dagger_{0\sigma} \hat{d}_\sigma) \, . \tag{19}$$

It is reasonable to assume that the potential energy is lower than the kinetic energy in realistic QIMs, meaning we choose sAIM parameters with $U_d/D < 1$ to keep the bandwidth the largest energy scale in the system. However, the sAIM transforms under RG such that $U_d/D \to U'_d/D' \gg 1$ and thus the interaction becomes the dominant energy scale in the system, this is known as the *local moment* (LM) FP. Under these circumstances one can perform a second order perturbation theory in

the impurity-bath hybridization V to obtain the famous Kondo model

$$\hat{H}^{\mathrm{K}}_{\mathrm{eff}} = \hat{H}^{\mathrm{bath}} + J\hat{\mathbf{S}}_d \cdot \hat{\mathbf{S}}_0 \ , \tag{20}$$

where $\hat{\mathbf{S}}_0 = \frac{1}{2}\sum_{\sigma,\sigma'} \boldsymbol{\sigma}_{\sigma\sigma'} c^\dagger_{0\sigma} c_{0\sigma'}$ is the conduction electron spin density at the impurity position. The Kondo model contains only terms that are RG-relevant to the flow of the sAIM and it encapsulates all its physics even for further scaling, away from the LM FP, this is the SWT. The SWT can also be applied to the non-renormalized theory, for which it yields $J_{\mathrm{SW}} = 8V^2/U_d$. However, this does not account for bandwidth renormalization $D \to D_{\mathrm{eff}}$. An approximate form of the bandwidth renormalization in the small U_d limit was found by Haldane [58], and discussed further by Krishnamurthy *et al* in [46], but still does not include subleading corrections from finite J. Recently, the authors determined the precise form numerically for all U_d and J_{SW} using an alternative machine learning scheme based on optimization with respect to the partition function [43]. There it was shown that while the J of the effective Kondo model, Eq. 20, can be tuned to reproduce the correct low-energy physics and Kondo scale T_{K}, a coinciding free energy does not imply that all local observables are simultaneously matched between bare and effective model since the RG-flow of the bare and effective models is not the same. Therefore, the KLD is a somewhat misleading measure of distinguishability as its minimum requires local observables to match, but matching observables does not imply a matching free energy or matching low-energy scales. The KLD is only then a good measure of distinguishability if the bare and effective model have the same RG-flow, which is not the case when only RG-relevant terms are included in the effective model. Therefore, in this chapter we go beyond the minimal models, like the Kondo model and introduce *minimally constrained* effective models. These models are completely general, determined only by the symmetries of the bare model and thus, also allow to include RG-marginal or -irrelevant terms. We show in the following sections that such an effective model has a RG-flow that is comparable to the bare model and it is hence capable of reproducing a special set of local observables as well as the universal low-energy Kondo physics.

The RG-analysis that yields the minimal model also provides insights in the reduction of degrees of freedom of the impurity; for sufficiently low temperatures ($T \lesssim U_d$) the impurity turns from a spin-full, fermionic orbital into a spin-1/2 magnetic moment. The only aspect that we adopt from the minimal model in the minimally constrained model is the effective impurity Fock space $\mathcal{H}^{\mathrm{imp}}_{\mathrm{eff}}$ that replaces the bare Fock space $\mathcal{H}^{\mathrm{imp}}_{\mathrm{bare}}$. To meaningfully define an impurity Hamiltonian, a Fock space $\mathcal{H}^{\mathrm{imp}}_{\mathrm{eff}}$ of a spin is not enough, at least one bath site has to be added to capture the impurity-bath coupling:

$$\mathcal{H}^{\mathrm{bath}}_0 \otimes \mathcal{H}^{\mathrm{imp}}_{\mathrm{eff}} = \mathcal{H}^{\mathrm{1e-imp}}_{\mathrm{eff}} \ ,$$

we refer to this as *single extended* impurity (1e − imp). To systematically generalise this we consider the semi-infinite Wilson chain Eq. 9. This allows the construction of an n-extended impurity

$$\mathcal{H}^{ne-\mathrm{imp}}_{\mathrm{eff}} = \mathcal{H}^{\mathrm{bath}}_{\mathrm{n}} \otimes \mathcal{H}^{\mathrm{bath}}_{\mathrm{n-1}} \otimes \cdots \otimes \mathcal{H}^{\mathrm{bath}}_0 \otimes \mathcal{H}^{\mathrm{imp}}_{\mathrm{eff}} \ ,$$

where $\mathcal{H}^{\mathrm{bath}}_i$ corresponds to the Fock space of Wilson site i. Note that every site that is absorbed in the extended impurity is removed from the bath

$$\hat{H}^{\mathrm{bath}}_{\mathrm{disc}} \mapsto \hat{\tilde{H}}^{\mathrm{bath}}_{\mathrm{disc}} = \sum_\sigma \sum_{i=n}^{\infty} \cdots \ .$$

The enlarged Fock space permits the inclusion of high order exchange terms with the bath and, due to the nature of the Wilson chain, RG-marginal and -irrelevant terms. It depends on the precise FP one is considering which eigenvalues the i^{th} bath site operator $\hat{c}^{(\dagger)}_{\sigma i}$ has in terms of the linearized RG transformation $\mathcal{R}^*_l$ (Eq. 17), but impurity operators have in general eigenvalues with $|\lambda^*| > 1$ and bath operators $|\lambda^*| < 1$, with $|\lambda^*|$ decreasing for increasing i. Thus, including bath sites allows one to construct effective interactions $\hat{h}_i$ that go beyond the FP point physics and are closer to the bare model in terms of the RG-flow. Thereby we systematically improve the quality of our effective model.

Following Ref. [46] we construct the minimally constrained effective Hamiltonian on the n-extended impurity Fock space, with the only constraints being the symmetries of the bare model $\{\mathrm{S}_1, \mathrm{S}_2, \dots\}$. In the following we only consider models invariant under $U(1) \equiv \mathrm{Q}$ and $SU(2) \equiv \mathrm{S}$ as they are most commonly encountered in quantum impurity models and in fact also in the sAIM, however our approach can be generalised to more symmetries. Since representations of Q and S on $\mathcal{H}^{\mathrm{1e-imp}}_{\mathrm{eff}}$ commute we can construct a basis on the effective Fock space and label it with the quantum numbers S for the non-abelian spin S quantum number, s_z for the abelian magnetic quantum number of the Cartan subalgebra of S, and Q for the abelian charge quantum number

$$\langle Q, S, s_z; m|\, Q', S', s'_z; m' \rangle = \delta_{QQ'}\delta_{SS'}\delta_{s_z s'_z}\delta_{mm'} \ , \tag{21}$$

where m is the *multiplet* index, as the same representation of S and Q can be repeated several times in the same Fock space. The multiplets are not unique: every linear combination of a multiplet of some collection of quantum numbers is again a multiplet with the same quantum numbers, hence multiplets introduce a gauge degree of freedom in the choice of the basis. Therefore, it is not possible to compare degenerate multiplets between bare and effective models. We show with the example of the sAIM that we can eliminate this ambiguity and enable meaningful analogies between bare and effective states. For now, assume we have chosen a multiplet gauge and thus, we can construct projectors on the quantum number subspaces

$$\hat{P}_{QS} = \sum_m |Q, S; m\rangle \langle Q, S; m| \ . \ , \tag{22}$$

where omitting s_z implies that we have summed over all $s_z = -S, -S+1 \ldots, S$. These projectors form the operator basis that we use to construct the effective impurity Hamiltonian

$$\hat{H}_{\text{eff}}^{n-\text{imp}}(\boldsymbol{\theta}) = \sum_{Q,S} \theta_{QS} \hat{P}_{QS} \ , \tag{23}$$

and according to our previous discussion, this leads to a convex KLD as the projectors clearly all mutually commute. This is the ideal case in terms of optimisation, but it is not the most general $SU(2) \times U(1)$ Hamiltonian on $\mathcal{H}_{\text{eff}}^{ne-\text{imp}}$, since different multiplets with the same QNs are treated independently in $\hat{H}_{\text{eff}}$, but may also be connected. With the example of the sAIM we make $\hat{H}_{\text{eff}}$ more general in the next section at the cost of the convexity of the KLD. However, here we first stick with Eq. 23. One of the advantages of the projectors $\hat{P}_{QS}$ is that $\hat{\Omega}_{\text{ad}}^\dagger \cdot \hat{P}_{QS} \cdot \hat{\Omega}_{\text{ad}}$ is trivial to compute and it makes the observables directly comparable between bare and effective model. This is because bare and effective model have the same symmetries by construction.

An important property of the projectors is that the sum over all projectors gives the identity on the effective Fock space $\mathcal{H}_{\text{eff}}^{ne-\text{imp}}$

$$\hat{\mathbb{1}}_{\text{eff}} = \sum_{Q,S} \hat{P}_{QS} \ , \tag{24}$$

from which follows that $\langle \hat{\mathbb{1}}_{\text{eff}} \rangle_{\text{eff}} = 1$. But the sum over all projectors is clearly not the identity on the bare Fock space and therefore $\langle \hat{\Omega}_{\text{ad}}^\dagger \cdot \hat{\mathbb{1}}_{\text{eff}} \cdot \hat{\Omega}_{\text{ad}} \rangle_{\text{bare}} \leq 1$. From this it follows that it is impossible to demand

$$\langle \hat{\Omega}_{\text{ad}}^\dagger \cdot \hat{P}_{QS} \cdot \hat{\Omega}_{\text{ad}} \rangle_{\text{bare}} = \langle \hat{P}_{QS} \rangle_{\text{eff}} \ , \tag{25}$$

for all Q and S combinations at the same time, as this would amount to demanding

$$\langle \hat{\mathbb{1}}_{\text{eff}} \rangle_{\text{eff}} \stackrel{!}{=} \langle \hat{\Omega}_{\text{ad}}^\dagger \cdot \hat{\mathbb{1}}_{\text{eff}} \cdot \hat{\Omega}_{\text{ad}} \rangle_{\text{bare}} \ , \tag{26}$$

which cannot be satisfied for a meaningful reduction in complexity of $\hat{H}_{\text{eff}}$ to $\hat{H}_{\text{bare}}$. Therefore at least one projector (or a multiplet within a projector) must be omitted in the Hamiltonian $\theta_{QS} = 0$, this choice is not arbitrary and an ideal choice can be made, as we show in Sec. IV.

III. SINGLE AND DOUBLE EXTENDED EFFECTIVE IMPURITY MODELS

The simplest effective Hamiltonians for the AIM are defined on a single extended (1e-imp) Fock space and the prime example is the Kondo model. However the Kondo model can only either reproduce the local observables or the universal temperature scale T_K and not both, as shown recently by the authors [43]. The minimally constrained models get increasingly capable of reproducing both by increasing the number of bath sites included in the extended impurity.

The construction of the 1e-imp minimally constrained model begins with creating the basis of the $\mathcal{H}_{\text{eff}}^{1e-\text{imp}}$ Fock space, which is straight forward as $M = 1$ for all QN combinations. The basis states are then turned into projectors associated with Q and S

$$\begin{aligned}
\hat{P}_{0,1} &= |\uparrow, \Uparrow\rangle \langle \uparrow, \Uparrow| + |\downarrow, \Downarrow\rangle \langle \downarrow, \Downarrow| + \frac{1}{2}(|\uparrow, \Downarrow\rangle \langle \uparrow, \Downarrow| \\
&\quad + |\downarrow, \Uparrow\rangle \langle \downarrow, \Uparrow| + |\downarrow, \Uparrow\rangle \langle \uparrow, \Downarrow| + |\uparrow, \Downarrow\rangle \langle \downarrow, \Uparrow|) \\
\hat{P}_{\pm 1,1/2} &= (|0\rangle \langle 0| + |\uparrow\downarrow\rangle \langle \uparrow\downarrow|) \otimes (|\Uparrow\rangle \langle \Uparrow| + |\Downarrow\rangle \langle \Downarrow|) \\
\hat{P}_{0,0} &= \frac{1}{2}(|\uparrow, \Downarrow\rangle \langle \uparrow, \Downarrow| + |\downarrow, \Uparrow\rangle \langle \downarrow, \Uparrow| - |\downarrow, \Uparrow\rangle \langle \uparrow, \Downarrow| \\
&\quad - |\uparrow, \Downarrow\rangle \langle \downarrow, \Uparrow|) \ ,
\end{aligned} \tag{27}$$

where we have assumed particle hole symmetry and $|\phi, \sigma\rangle \equiv |\phi\rangle_0 \otimes |\sigma\rangle_d$ with $\phi = \{0, \uparrow, \downarrow, \uparrow\downarrow\}$ for the local bath orbital and $\sigma = \{\Uparrow, \Downarrow\}$ for the impurity spin. The corresponding projectors, that are measured in the bare AIM must include an additional multiplier $\hat{\Omega}_{\text{ad}} = \hat{n}_\uparrow^d + \hat{n}_\downarrow^d - 2\hat{n}_\uparrow^d \hat{n}_\downarrow^d$ to project out impurity states with 0 and 2 electrons $\hat{P}_{QS} \to \hat{\Omega}_{\text{ad}} \cdot \hat{P}_{QS} \cdot \hat{\Omega}_{\text{ad}}$. Thus, the representation of $\hat{P}_{QS}$ in second quantized operators is different for bare and effective model. The effective Hamiltonian $\hat{H}_{\text{eff}}^{1n-\text{mol}} = \hat{\tilde{H}}_{\text{bath}} + \theta_{0,1}\hat{P}_{0,1} + \theta_{0,0}\hat{P}_{0,0} + \theta_{\pm 1,1/2}\hat{P}_{\pm 1,1/2}$ reads in second quantized notation

$$\hat{P}_{0,1} = \frac{3}{4}(\hat{n}_\uparrow^0 + \hat{n}_\downarrow^0 - 2\hat{n}_\uparrow^0 \hat{n}_\downarrow^1) + \hat{\mathbf{S}}_d \cdot \hat{\mathbf{S}}_0 \tag{28}$$

$$P_{0,0} = \frac{1}{4}(\hat{n}_\uparrow^1 + \hat{n}_\downarrow^1 - 2\hat{n}_\uparrow^1 \hat{n}_\downarrow^1) - \mathbf{S}_1 \cdot \mathbf{S}_2 \tag{29}$$

$$\hat{P}_{\pm 1,1/2} = \hat{\mathbb{1}}_{\text{eff}} - \hat{n}_\uparrow^0 - \hat{n}_\downarrow^0 + 2\hat{n}_\uparrow^0 \hat{n}_\downarrow^0 \ , \tag{30}$$

where $\hat{n}_0 = \sum_\sigma \hat{c}_{0\sigma}^\dagger \hat{c}_{0\sigma}$. The projectors Eq. 28 combined form the most general SU(2) and particle-hole symmetric Hamiltonian on $\mathcal{H}_{\text{eff}}^{1e-\text{imp}}$. Recombining the projectors yields a different parametrization with the same expressibility

$$\hat{H}_{\text{eff}}^{JU} = \hat{\tilde{H}}_{\text{bath}} + J\hat{\mathbf{S}}_d \cdot \hat{\mathbf{S}}_0 + \tfrac{1}{2}U_0(\hat{n}_0 - \hat{\mathbb{1}}_{\text{eff}})^2 + L\hat{\mathbb{1}}_{\text{eff}} \ . \tag{31}$$

This parametrization Eq. 31 is still meaningful in bare and effective model, because of the simplicity of the AIM. However, note that Eq. 31 includes the identity operator. As discussed before the identity is an unmatchable observable between bare and effective model and since it does not have any impact on the physics we set $L = 0$ and find $J = \theta_{0,1},\ U_0 = 2\theta_{\pm 1,1/2} - \frac{3}{2}\theta_{0,1}$. Later in the paper we compare in detail the performance of Eq. 31 and Eq. 28. The parametrization Eq. 31 shows that the $\hat{H}_{\text{eff}}^{1n-\text{mol}}$ is equivalent to the LM FP Hamiltonian including the leading irrelevant perturbation, as shown in the RG-Analysis by Krishnamurty et. al. [46]. This exemplifies how the extended impurity approach allows to include irrelevant terms in the Hamiltonian.

We can use the same procedure as for the 1e-imp to derive the 2e-imp minimally constrained Hamiltonian.

Constructing the Q and S basis of $\mathcal{H}_{\rm eff}^{\rm 2e-imp}$ one finds a total of eight Q and S combinations, but a total of 15 s_z-multiplets. Since multiplets are arbitrary they can not be compared between bare and effective model. Thus, we need additional operators that commute with the representations of Q and S such that a common eigenbasis can be formed and then allows us to identify every multiplet. Choosing specific operators creates combinations of multiplets that are not all equally suited to express the physics of the problem, we want all terms in the Hamiltonian to commute there are no terms in the Hamiltonian between QN subspaces. Hence one has to use physical intuition to make a good choice for the additional operators, however if we sacrifice the convexity in parts we can introduce terms that hybridize QN subspaces and thus again yield the most general SU(2) symmetric Hamiltonian. First we consider no such terms though, thus we introduce the hopping operator between the first and second Wilson chain sites

$$\hat{T} = \sum_{\sigma} (\hat{c}^{\dagger}_{\sigma 0}\hat{c}_{\sigma 1} + \hat{c}^{\dagger}_{\sigma 1}\hat{c}_{\sigma 0}) \, . \tag{32}$$

The $\hat{T}$ operator has the advantage that it is supported on the bath Fock-space and is thus comparable between effective and bare Fock space as it does not depend on the impurity Fock-space. Then there is a clear physical motivation why one would want to include this term, that is because it is also present in the bare Hamiltonian. Forming the common eigenspaces of S, Q and now also $\hat{T}$ we find that only one multiplet remains degenerate

$$\begin{aligned}
|0,\tfrac{1}{2},0,\tfrac{1}{2};0\rangle &= \sqrt{\frac{2}{3}}\,|\uparrow,\uparrow,\Downarrow\rangle - \sqrt{\frac{1}{6}}(|\downarrow,\uparrow,\Uparrow\rangle + |\uparrow,\downarrow,\Uparrow\rangle) \\
|0,\tfrac{1}{2},0,-\tfrac{1}{2};0\rangle &= -\sqrt{\frac{2}{3}}\,|\downarrow,\downarrow,\Uparrow\rangle + \sqrt{\frac{1}{6}}(|\downarrow,\uparrow,\Downarrow\rangle + |\uparrow,\downarrow,\Downarrow\rangle) \\
|0,\tfrac{1}{2},0,\tfrac{1}{2};1\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\downarrow,0,\Uparrow\rangle - |0,\uparrow\downarrow,\Uparrow\rangle) \\
|0,\tfrac{1}{2},0,-\tfrac{1}{2};1\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\downarrow,0,\Downarrow\rangle - |0,\uparrow\downarrow,\Downarrow\rangle) \, ,
\end{aligned} \tag{33}$$

where we have adopted the notation $|Q,S,T,s_z;m\rangle$ with T being the hopping eigenvalue. This last ambiguity can be resolved by introducing on more operator

$$\hat{W} = \hat{n}^0_{\uparrow}\hat{n}^0_{\downarrow} + \hat{n}^1_{\uparrow}\hat{n}^1_{\downarrow} \, , \tag{34}$$

which we only use in the $S = 1/2$ and $Q = 0$ subspace. Since, the hopping eigenvalues in this subspace are $T \in \{-2,0,2\}$ we do not introduce an additional quantum number for $\hat{W}$, but combine the T and W quantum numbers to $T \in \{-2,0,1,2\}$ such that $|0,\frac{1}{2},0;0\rangle \to |0,\frac{1}{2},0\rangle$ and $|0,\frac{1}{2},0;1\rangle \to |0,\frac{1}{2},1\rangle$. All states and quantum numbers for this particular basis are given in Appendix B. With the basis being now complete and unique we have again a projector basis $\hat{P}_{QST}$ that allows us to build the effective Hamiltonian $\hat{H}^{\rm 2e-imp}_{\rm eff} = \hat{\tilde{H}}_{\rm bath} + \sum_{\{QST\}:{\rm ad}} \theta_{QST}\hat{P}_{QST}$ while yielding a convex KLD.

In building the 2e-imp basis we have prioritised the convexity over generality of the Hamiltonian, we can also prioritise generality over convexity. The generality is limited because Q,S-multiplets are not connected by any term in the Hamiltonian despite Q and S symmetry in principle allowing such transitions. To get the most general Q- and S-symmetric Hamiltonian we can include the terms

$$\hat{P}_{QS;t,t'} = |Q,S,T=t\rangle\langle Q,S,T=t'| + h.c. \, , \tag{35}$$

that connect states with the same Q,S but different T and W eigenvalues. In the next section we show numerical results for the 1e-imp and 2e-imp, where we also compare the performance of the convexity prioritised and generality prioritised 2e-imp Hamiltonians.

IV. QUANTIFYING THE LEARNING PROGRESS

In the previous section we have introduced the KLD loss function Eq. 4 that quantifies the distinguishability between bare and effective model. In general ML applications, the value of the loss function is a useful metric to judge the training progress and performance of the model. In the case of the KLD for impurity models it is impossible to compute the actual value and thus, the training progress cannot be known. However, to optimize the KLD it is not necessary to evaluate it: as we know its gradient Eq. 5 we can perform a *gradient descent* (GD) optimization. Since the KLD can be shown to be convex for the minimally constrained models, the gradient vanishes only at saddle-points or global minima of the KLD. Thus, the gradient is itself a good replacement for the loss function to determine whether the training of a model has converged. Due to the special form of the gradient an optimum can only be reached when expectation values associated to the effective Hamiltonian interactions coincide in the bare and effective system

$$\langle \hat{\Omega}^{\dagger}_{\rm ad} \nabla_{\theta} \hat{H}_{\rm eff} \hat{\Omega}_{\rm ad} \rangle_{\rm bare} = \langle \nabla_{\theta} \hat{H}_{\rm eff} \rangle_{\rm eff} \, . \tag{36}$$

It is clearly a desirable property for an effective model to reproduce bare thermodynamic observables. However the fact that the bare and effective models are connected through RG implies that if their free energy for a single temperature coincides, then all other energy scales of the system will coincide too (within the validity range of the effective model). For impurity models such a scale would be the Kondo temperature T_K, which is crucial to infer in order to capture the low temperature behaviour of molecular junctions. The accuracy of the effective model can be related to the Kondo temperature in terms of the relative error

$$\Delta(T_K^{\rm eff}, T_K^{\rm bare}) = \left| 1 - \frac{\max(T_K^{\rm eff}, T_K^{\rm bare})}{\min(T_K^{\rm eff}, T_K^{\rm bare})} \right| \, . \tag{37}$$

We use $\Delta(T_K^{\rm eff}, T_K^{\rm bare})$ in the following as one of the metrics to quantify the performance of a trained effective

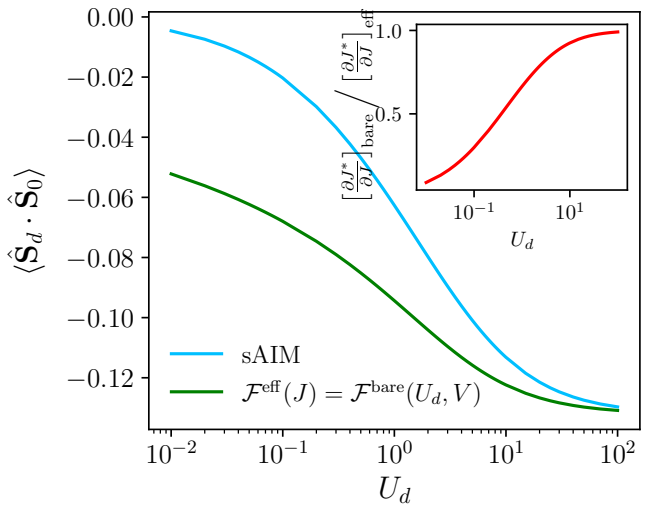

FIG. 2. Impurity-bath spin-spin expectation value $\langle \hat{\mathbf{S}}_d \cdot \hat{\mathbf{S}}_0 \rangle$ at zero temperature for the sAIM Eq. 18 (blue line) compared with the corresponding value in the minimal model Eq. 20, optimized by F-learning (green). The inset shows their ratio, which captures the different renormalizations along the two paths to the same SC fixed point. Plotted as a function of U_d for $8V^2/U_d = 0.3$.

model. Note that T_K is not directly optimized during the UML procedure, so this is a stringent and nontrivial metric. When learning a representation for realistic molecular junctions T_K^{bare} cannot be inferred and therefore the approximation error of calculated observables has to suffice

$$\Delta(\langle \hat{h}_i \rangle_{\text{eff}}, \langle \hat{h}_i \rangle_{\text{bare}}) = \left| 1 - \frac{\max(\langle \hat{h}_i \rangle_{\text{eff}}, \langle \hat{h}_i \rangle_{\text{bare}})}{\min(\langle \hat{h}_i \rangle_{\text{eff}}, \langle \hat{h}_i \rangle_{\text{bare}})} \right| . \quad (38)$$

While energy scales such as T_K and the free energy are invariant under the Wilsonian RG, thermodynamic observables flow under the RG transformation. To understand how observables flow it is instructive to interpret the RG procedure as a reparametrization of the free energy. In this picture the renormalization is achieved by replacing the bare coordinate system with a new coordinate system that leaves $\mathcal{F}$ invariant $\mathbf{C} \to \mathbf{C}'(\mathbf{C})$ (for simplicity we drop the explicit dependence of the renormalized parameters on the bare parameters from the notation). In this picture thermodynamic observables can be interpreted as covariant tensors that can be transformed as such between coordinate systems [56]

$$\begin{aligned} \langle \hat{h}_i \rangle_{\text{eff}} &= \frac{\partial}{\partial \theta_i} \mathcal{F}(\boldsymbol{\theta}) \\ &= \sum_j \frac{\partial \theta_j^*}{\partial \theta_i} \frac{\partial}{\partial \theta_j^*} \mathcal{F}(\boldsymbol{\theta}^*) = \sum_j \chi^*_{\theta_j} \frac{\partial \theta_j^*}{\partial \theta_i} , \end{aligned} \quad (39)$$

with $\boldsymbol{\theta}^*$ being the FP couplings and $\chi^*_{\theta_i}$ the FP expectation value. The matrix $\frac{\partial \theta_j^*}{\partial \theta_i}$ can be obtained from the β-function (Eq. 15) and allows us to make a connection between the FP value of an observable and the unrenormalized expectation values. Note that we have replaced $\mathbf{C}$ with $\boldsymbol{\theta}$ to emphasize the effective model parameters (Eq. 11). Eq. 39 shows that the bare expectation value is a linear combination of all possible FP expectation values, weighted by the dependence of the FP parameter on the bare parameters.

As an example we consider the Kondo model, for which it is well known that the SC FP is reached for a renormalized coupling of $J^* = \mathcal{O}(1)$. Therefore, the electrons in the system lock into a singlet to screen the impurity spin. The impurity-bath spin correlator $\langle \hat{\mathbf{S}}_d \cdot \hat{\mathbf{S}}_0 \rangle = \chi^*_J$ athe FP follows as:

$$\partial_J \mathcal{F}(J^*) = \chi^*_J = -3/4 .$$

Despite that the same mechanism dominates the bare Kondo model, one does not measure the FP value χ^*_J but

$$\langle \hat{\mathbf{S}}_d \cdot \hat{\mathbf{S}}_0 \rangle = \chi^*_J \frac{\partial J^*}{\partial J} .$$

The scaling laws of the Kondo model are known analytically from several different approaches [59], however only to a certain order in perturbation theory and therefore the estimation for $\partial_J J^*$ vanishes as $D \to 0$. This prohibits the exact analytic calculation of $\langle \hat{\mathbf{S}}_d \cdot \hat{\mathbf{S}}_0 \rangle$ from χ^*_J (the true scaling of bare couplings is far beyond tractability).

Despite that, it is still insightful to express the expectation values of effective interactions $\hat{h}_i$ (or $\hat{P}_i$) in the bare model in this formulation, even if in the end we evaluate them numerically. Because the effective interactions do not occur in the bare Hamiltonian they have to be added by means of source terms, which we also denote with θ_i. The thermodynamic expectation value can now be calculated by differentiating the free energy with respect to the source term

$$\begin{aligned} \langle \hat{h}_i \rangle_{\text{bare}} &= \frac{\partial}{\partial \theta_i} \mathcal{F}(\mathbf{C}; \theta_i)\Big|_{\theta_i=0} \\ &= \left[\frac{\partial \theta_i^*}{\partial \theta_i} \frac{\partial}{\partial \theta_i^*} + \sum_j \frac{\partial C_j^*}{\partial \theta_i} \frac{\partial}{\partial C_j^*} \right] \mathcal{F}(\mathbf{C}^*; \theta_i^*)\Big|_{\theta_i=0} \\ &= \chi^*_{\theta_i} \frac{\partial \theta_i^*}{\partial \theta_i}\Big|_{\theta_i=0} + \sum_j \chi^*_{C_j} \frac{\partial C_j^*}{\partial \theta_i}\Big|_{\theta_i=0} , \end{aligned} \quad (40)$$

where $\mathbf{C}^*$ are the bare FP parameters. The effective model is chosen such that as $D \to 0$ it approaches the same FP as the bare model. Thus, the bare FP parametrization involves the same interactions and parameters as the effective model

$$\langle \hat{h}_i \rangle_{\text{bare}} = \frac{\partial}{\partial \theta_i} \mathcal{F}(\mathbf{C}; \theta_i)\Big|_{\theta_i=0} = \sum_j \chi^*_{\theta_j} \frac{\partial \theta_j^*}{\partial \theta_i}\Big|_{\theta_i=0} , \quad (41)$$

however here we have $\theta_j^*(\mathbf{C})$, while in we have Eq. 39 $\theta_j^*(\boldsymbol{\theta})$. The FP expectation values $\chi^*_{\theta_j}$ are identical for

bare and effective models, but the bare expectation value differs from the effective one only by the different dependence of the FP parameters on the source term. Therefore, matching observables ($\langle\hat{h}_i\rangle_{\text{bare}} = \langle\hat{h}_i\rangle_{\text{eff}}$) actually means that

$$\left[\frac{\partial\theta_j^*}{\partial\theta_i}\right]_{\text{eff}} = \left[\frac{\partial\theta_j^*}{\partial\theta_i}\Big|_{\theta_i=0}\right]_{\text{bare}} , \tag{42}$$

is matched between bare and effective model. If both models would flow in the same manner to the FP this would imply that observables and energy scales can be matched at the same time. However, the way in which bare and effective model flow to the FP is not in general the same. The bare RG-flow typically involves more FPs than the effective RG-flow, which by construction is simpler. For instance, the sAIM visits the *free orbital* FP, which the Kondo model does not. As a consequence when bare and effective observables coincide, the free energies are not necessarily agreeing, and vice versa.

For the case of the sAIM and the Kondo model Eq. 20 acting as minimal model, it is found that the expectation value of the effective interaction in the effective model is upper bounded by the expectation value in the bare model

$$\langle\hat{\mathbf{S}}_d\cdot\hat{\mathbf{S}}_0\rangle_{\text{eff}} \leq \langle\hat{\mathbf{S}}_d\cdot\hat{\mathbf{S}}_0\rangle_{\text{bare}} . \tag{43}$$

We demonstrate this fact for a wide range of sAIM parameters in Fig. 2. We can rewrite this upper bound using the FP expectation value and the FP coupling derivative as

$$\chi_J^*\left[\frac{\partial J^*}{\partial J}\right]_{\text{eff}} \leq \chi_J^*\left[\frac{\partial J^*}{\partial J}\right]_{\text{bare}} \tag{44}$$

and using $\chi_J^* = -\frac{3}{4}$ if follows immediately that

$$\left[\frac{\partial J^*}{\partial J}\right]_{\text{bare}} \Big/ \left[\frac{\partial J^*}{\partial J}\right]_{\text{eff}} \leq 1 , \tag{45}$$

which is exactly what we observe in Fig. 2 and also for models beyond the minimal model. In the following we show how this discrepancy affects the ability of various effective models to capture the Kondo temperature and how the approximation can be improved by including RG-irrelevant terms.

In equation Eq. 26 we have demonstrated the importance or relating the identity operator of the effective impurity Fock space to the bare impurity Fock space. However, the identity operator is special in the sense that it does not obey Eq. 39

$$\langle\hat{\Omega}_{\text{ad}}\rangle_{\text{eff}} = \langle\mathbb{1}\rangle_{\text{eff}} = 1 \neq \sum_j \chi_j^*\frac{\partial\theta_j^*}{\partial\theta_i} = 0 , \tag{46}$$

since the identity can not influence the RG-flow as it has no influence on the Hamiltonian and therefore $\partial\theta_j^*/\partial\theta_i = 0$. In the bare Hilbert space, $\hat{\Omega}_{\text{ad}}$ behaves the same as a regular observable, for which Eq. 39 holds. Thus, it follows from Eq. 42 that knowledge of $\langle\hat{\Omega}_{\text{ad}}\rangle_{\text{bare}}$ does not allow to establish a relationship between $\langle\hat{h}_i\rangle_{\text{eff}}$ and $\langle\hat{h}_i\rangle_{\text{bare}}$ through rescaling $\langle\hat{h}_i\rangle_{\text{bare}}$ with $\langle\hat{\Omega}_{\text{ad}}\rangle_{\text{bare}}$. Hence, the importance of Eq. 26 comes not from accounting for scale differences, but from demonstrating that the optimization problem Eq. 4 is *overdetermined* when using the projectors $\hat{P}_{QS}$ as effective interactions and this holds at all scales due to Eq. 46.

V. RESULTS

As proof of principle we only study mappings for the AIM, which can be solved completely using Wilson's NRG (see Appendix C for the details of the calculations). We can use the analytical mapping to the minimal model Eq. 20 with the coupling J_{SW} as comparison for our results. This mapping gives us also a reference value for the Kondo temperature

$$T_K(D, J_{SW}) = \alpha D\sqrt{\rho_0 J_{SW}}\exp[-1/\rho_0 J_{SW} + \gamma\rho_0 J_{SW}] ,$$

with $\rho_0 = 1/2D$ the density of states of the bath at the Fermi level, $\gamma = \pi^2/4$ and $\alpha = \mathcal{O}(1)$. Our tests consist of performing the minimization of the KLD for the minimal model, the 1e-imp and the 2e-imp for an AIM with $8V^2/U_d = 0.3D$ which leads to a constant $T_K(D, J_{SW})$. The interaction strengths are drawn from between $0.01 \leq U_d/D \leq 100$ such as to cover the cases $U_d/D \ll 1$ and $U_d/D \gg 1$. The true Kondo temperature for the bare and effective models is extracted from the impurity entropy $S(T)$ [60], as shown in Fig. 3 d) for different optimized effective models and the corresponding bare model. In the following we us the definition $S(T_K) = 0.5$.

The first example is the minimization of the KLD for the minimal model Eq. 20, which contains only the RG-relevant spin-exchange term. The associated coupling J gets strongly renormalized and as a result $\frac{\partial J^*}{\partial J}$ has a different functional form for bare and effective model depending on whether $U_d/D \ll 1$ or $U_d/D \gg 1$. In the case of $U_d/D \gg 1$ the AIM is effectively already in a LM fixed point and thus the flow of the AIM and the minimal model are coinciding, which means that because of matching the spin-spin exchange term the minimal model is also able to match T_K. This behaviour is clearly shown in Fig. 3 a) in green. When $U_d/D = \mathcal{O}(1)$ the free orbital fixed point plays a significant contribution in the flow of the AIM, which explains the behaviour in Fig. 3 a) as matching the observable is now no longer equal to matching T_K. Thus, the minimal model is not fit to yield good approximations of the T_K^{bare} based on knowledge of the observables. As a result the relative error $\Delta(T_K^{\text{eff}}, T_K^{\text{bare}})$ diverges for $U_d \to 0$.

To understand more realistic scenarios only the parameter range $U_d/D < 1$ is of importance, which is

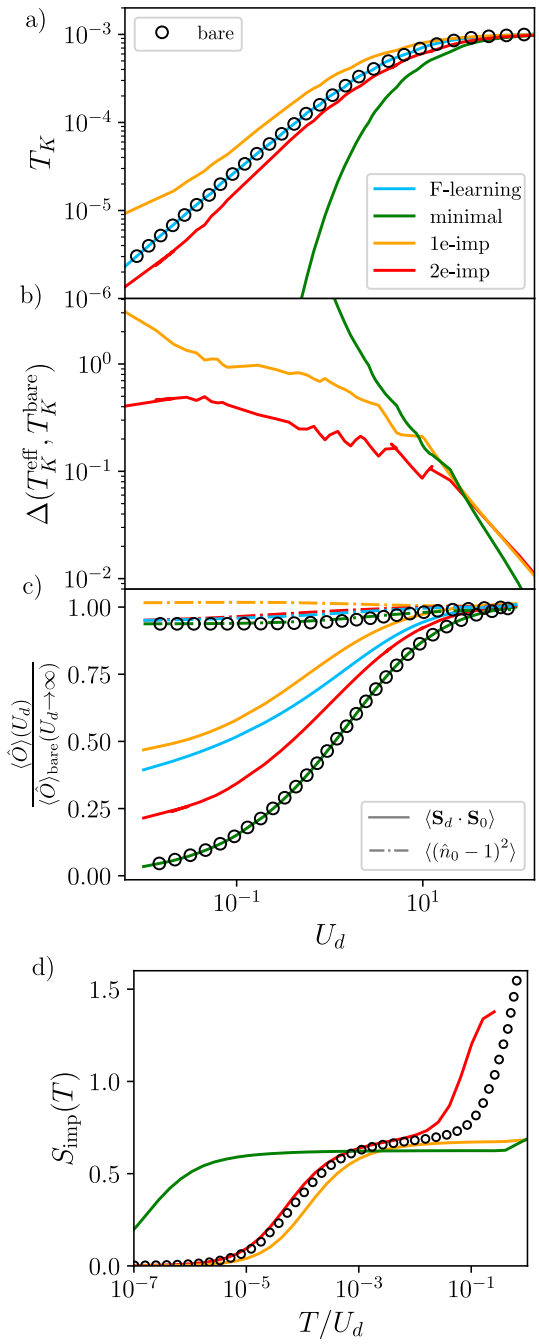

FIG. 3. Performance metrics for the optimized effective models. Here we compare the bare sAIM (points) with the results of F-learning (blue lines), and UML employing the minimal (green), 1e (orange) and 2e (red) models. As a figure of merit, panel a) shows the Kondo scale T_K for these models, while b) gives the relative error of the effective $T_K^{\rm eff}$ with respect to the bare $T_K^{\rm bare}$. Panel c) shows the dependence of RG-relevant and -irrelevant operators on U_d. To illustrate the relative quality of the various effective models, in panel d) we show their impurity entropy (lines), compared with benchmark results for the sAIM (points), for representative bare $U_d = 0.4$. Throughout we have used $8V^2/U_d = 0.3$.

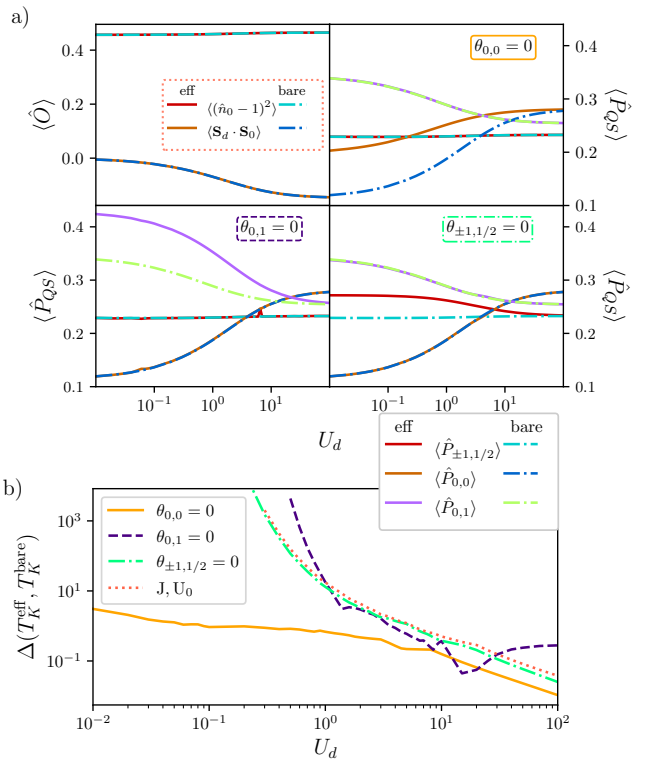

FIG. 4. Panel a) and b) compare different configurations of Eq. 28 as effective Hamiltonians. The top left panel of a) shows the observables for the $\hat{H}_{\rm eff}^{JU}$ model, while the others show the observables of the $\hat{H}_{\rm eff}^{\rm 1e-imp}$ for different eliminated interactions. The panels in a) are labelled by the eliminated interaction $\theta_{QS} = 0$. Panel b) compares the relative error of T_K for these four models.

where the most improvement is needed over the minimal model. This can be achieved with the most general SU(2) symmetric Hamiltonian we can define on $\mathcal{H}_0^{\rm bath} \otimes \mathcal{H}_{\rm eff}^{\rm mol}$. However, keeping all projectors in Eq. 28 can not converge the KLD. At least one of the projectors must be eliminated. To help to choose which one we introduce the Frobenius scalar product (FSP) on the operator space $\langle \hat{A}, \hat{B}\rangle = {\rm tr}\left[\hat{A}^\dagger \cdot \hat{B}\right]$ and the corresponding norm $\|\hat{A}\| = \sqrt{\langle \hat{A}^\dagger, \hat{A}\rangle}$. Using the FSP any projector $\hat{h}_i$ can be decomposed into its contribution from the FP Hamiltonian $\langle \hat{h}_i, \hat{H}^*\rangle/\|\hat{H}^*\|^2 = \alpha_i$ and an RG-irrelevant operator $\langle \hat{h}_i, \hat{\Lambda}_i\rangle/\|\hat{\Lambda}_i\|^2 = \beta_i$, with $\langle \hat{\Lambda}_i, \hat{H}^*\rangle = 0$. The operator $\hat{\Lambda}_i$ can easily be found after one round of the Gram-Schmidt procedure $\hat{\Lambda}_i = \hat{h}_i - \alpha_i \hat{H}^*$. The decomposition of the projector $\theta_i \hat{h}_i = \theta_i(\alpha_i \hat{H}^* + \beta_i \hat{\Lambda}) = \gamma_i \hat{H}^* + \lambda_i \hat{\Lambda}_i$ can be used

for a substitution

$$\frac{\partial}{\partial\theta_i}\mathcal{F}(\{K\},\theta_i)\Big|_{\theta_i=0} = \Big[\beta_i \underbrace{\frac{\partial}{\partial\lambda_i}}_{\text{RG irrelevant}} + \alpha_i \underbrace{\frac{\partial}{\partial\gamma_i}}_{\text{RG relevant}}\Big]\mathcal{F}(\{K\},\lambda_i,\gamma_i)\Big|_{\gamma_i=\lambda_i=0}, \tag{47}$$

which tells us that the larger α_i the stronger $\langle\hat{h}_i\rangle$ flows under RG. For the example of the minimal model it is trivially $\beta_i = 0$ and the only contribution to the flow is RG-relevant. The projectors $\hat{P}_{QS}$ (Eq. 28) on the other hand have a non-trivial decomposition. To render the overlap parameter α_i comparable we introduce $\tilde{\alpha}_i = |\langle\hat{h}_i, \hat{H}^*\rangle|/\|\hat{H}^*\|\|\hat{h}_i\|$, which is normalized with respect to the effective interaction. This provides a means to compare the strength of the renormalization of the effective interactions and therefore to decide which interaction to eliminate. For the effective interactions of $\hat{H}^{JU}_{\text{eff}}$ and $\hat{H}^{1e-imp}_{\text{eff}}$ the FSP with $\hat{H}^* = \hat{\mathbf{S}}_d \cdot \hat{\mathbf{S}}_0$ is:

	$\hat{\mathbf{S}}_d \cdot \hat{\mathbf{S}}_0$	$(\hat{n}_0 - 1)^2$	$\hat{P}_{0,1}$	$\hat{P}_{0,0}$	$\hat{P}_{\pm1,1/2}$
$\tilde{\alpha} \approx$	1	0	0.5	0.87	0

Based on the values of $\tilde{\alpha}_i$ the ideal Hamiltonian is $\hat{H}^{1e-imp}_{\text{eff}}$ with $\theta_{0,0} = 0$, as that excludes the term that gets renormalized the strongest. We put this hypothesis to the test by optimizing all four possible Hamiltonians, starting with $\hat{H}^{JU}_{\text{eff}}$. In $\hat{H}^{JU}_{\text{eff}}$ we have already set $L = 0$, which means that the minimization of the KLD can be converged. As Fig. 4 a) shows, the observables can be matched for any U_d, however Fig. 4 b) shows that the relative error $\Delta(T_K^{\text{eff}}, T_K^{\text{bare}})$ of the Kondo temperature is roughly equal to what can be obtained with the minimal model. This shows that the naive inclusion of RG-irrelevant terms does not improve the inference of energy scales, as is to be expected from Eq. 47. A very similar result in terms of $\Delta(T_K^{\text{eff}}, T_K^{\text{bare}})$ is produced by optimizing $\hat{H}^{1e-imp}_{\text{eff}}$ with $\theta_{1,0} = 0$. However, Fig. 4 a) shows that the observables that get optimized match in bare and effective model for all U_d, while $\langle\hat{P}_{0,1}\rangle$ only matches if $U_d/D \to \infty$. That implies that in this limit $\langle\hat{\mathbb{I}}_{\text{eff}}\rangle_{\text{eff}} \to \langle\hat{\Omega}^\dagger_{\text{ad}} \cdot \hat{\mathbb{I}}_{\text{eff}} \cdot \hat{\Omega}_{\text{ad}}\rangle_{\text{bare}}$, which is clearly the case as double and zero-occupation states of the impurity are strongly suppressed. Also having $\theta_{\pm,1/2} = 0$ does not improve $\Delta(T_K^{\text{eff}}, T_K^{\text{bare}})$, but keeps it roughly at the same quality as the minimal model. Only when eliminating the strongest flowing observable $\theta_{0,0} = 0$ the relative error seizes to diverge for $U_d/D \to 0$, as Fig. 4 a) shows. The results of minimizing the KLD with $\theta_{0,0} = 0$ are shown as orange line in Fig. 3 a) which follows the bare T_K curve orders of magnitude closer than the minimal model curve (green). The corresponding coupling parameters for the optimized 1e-imp model can be mapped to the parameters of Eq. 31. The coupling of the RG-relevant term can be isolated and compared to the coupling for the minimal

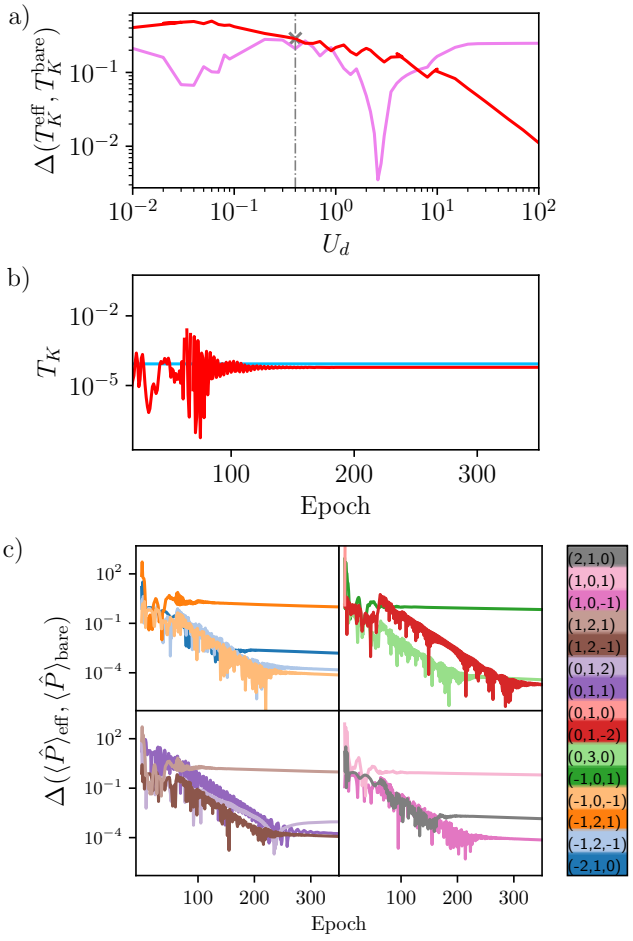

FIG. 5. Panel a) compares the 2e-imp effective Hamiltonian with (red) and without (purple) the off-diagonal terms Eq. 35. For the 2e-imp with the off-diagonal terms, panel b) shows the evolution of T_K with respect to parameter update steps (epochs). The four panels of c) show the evolution of the relative error of observables with respect to the epoch. The observables are colour coded with the quantum numbers (Q, S, T). Plotted for $U_d = 0.4$.

model, which was tuned to reproduce the bare free energy [43]. Both couplings show a great correspondence, meaning that 1e-imp provides a good balance between capturing thermodynamic observables and energy scales. Just like for the minimal model the limit $U_d/D \gg 1$ yields the SW result where $J \to 0.3$ and thus, the RG irrelevant term $U_0 \to 0$ vanishes.

This shows that the FSP provides a reliable way to identify ill-fitted interactions for the optimization. It also exemplifies that observables with at most a *moderate* flow ($\tilde{\alpha} \approx 0.5$) are best suited for the KLD distinguishability measure. Thus, 1e-imp with $\theta_{0,0} = 0$ shows an enormous improvement over the observable-matched minimal model. We now show that the performance of the effective model can be systematically improved by including more RG-irrelevant terms in the form of the

2e-imp Hamiltonian.

The projectors of the 2e-imp effective Hamiltonian are constructed from the basis $|Q, S, T, s_z; m\rangle$ Appendix B, which are defined on $\mathcal{H}_1^{\text{bath}} \otimes \mathcal{H}_0^{\text{bath}} \otimes \mathcal{H}_{\text{eff}}^{\text{mol}}$. In this basis there are a total of 15 s_z-multiplets with several degenerate spin-charge subspaces and in particular a four-fold degeneracy of the subspace $Q = 0,\ S = 1/2$, as shown in Eq. 33. Transitions between multiplets are not prohibited by symmetries. Thus not including such transitions means limiting the expressiveness of the model, while however guaranteeing a convex loss function. The impact of including the off-diagonal terms Eq. 35 can be quantified using the FSP. An optimized representation of any operator $\hat{X}$ using $\hat{H}_{\text{eff}}^{n-\text{mol}}(\{\theta\})$ can be found by minimizing the normalized FSP [61]

$$\mathcal{L}(\{\theta\}, \hat{X}) = 1 - \frac{\left|\left\langle \sum_i \theta_i \hat{P}_i, \hat{X} \right\rangle\right|}{\left\|\sum_i \theta_i \hat{P}_i\right\| \cdot \|\hat{B}\|} \,. \tag{48}$$

In the following optimization of Eq. 48 are carried out using vanilla gradient descent. To satisfy $T_K^{\text{bare}} \approx T_K^{\text{eff}}$ in the limit $U_d/D \gg 1$ the effective Hamiltonian must be able to represent the Kondo Hamiltonian Eq. 20. With $\hat{X} = \hat{\mathbf{S}}_d \cdot \hat{\mathbf{S}}_0$ it can be determined that including the off-diagonal terms optimizing Eq. 48 yields $\mathcal{L}(\{\theta_{\text{op}}\}, \hat{\mathbf{S}}_d \cdot \hat{\mathbf{S}}_0) = 0$, whereas without the off-diagonal terms the representation is limited to $\mathcal{L}(\{\theta_{\text{op}}\}, \hat{\mathbf{S}}_d \cdot \hat{\mathbf{S}}_0) \approx 0.3$. This is an enormous difference in expressiveness and since the $U_d/D \gg 1$ limit is important to satisfy, we prioritise generality over convexity in the case of 2e-imp and include the off-diagonal elements. Note that choosing $\hat{T}$ as a defining operator for the 2e-imp basis means that with and without off-diagonal elements we can find $\{\theta_{\text{op}}\}$ such that $\mathcal{L}(\{\theta_{\text{op}}\}, \hat{T}) = 0$. Including the off-diagonal terms leads to a total of 25 interactions in the 2e-imp Hamiltonian. Again, the interaction with the largest $\tilde{\alpha}$ is omitted from the Hamiltonian. The minimization of Eq. 4 can still be carried out using GD, however local minima can be encountered. We have found that running in to local minima can be avoided by initializing the optimization with the expectation values of the bare model $\langle \hat{P}_{QS;t,t'} \rangle_{\text{bare}} \xrightarrow{\text{initialize}} \theta_{QS;t,t'}$. This has the advantage that symmetries of the model are automatically respected and small bare expectation values lead directly to small effective parameters.

We have established that 2e-imp model is with the off-diagonal terms capable of representing every possible real SU(2) symmetric Hamiltonian on the effective Fock space. However, it is incapable to represent all possible SU(2) symmetric Hamiltonians on the bare Fock space. As a result we find that minimizing the KLD is only approximately possible, since there are observables that can not be simultaneously matched, despite that one observables has been omitted. We give an example for this in Fig. 5 c), where the relative error of the expectation values of all diagonal projectors is shown for every step of the optimization. The figure shows that most observables converge exponentially fast, while some are seemingly stationary. While this means that the optimization can not reach the minimum of the KLD, the parameters however converge and so does the T_K as seen in Fig. 5 b). The results of the minimizations for all U_d are shown in Fig. 3, we find an improvement of the relative error of the Kondo temperature (Fig. 3 b)) of an order of magnitude with respect to the 1e-imp result in the $U_d/D \ll 1$ regime. This result can be compared to the convex 2e-imp optimization problem without off-diagonal terms as done in Fig. 5 b). As expected the more general effective Hamiltonian performs better for $U_d/D \gg 1$, however the less general Hamiltonian yields improved accuracy at $U_d/D \ll 1$. This can be attributed to the off-diagonal elements that include terms that are more RG-relevant, which is clearly necessary to perform well at $U_d/D \gg 1$ but in the $U_d/D \ll 1$ regime these observables are stronger renormalized and hence impact the performance. However, the impact on the performance of 2e-imp is small compared to the improvement over 1e-imp, as shown in Fig. 5 b).

An other metric that indicates how well an effective model captures the bare model are observables that get not explicitly matched during the optimization process. The spin exchange term is very interesting in that regard, as it is the only RG-relevant term in the system and thus experiences the strongest renormalization. As expected, Fig. 3 c) shows that at $U_d/D \gg 1$ all effective models coincide with the bare model. For $U_d/D \ll 1$ the minimal model is trivially still perfectly matching the bare model. However, we find that while 1e-imp performs worse than the $\mathcal{Z}$-matched minimal model, the 2e-imp greatly improves over both. An other insightful observable is the density operator on the 0-orbital $(1 - \hat{n}_0)^2$, which is RG-irrelevant and does only get weakly renormalized. In fact 2e-imp and the $\mathcal{Z}$-matched minimal model reproduce the bare value in the $U_d/D \gg 1$ and $U_d/D \ll 1$ regimes. The most significant deviations occur at $U_d/D \approx 1$, they are however still small. The advantage of this analysis is that other than the T_K, observables are necessarily computable without solving the entire bare model and thus, allow cross comparison with $\Delta(\langle \hat{h}_i \rangle_{\text{eff}}, \langle \hat{h}_i \rangle_{\text{bare}})$.

This shows that energy scales and general thermodynamic observables can be captured by matching only a specific set of observables. The quality of the approximation can be improved systematically by incorporating more RG-irrelevant terms in the effective Hamiltonian.

At the example of the single Anderson impurity model we have shown how to compensate the renormalization of the effective observables compared to the bare observables. The mapping of the AIM on the 2e-imp model is used to exemplify the mapping procedure for more complicated models. The method becomes especially interesting when models are consider that lie beyond the capabilities of NRG. The learning process can be carried out with only thermodynamic observables, thus every numerical method that can compute static observables can be used to minimize the KLD – this does also include ab-

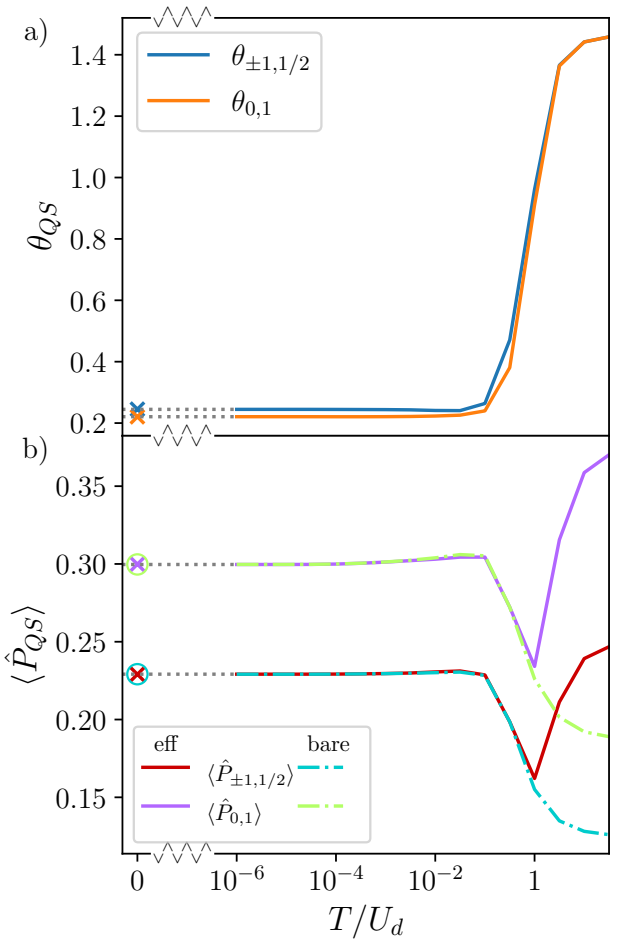

FIG. 6. Panel a) shows the evolution of optimized parameters of $\hat{H}_{\text{eff}}^{\text{1e-imp}}$ with temperature for the sAIM. The corresponding observables for the sAIM and optimized effective model are shown in b). The temperature axis is logarithmic, but includes a cut to $T = 0$ to indicate that optimization can be done using only ground state properties if required.

initio methods. While some ab-intio methods are based on the variational principle and thus operate at $T/D = 0$ (that is on the ground state [62, 63]), methods based on imaginary time evolution favour high temperatures $T/D \gtrsim 10^{-2}$ [64, 65]. Thus, it is important to verify that the here presented learning method can operate at high temperatures and at the level of the ground state to be compatible with a large variety of solvers. Regarding high temperatures, for the example of the AIM the learning can be performed once the bare model is in the local moment phase, this happens for $T/U_d = \mathcal{O}(0.1)$. In more generally terms, that are also applicable to more realistic molecules, this translates to $T/(E_{\text{ex},1} - E_g) = \mathcal{O}(0.1)$. Once the AIM is in the local moment phase we find numerically that the observables become temperature independent and can be computed with the ground state only. This is shown in Fig. 6, where the result of the trained parameters, the bare and effective observables are compared. Fig. 6 reveals a useful feature, that is for $T/U_d > 1$ the bare and effective observables can not be matched and thus, the algorithm is protected from making wrong predictions. However, to properly assure that the temperature is low enough to perform the learning a convergence test should reveal a weak or strong temperature dependence of the observables in question.

VI. OUTLOOK AND DISCUSSION

In this work we demonstrate that through the combination of the omission of the identity from the RG-flow and through the RG-analysis of all possible impurity quantum-number subspace projectors, low-energy scales of impurity models can be inferred by knowing only the corresponding thermodynamic observables.

To apply this concept in future to realistic molecular devices it is necessary to compute $\langle \hat{P}_{QS} \rangle_{\text{bare}}$. For methods that operate in second quantization we present here a straightforward way to construct the bare projectors $\hat{P}_{QS}$ that project onto the quantum number states $|Q, S\rangle$. First we introduce the auxiliary operators

$$\hat{X}_Q = \hat{Q} - Q \tag{49}$$

$$\hat{Y}_S = \hat{\mathbf{S}}^2 - S(S+1)\,, \tag{50}$$

where $\hat{\mathbf{S}}^2$ is the total spin operator of the extended impurity and $\hat{Q}$ the total charge operator of the extended impurity. The operator $\hat{X}_Q$ eliminates all contributions with charge Q to a state $|n\rangle$ and similarly $\hat{Y}_S$ eliminates all contributions with total spin S. Using these operators we can now write the projector onto the multiplet subspace of Q and S as

$$\hat{P}_{QS} = \frac{1}{N_{QS}} \prod_{Q' \neq Q} \hat{X}_{Q'} \times \prod_{S' \neq S} \hat{Y}_{S'}\,, \tag{51}$$

where

$$N_{QS} = \prod_{Q' \neq Q} (Q - Q') \prod_{S' \neq S} (S(S+1) - S'(S'+1))\,, \tag{52}$$

is the normalization constant.

However, when one plans to use the full orbital description of all atoms in the system and compute $\langle \hat{\Omega}_{\text{ad}}^\dagger \cdot \hat{P}_{QS} \cdot \hat{\Omega}_{\text{ad}} \rangle$ using ab-initio methods, then the formulation Eq. 51 of $\hat{P}_{QS}$ is inappropriate. Observables and energies for ab-initio methods are typically expressed as reduced density matrices (RDM) [66]. For a given basis of molecular orbitals $\{\phi_i\}_{i=1}^M$ we can express the spinless single particle RDM (1-RDM) as

$$\rho_1(r'; r) = \sum_{ij}^{M} {}^1D_j^i \phi_i(r') \phi_j(r)\,, \tag{53}$$

where ${}^1D^i_j$ the 1-RDM. The 1-RDM essentially expresses a single particle wave-function

$$\psi(r) = \sum_i^M d_i \phi_i(r) \ , \tag{54}$$

which then gives

$${}^1D^i_j = d_i \times d_j = \langle \psi | \hat{c}^\dagger_i \hat{c}_j | \psi \rangle \ . \tag{55}$$

The 1-RDM, is not a many body object, although it is related to the 2-RDM ${}^2D$ as [66]

$${}^2D^{pq}_{st} = 2\ {}^1D^s_p \wedge {}^1D^q_t + \ {}^2\Delta^{pq}_{st} \tag{56}$$

where ${}^2\Delta$ is a so-called cumulant. The cumulant encodes the relations between multiple determinants. The tensor $D^{pq}_{st} = \langle \psi | \hat{c}^\dagger_p \hat{c}^\dagger_q \hat{c}_s \hat{c}_t | \psi \rangle$ encodes two-body correlations. The construction of the N-RDM is possible for an arbitrary number for particles N, however it becomes prohibitively expensive rather quickly.

It becomes immediately apparent by looking at Eq. 51 that $\hat{P}_{QS}$ is an N-body operator, when N is the number of particles in the extended bare impurity Hilbert space. As just discussed, that is not tractable in the ab-initio framework. This means the projectors $\hat{P}_{QS}$ need to be simplified to be compatible with ab-initio methods. In the following we outline a possible method to achieve the reduction of the complexity of $\hat{P}_{QS}$ by approximating it with two-body operators.

Using the Frobenius scalar product, we have introduced the loss function Eq. 48, which can be used to find the best possible representation of an operator $\hat{X}$ using a linear combination of a specific set of operators $\{\hat{h}_i\}$. This can also be used to variationally find an optimal decomposition of $\hat{P}_{QS}$ in to 1- and 2-body operators. It is straightforward to write down a basis of operators for the bare Hilbert space

$$\hat{A} \in \Xi_n \in \{\{\hat{c}^\dagger_i \hat{c}_j\},\ \{\hat{c}^\dagger_i \hat{c}^\dagger_j \hat{c}_k \hat{c}_l\},\ \{\hat{c}^\dagger_i \hat{c}^\dagger_j \hat{c}^\dagger_k \hat{c}^\dagger_l \hat{c}_m \hat{c}_n \hat{c}_o \hat{c}_p\}, \dots\} \ , \tag{57}$$

retaining only the 1- and 2-body terms, this can serve as the basis $\{\hat{h}_i\}$ for the approximate representation of $\hat{P}_{QS}$. It is important to note that the basis Eq. 57 is over-complete and therefore any optimal representation of $\hat{P}_{QS}$ according to Eq. 48 is not necessarily unique. If the basis Ξ_n is complete the decomposition is unique and can be achieved like a regular spectral decomposition of a vector using the Frobenius scalar product

$$\theta_i = \frac{\langle \hat{P}_{QS}, \hat{h}_i \rangle}{\| \hat{h}_i \|},\ \hat{h}_i \in \Xi_n \ . \tag{58}$$

However, constructing a complete operator basis on the bare Hilbert space can be challenging. The variational approach can find optimal representations of any operator even for an over-complete basis and the value of the optimized loss function can serve as an error estimation for the decomposition, making this a controlled approach to obtain a 2-body approximation of the operators $\hat{P}_{QS}$. However, tests of this approach exceed the scope of this manuscript and the integration of the above framework with ab-initio methods is left for future work.

VII. CONCLUSION

In this paper we have introduced the "UML" framework, based on unsupervised ML techniques, to optimize a variational Hamiltonian. In a well-defined sense, the optimized effective Hamiltonian is the "best fit" description at a given complexity level to capture the physics of some bare quantum impurity problem. We show that the loss function is convex for the considered effective models, rendering the optimization with GD as efficient as possible. A result of the optimization procedure is that the thermal expectation values of the effective interactions in the effective model get matched to the corresponding bare expectation values. Care must be taken however, since thermodynamic observables are not invariant under RG and so bare an effective models with different RG flows may have different static expectation values. Thus, the naive matching of observables does not imply that low-energy scales, such as the Kondo temperature, are correctly recovered. We confirm this directly, already at the simplest possible example of the Anderson-Kondo model mapping. To mitigate this we identify and construct the best observables to match, based on an analysis of their RG relevance. We show that this can be done while maintaining convexity of the optimization, and that low-energy scales can be recovered. As shown by the RG-analysis, the accuracy of these effective models can be systematically improved by increasing the complexity of the effective impurity. Although this clearly offers no advantage at the level of the single impurity Anderson model, the approach is general and systematic, meaning that complex systems beyond reach of existing methods can be boiled down to tractable effective models.

An important aspect of the UML framework is that it can be performed at high temperatures without compromising precision, which allows to combine QMC methods with NRG for the optimization process. The bare model can be solved in QMC and the necessary observables can be obtained for the optimization of the effective model, which is solved repeatedly and cheaply during the optimization process using NRG. This opens the door to applying the algorithm to vastly more complex systems that are inaccessible to brute-force NRG calculations. CT-QMC [67] methods can be used to obtain the effective model for complex systems at (relatively) high temperatures, and then these effective models can be solved completely and cheaply down to $T = 0$ using NRG. In principle, *ab-initio* methods may also be used to compute approximations of the required observables to perform the UML. This may permit the future study of hitherto intractable systems.

ACKNOWLEDGMENTS

The authors would like to thank Sudeshna Sen, Alexander Munoz, Lucas K. Wagner and Denjoe O'Connor for the insightful discussion. We acknowledge funding from the Irish Research Council Laureate Awards 2017/2018 through grant IRCLA/2017/169 (JBR and AKM), from Science Foundation Ireland through grant 21/RP-2TF/10019 (AKM) and from the Helmholtz Initiative and Networking Fund, grant no. VH-NG-1711 (JBR).

Appendix A: Convexity of the loss function

For a general Hamiltonian, that can be bi-partitioned $\hat{H} = \hat{H}_0 + \hat{H}_1$, one can compute the partition function as an expansion in powers of $\hat{H}_1$ by [58],

$$\mathcal{Z} = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} T_\tau \int_0^\beta \mathrm{d}\tau_1 ... \int_0^\beta \mathrm{d}\tau_n \times \mathrm{tr}\left[e^{-\beta \hat{H}_0} \hat{H}_1(\tau_n) ... \hat{H}_1(\tau_1)\right] , \quad \text{(A1)}$$

where τ is the imaginary time. This expansion is discussed extensively in the context of continuous time quantum Monte Carlo (CT-QMC) [67–69]. In the case of hybridization expansion CT-QMC [69] $\hat{H}_1 = \hat{H}^{\mathrm{hyb}}$ describes the hybridization between a non-interacting bath $\hat{H}^{\mathrm{bath}}$ and an interacting quantum impurity $\hat{H}^{\mathrm{imp}}$. For the following discussion, we therefore consider Hamiltonians of the form,

$$\hat{H}_0 = \hat{H}^{\mathrm{bath}} + \hat{H}^{\mathrm{imp}} \quad \text{(A2a)}$$

$$\hat{H}_1 = \hat{H}^{\mathrm{hyb}} = \sum_k \sum_\sigma V_k^\sigma \hat{d}_\sigma^\dagger \hat{c}_{k\sigma} + \mathrm{H.c.} , \quad \text{(A2b)}$$

where we assume that the hybridization tensor is diagonal in the spin quantum number σ [69]. For this type of Hamiltonian Eq. A1 becomes,

$$\mathcal{Z} = \sum_{n=0}^{\infty} \int_0^\beta \mathrm{d}\tau_1 ... \int_{\tau_{n-1}}^\beta \mathrm{d}\tau_n \int_0^\beta \mathrm{d}\tau_1' ... \int_{\tau_{n-1}'}^\beta \mathrm{d}\tau_n' \sum_{\substack{a_1...a_n \\ a_1'...a_n'}} \sum_{\substack{k_1...k_n \\ k_1'...k_n'}} V_{k_1}^{a_1} V_{k_1'}^{a_1'*} ... V_{k_n}^{a_n} V_{k_n'}^{a_n'*} \quad \text{(A3)}$$
$$\mathrm{tr}\left[T_\tau e^{-\beta \hat{H}^{\mathrm{bath}}} c_{k_n' a_n'}^\dagger(\tau_n') \hat{c}_{k_n a_n}(\tau_n) ... \hat{c}_{k_1' a_1'}^\dagger(\tau_1') \hat{c}_{k_1 a_1}(\tau_1)\right]$$
$$\times \mathrm{tr}\left[T_\tau e^{-\beta \hat{H}^{\mathrm{imp}}} \hat{d}_{a_k'}(\tau_n') \hat{d}_{a_n}^\dagger(\tau_n) ... \hat{d}_{a_1'}(\tau_1') \hat{d}_{a_1}^\dagger(\tau_1)\right] .$$

Eq. A3 can be interpreted as sum over all possible diagrams obtained by allowing electrons to hop between the impurity and the bath. Since the bath is non-interacting it can be integrated out and using Wick's theorem the *antiperiodic hybridization function* can be obtained,

$$\det_{ij}\left[V_{k_i}^{\sigma_i} V_{k_j'}^{\sigma_j'*} \mathrm{tr}\left(T_\tau e^{-\beta \hat{H}^{\mathrm{bath}}} \hat{c}_{k_i\sigma_i}^\dagger(\tau_i) \hat{c}_{k_j'\sigma_j'}(\tau_j')\right)\right]$$
$$= \mathcal{Z}_{\mathrm{bath}} \det_{ij}\left[V_{k_i}^{\sigma_i} V_{k_j'}^{\sigma_j'*} \langle T_\tau \hat{c}_{k_i\sigma_i}^\dagger(\tau_i) \hat{c}_{k_j'\sigma_j'}(\tau_j')\rangle_{\mathrm{bath}}\right]$$
$$= \mathcal{Z}_{\mathrm{bath}} \det(\Delta^{(x)}) , \quad \text{(A4)}$$

where $x = (n, \{k_i, k_i', a_i, a_i', \sigma_i, \sigma_i', \tau_i, \tau_i'\}_{i=1}^n)$ denotes an impurity diagram in terms of the sequence of impurity operators [58] such that,

$$\int_0^\beta \mathrm{d}\tau_1 ... \int_{\tau_{n-1}}^\beta \mathrm{d}\tau_k \sum_{a_1...a_n} \sum_{k_1...k_n} \mapsto \sum_x . \quad \text{(A5)}$$

Following the approach of Ref. [52], we bring the impurity operators into the eigenbasis $\{(E_\alpha, |\alpha\rangle)\}$ of $\hat{H}^{\mathrm{imp}}$ and thus, $e^{-\tau \hat{H}^{\mathrm{imp}}}$ can be trivially evaluated, with

$$d_a^\dagger(\tau) = e^{-\tau_i \hat{H}^{\mathrm{imp}}} d_a^\dagger e^{\tau_i \hat{H}^{\mathrm{imp}}} = \sum_{\alpha,\alpha'} e^{\tau(E_{\alpha'} - E_\alpha)} |\alpha\rangle\langle\alpha'| \langle\alpha| d_a^{(\dagger)} |\alpha'\rangle \quad \text{(A6)}$$

and instead of occupation diagrams we obtain diagrams involving impurity eigenstates $\{\alpha\}_x \equiv \{\alpha_1 ... \alpha_k, \alpha_1' ... \alpha_k'\}_x$. Thus, Eq. A1 can be rewritten in terms of the weights of the sum of the eigenstate diagrams

$$\mathcal{Z}/\mathcal{Z}_{\mathrm{bath}} = \sum_x \sum_{\{\alpha\}} \Lambda_{\{\alpha\}_x} \det(\Delta_x) e^{-\langle \hat{H}^{\mathrm{imp}} \rangle_{\{\alpha\}_x}} , \quad \text{(A7)}$$

with $\Lambda_{\{\alpha\}_x}$ denoting the contribution from impurity operators in the impurity eigenbasis

$$\Lambda_{\{\alpha\}_x} = \prod_{i=1}^n \langle\alpha_i| d_{a_i'} |\alpha_i'\rangle \langle\alpha_i'| d_{a_i}^\dagger |\alpha_{i+1}\rangle \quad \text{(A8)}$$

and $\langle \hat{H}_{\mathrm{imp}} \rangle_{\{\alpha\}_x}$ the average impurity energy over a diagram

$$\langle \hat{H}^{\mathrm{imp}} \rangle_{\{\alpha\}_x} = \quad \text{(A9)}$$
$$\sum_{i=1}^n \left[E_{\alpha_i}(\tau_i' - \tau_{i-1}) + E_{\alpha_i'}(\tau_i - \tau_i')\right] + E_{\alpha_n}(\beta - \tau_n) .$$

Having reformulated the expansion of the impurity partition Eq. A3 as Eq. A7 we now look to prove the convexity of KLD loss function Eq. 4.

To show the convexity of Eq. 4 it does suffice to show that the free energy $\log(\mathcal{Z})$ is convex in the variational parameters $\boldsymbol{\theta}$. The core assumption of our calculation is that the effective impurity Hamiltonian is constructed as follows:

$$H^{\mathrm{imp}} = \sum_i \theta_i \hat{h}_i, \qquad [\hat{h}_i, \hat{h}_j] = 0 , \quad \text{(A10)}$$

which is the case for the effective models in the main text. For the following proof it is convenient to define a *weight* w for eigenstate diagrams to reformulate Eq. A7,

$$\mathcal{Z} = \mathcal{Z}_{\text{bath}} \sum_x \sum_{\{\alpha\}_x} w(\{\alpha\}_x) \,, \tag{A11}$$

$$w(\{\alpha\}_x) = e^{-\langle \hat{H}^{\text{imp}} \rangle_{\{\alpha\}_x}} \Lambda_{\{\alpha\}_x} \det(\Delta_x) \,, \tag{A12}$$

where we also assume that $w(\{\alpha\}_x) > 0$, such that w acts as a probability distribution upon normalization. To show that $\log(\mathcal{Z})$ is convex we need to show its Hessian is positive semi-definite,

$$\partial_{\theta_i} \partial_{\theta_j} \log(\mathcal{Z}) \succcurlyeq 0 \,. \tag{A13}$$

In the following, we use the shorthand notation ∂_i for ∂_{θ_i} for concision. It is straightforward to compute the first order derivative using the fact that the trace is invariant under circular shifts,

$$\begin{aligned}
\partial_i \ln(\mathcal{Z}) &= -\frac{1}{\mathcal{Z}} \int_0^\beta d\tau \, \text{tr} \left[T_\tau e^{-\int_0^\tau d\tau' \hat{H}} \, \hat{h}_i(\tau) \, e^{-\int_\tau^\beta d\tau' \hat{H}} \right] \\
&= -\frac{1}{\mathcal{Z}} \int_0^\beta d\tau \, \text{tr} \left[e^{-\int_0^\tau d\tau' \hat{H} - \int_\tau^\beta d\tau' \hat{H}} \, \hat{h}_i(\tau) \right] \\
&= -\frac{1}{\mathcal{Z}} \int_0^\beta d\tau \sum_\alpha e^{-\beta E_\alpha} \langle \alpha | \, e^{-\tau E_\alpha} \hat{h}_i \, e^{\tau E_\alpha} \, | \alpha \rangle \\
&= -\beta \, \text{tr}[\hat{\rho} \, \hat{h}_i] = -\beta \langle \hat{h}_i \rangle \,.
\end{aligned}$$

However we can also carry out the derivative directly on the diagrammatic expansion of the partition function,

$$\begin{aligned}
\partial_i \log(\mathcal{Z}) &= \frac{1}{\mathcal{Z}} \partial_i \left[\mathcal{Z}_{\text{bath}} \sum_x \sum_{\{\alpha\}_x} e^{-\langle \hat{H}^{\text{imp}} \rangle_{\{\alpha\}_x}} \Lambda_{\{\alpha\}_x} \det(\Delta_x) \right] \\
&= \frac{1}{\mathcal{Z}} \left[\mathcal{Z}_{\text{bath}} \sum_x \sum_{\{\alpha\}_x} \partial_i w(\{\alpha\}_x) \right] \,.
\end{aligned} \tag{A14}$$

It holds that a given eigenbasis $\{(E_\alpha, |\alpha\rangle)\}$ of $\hat{H}^{\text{imp}}$ depends in the following way on the parameters $\boldsymbol{\theta}$,

$$\partial_i E_\alpha = \langle \alpha | \, \hat{h}_i \, | \alpha \rangle \,, \tag{A15}$$

$$\partial_i \, | \alpha \rangle = \sum_{\alpha \neq \beta} \frac{\langle \beta | \, \hat{h}_i \, | \alpha \rangle}{E_\alpha - E_\beta} \, | \beta \rangle \,. \tag{A16}$$

With the assumptions Eq. A10 in place all operators in the Hamiltonian are mutually commuting, implying that they share a common set of eigenvectors. This property allows us to trivially evaluate these derivatives,

$$\partial_i E_\alpha = \epsilon_i^\alpha \tag{A17}$$

$$\partial_i \, | \alpha \rangle = 0 \,, \tag{A18}$$

with $\hat{h}_i \, | \alpha \rangle = \epsilon_i^\alpha \, | \alpha \rangle$. With these relationships it is straightforward to show that,

$$\partial_j \Lambda_{\{\alpha\}_x} = 0 \tag{A19}$$

$$\partial_j \langle \hat{H}^{\text{imp}} \rangle_{\{\alpha\}_x} = \langle \hat{h}_j \rangle_{\{\alpha\}_x} \,. \tag{A20}$$

This allows the calculation of the derivative of w, which amounts to,

$$-\partial_i w(\{\alpha\}_x) = \langle \hat{h}_i \rangle_{\{\alpha\}_x} w(\{\alpha\}_x) \,. \tag{A21}$$

This shows that w is a generating functional of moments of $\langle \hat{h}_i \rangle_{\{\alpha\}_x}$. We therefore write,

$$-\partial_i \log(\mathcal{Z}) = -\frac{1}{\mathcal{Z}} \left[\mathcal{Z}_{\text{bath}} \sum_x \sum_{\{\alpha\}_x} \partial_i w(\{\alpha\}_x) \right] \tag{A22}$$

$$= \mathbb{E}_w \left[\langle \hat{h}_i \rangle_{\{\alpha\}_x} \right] \,. \tag{A23}$$

Accordingly the second order derivative becomes,

$$\partial_j \partial_i \log(\mathcal{Z}) = \mathbb{E}_w \left[\langle \hat{h}_i \rangle_{\{\alpha\}_x} \langle \hat{h}_j \rangle_{\{\alpha\}_x} \right] \tag{A24}$$

$$- \mathbb{E}_w \left[\langle \hat{h}_i \rangle_{\{\alpha\}_x} \right] \mathbb{E}_w \left[\langle \hat{h}_j \rangle_{\{\alpha\}_x} \right] \,, \tag{A25}$$

which can be identified as the covariance. Since the covariance is a positive-semidefinite matrix, we have,

$$\partial_j \partial_i \log(\mathcal{Z}) = \text{cov}_w \left[\langle \hat{h}_i \rangle_{\{\alpha\}_x}, \, \langle \hat{h}_j \rangle_{\{\alpha\}_x} \right] \succcurlyeq 0 \,, \tag{A26}$$

making the Hessian of $\log(\mathcal{Z})$ positive semi-definite and the same applies to the Hessian of the KLD loss function Eq. 4 under the aforementioned assumptions Eq. A10.

Appendix B: Double extended basis

To construct the projectors of the 2e-imp effective Hamiltonian, we require the basis $|Q, S, T, s_z; m\rangle$ on the double extended Fock space $\mathcal{H}_1^{\text{bath}} \otimes \mathcal{H}_0^{\text{bath}} \otimes \mathcal{H}_{\text{eff}}^{\text{imp}}$. The first step in the construction is to determine the charge $\hat{Q}$ and total spin $\hat{\mathbf{S}}^2$ eigenstates $|Q, S\rangle$. For the occupation basis we use the labelling convention:

$$|\text{orbital 2 (L)}, \text{orbital 1 (R)}, \text{local moment (M)}\rangle \,. \tag{B1}$$

Thus, the Hilbert space is spanned by the states:

$\lvert -2, 1/2, s_z; m\rangle$	$\lvert -1, 1, s_z; m\rangle$	$\lvert -1, 0, s_z; m\rangle$
$\lvert 0, 0, \Uparrow\rangle$	$\lvert\uparrow, 0, \Uparrow\rangle$	$1/\sqrt{2}(\lvert\uparrow, 0, \Downarrow\rangle - \lvert\downarrow, 0, \Uparrow\rangle)$
$\lvert 0, 0, \Downarrow\rangle$	$\lvert\downarrow, 0, \Downarrow\rangle$	$1/\sqrt{2}(\lvert 0, \uparrow, \Downarrow\rangle - \lvert 0, \downarrow, \Uparrow\rangle)$
	$1/\sqrt{2}(\lvert\uparrow, 0, \Downarrow\rangle + \lvert\downarrow, 0, \Uparrow\rangle)$	
	$\lvert 0, \uparrow, \Uparrow\rangle$	
	$\lvert 0, \downarrow, \Downarrow\rangle$	
	$1/\sqrt{2}(\lvert 0, \uparrow, \Downarrow\rangle + \lvert 0, \downarrow, \Uparrow\rangle)$	

$\lvert 0, 3/2, s_z; m\rangle$	$\lvert 0, 1/2, s_z; m\rangle$	$\lvert 0, 1/2, s_z; m\rangle$
$\lvert\uparrow, \uparrow, \Uparrow\rangle$	$\sqrt{2/3}\lvert\uparrow, \uparrow, \Downarrow\rangle - \sqrt{1/6}(\lvert\downarrow, \uparrow, \Uparrow\rangle + \lvert\uparrow, \downarrow, \Uparrow\rangle)$	$\lvert\uparrow\downarrow, 0, \Uparrow\rangle$
$1/\sqrt{3}(\lvert\downarrow, \uparrow, \Downarrow\rangle + \lvert\downarrow, \downarrow, \Uparrow\rangle + \lvert\uparrow, \downarrow, \Downarrow\rangle)$	$-\sqrt{2/3}\lvert\downarrow, \downarrow, \Uparrow\rangle + \sqrt{1/6}(\lvert\downarrow, \uparrow, \Downarrow\rangle + \lvert\uparrow, \downarrow, \Downarrow\rangle)$	$\lvert\uparrow\downarrow, 0, \Downarrow\rangle$
$1/\sqrt{3}(\lvert\downarrow, \uparrow, \Uparrow\rangle + \lvert\uparrow, \uparrow, \Downarrow\rangle + \lvert\uparrow, \downarrow, \Uparrow\rangle)$		$\lvert 0, \uparrow\downarrow, \Uparrow\rangle$
$\lvert\downarrow, \downarrow, \Downarrow\rangle$		$\lvert 0, \uparrow\downarrow, \Downarrow\rangle$
		$\sqrt{1/2}(\lvert\downarrow, \uparrow, \Uparrow\rangle - \lvert\uparrow, \downarrow, \Uparrow\rangle)$
		$\sqrt{1/2}(\lvert\downarrow, \uparrow, \Downarrow\rangle - \lvert\uparrow, \downarrow, \Downarrow\rangle)$

The states $|2, 1/2, s_z; m\rangle, |1, 1, s_z; m\rangle$ and $|1, 0, s_z; m\rangle$ can be obtained straightforwardly by replacing all 0 entries with $\uparrow\downarrow$. It is important to note that using only Q and S quantum numbers, the $Q = -1, S = 1$ subspace is twofold degenerate ($m \leq M = 2$) and the $Q = 0, S = 1/2$ subspace is threefold degenerate ($m \leq M = 3$). Therefore, using only the projectors onto the Q and S quantum number subspaces limits the expressibility of our effective model. Full tunability of the model requires that we can address each basis state individually. The next step is thus to lift these degeneracies. We may do this by diagonalizing the hopping operator in the degenerate subspaces,

$$\hat{T} = \sum_{\sigma} \hat{c}^{\dagger}_{L\sigma}\hat{c}_{R\sigma} + \hat{c}^{\dagger}_{R\sigma}\hat{c}_{L\sigma} \ . \tag{B2}$$

In the $|-1(1), 1\rangle$ and $|-1(1), 0\rangle$ sub-spaces we can find a common eigenbasis $|Q, S, T, s_z; m\rangle$ of $\hat{T}$ and $\hat{\mathbf{S}}^2$ that allows to completely lift the multiplet degeneracy:

$\lvert -1, 1, +1, s_z; m\rangle$	$\lvert -1, 1, -1, s_z; m\rangle$
$1\sqrt{2}(\lvert\uparrow, 0, \Uparrow\rangle + \lvert 0, \uparrow, \Uparrow\rangle)$	$1\sqrt{2}(\lvert\uparrow, 0, \Uparrow\rangle - \lvert 0, \uparrow, \Uparrow\rangle)$
$1/\sqrt{4}(\lvert 0, \uparrow, \Downarrow\rangle + \lvert 0, \downarrow, \Uparrow\rangle + \lvert\uparrow, 0, \Downarrow\rangle + \lvert\downarrow, 0, \Uparrow\rangle)$	$1/\sqrt{4}(\lvert 0, \uparrow, \Downarrow\rangle + \lvert 0, \downarrow, \Uparrow\rangle - \lvert\uparrow, 0, \Downarrow\rangle - \lvert\downarrow, 0, \Uparrow\rangle)$
$1\sqrt{2}(\lvert\downarrow, 0, \Downarrow\rangle + \lvert 0, \downarrow, \Downarrow\rangle)$	$1\sqrt{2}(\lvert\downarrow, 0, \Downarrow\rangle - \lvert 0, \downarrow, \Downarrow\rangle)$

Here we have used that

$$\hat{T}\,|0, \downarrow\rangle = -\,|\downarrow, 0\rangle \ .$$

Thus, the symmetric combination of $|-1(1), 1(0), s_z; m\rangle$ states is associated to the negative $T = -1$ eigenvalue of $\hat{T}$ and the asymmetric combination to the positive eigenvalue $T = 1$ of $\hat{T}$.

In the $Q = 0$ subspace one can show that the hopping has no effect on the quadruplet and the doublet. However the $Q = 0, S = 1/2$ subspace behaves non-trivially under the hopping:

$$\hat{T}\,|\uparrow\downarrow, 0, \Uparrow\rangle = |\downarrow, \uparrow, \Uparrow\rangle - |\uparrow, \downarrow, \Uparrow\rangle$$
$$\hat{T}\,|0, \uparrow\downarrow, \Uparrow\rangle = |\downarrow, \uparrow, \Uparrow\rangle - |\uparrow, \downarrow, \Uparrow\rangle$$
$$\hat{T}(|\downarrow, \uparrow, \Uparrow\rangle - |\uparrow, \downarrow, \Uparrow\rangle)/\sqrt{2} = \sqrt{2}(|\uparrow\downarrow, 0, \Uparrow\rangle + |0, \uparrow\downarrow, \Uparrow\rangle)$$

We can combine these states such that they become eigenstates of $\hat{T}$:

$\lvert 0, 1/2, 0, s_z; m\rangle$	$\lvert 0, 1/2, +2, s_z; m\rangle$	$\lvert 0, 1/2, -2, s_z; m\rangle$
$\sqrt{\frac{2}{3}}\lvert\uparrow, \uparrow, \Downarrow\rangle - \sqrt{\frac{1}{6}}(\lvert\downarrow, \uparrow, \Uparrow\rangle + \lvert\uparrow, \downarrow, \Uparrow\rangle)$	$\frac{1}{\sqrt{4}}(\lvert\downarrow, \uparrow, \Uparrow\rangle - \lvert\uparrow, \downarrow, \Uparrow\rangle + \lvert\uparrow\downarrow, 0, \Uparrow\rangle + \lvert 0, \uparrow\downarrow, \Uparrow\rangle)$	$\frac{1}{\sqrt{4}}(\lvert\downarrow, \uparrow, \Uparrow\rangle - \lvert\uparrow, \downarrow, \Uparrow\rangle - \lvert\uparrow\downarrow, 0, \Uparrow\rangle - \lvert 0, \uparrow\downarrow, \Uparrow\rangle)$
$-\sqrt{\frac{2}{3}}\lvert\downarrow, \downarrow, \Uparrow\rangle + \sqrt{\frac{1}{6}}(\lvert\downarrow, \uparrow, \Downarrow\rangle + \lvert\uparrow, \downarrow, \Downarrow\rangle)$	$\frac{1}{\sqrt{4}}(\lvert\downarrow, \uparrow, \Downarrow\rangle - \lvert\uparrow, \downarrow, \Downarrow\rangle + \lvert\uparrow\downarrow, 0, \Downarrow\rangle + \lvert 0, \uparrow\downarrow, \Downarrow\rangle)$	$\frac{1}{\sqrt{4}}(\lvert\downarrow, \uparrow, \Downarrow\rangle - \lvert\uparrow, \downarrow, \Downarrow\rangle - \lvert\uparrow\downarrow, 0, \Downarrow\rangle - \lvert 0, \uparrow\downarrow, \Downarrow\rangle)$
$\frac{1}{\sqrt{2}}(\lvert\uparrow\downarrow, 0, \Uparrow\rangle - \lvert 0, \uparrow\downarrow, \Uparrow\rangle)$		
$\frac{1}{\sqrt{2}}(\lvert\uparrow\downarrow, 0, \Downarrow\rangle - \lvert 0, \uparrow\downarrow, \Downarrow\rangle)$		

The *fourfold* degenerate multiplet is now almost completely lifted. Only the $|0, 1/2, 0, s_z; m\rangle$ subspace is still twofold degenerate. This last degeneracy can be lifted by introducing a further operator,

$$\hat{W} = \hat{n}^L_\uparrow \hat{n}^L_\downarrow + \hat{n}^R_\uparrow \hat{n}^R_\downarrow \,,$$

which has the eigenvales $W \in \{0, 1\}$ in the $|0, 1/2, 0, s_z; m\rangle$ subspace. Instead of introducing a new label for only this subspace, the eigenvalues for $\hat{T}$ and $\hat{W}$ are added together in the T label. This gives the states:

$\|0, 1/2, 1, s_z; m\rangle$	$\|0, 1/2, 0, s_z; m\rangle$
$\sqrt{\frac{2}{3}}\,\|\uparrow,\uparrow,\Downarrow\rangle - \sqrt{\frac{1}{6}}(\|\downarrow,\uparrow,\Uparrow\rangle + \|\uparrow,\downarrow,\Uparrow\rangle)$	$\frac{1}{\sqrt{2}}(\|\uparrow\downarrow,0,\Uparrow\rangle - \|0,\uparrow\downarrow,\Uparrow\rangle)$
$-\sqrt{\frac{2}{3}}\,\|\downarrow,\downarrow,\Uparrow\rangle + \sqrt{\frac{1}{6}}(\|\downarrow,\uparrow,\Downarrow\rangle + \|\uparrow,\downarrow,\Downarrow\rangle)$	$\frac{1}{\sqrt{2}}(\|\uparrow\downarrow,0,\Downarrow\rangle - \|0,\uparrow\downarrow,\Downarrow\rangle)$

We have thereby succeeded in reducing the dimension of each subspace spanned by these labels to $M = 1$, as desired.

Appendix C: Details of NRG calculations

All calculations in this work were carried out using Wilson's NRG method [60, 70] using the full density matrix approach [71, 72], which allows the precise calculation of the static observables discussed here. The impurity entropy, which is used to estimate T_K was computed using standard thermodynamic NRG [70]. For the simulation of bare and effective models, the number of kept states is $M_K = 2000$ and we use a Wilson chain discretization parameter of $\Lambda = 2.5$ at a Wilson chain length of $N = 40$. Total charge and spin projection quantum numbers are exploited.

[1] D. J. Thouless, *The quantum mechanics of many-body systems* (Courier Corporation, 2014).
[2] J. Hubbard, Electron correlations in narrow energy bands, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **276**, 238 (1963).
[3] J. Kanamori, Electron correlation and ferromagnetism of transition metals, Progress of Theoretical Physics **30**, 275 (1963).
[4] F. Aryasetiawan, M. Imada, A. Georges, G. Kotliar, S. Biermann, and A. Lichtenstein, Frequency-dependent local interactions and low-energy effective models from electronic structure calculations, Physical Review B **70**, 195104 (2004).
[5] E. Şaşıoğlu, C. Friedrich, and S. Blügel, Effective coulomb interaction in transition metals from constrained random-phase approximation, Physical Review B **83**, 121101 (2011).
[6] N. P. Bauman and K. Kowalski, Coupled cluster downfolding theory: towards universal many-body algorithms for dimensionality reduction of composite quantum systems in chemistry and materials science, Materials Theory **6**, 1 (2022).
[7] H. Zheng, H. J. Changlani, K. T. Williams, B. Busemeyer, and L. K. Wagner, From real materials to model hamiltonians with density matrix downfolding, Frontiers in Physics **6**, 43 (2018).
[8] R. Pariser and R. G. Parr, A semi-empirical theory of the electronic spectra and electronic structure of complex unsaturated molecules. i., The Journal of Chemical Physics **21**, 466 (1953); A semi-empirical theory of the electronic spectra and electronic structure of complex unsaturated molecules. ii, **21**, 767 (1953).
[9] H. C. Nguyen, R. Zecchina, and J. Berg, Inverse statistical problems: from the inverse ising problem to data science, Advances in Physics **66**, 197 (2017).
[10] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning* (MIT Press, 2017).
[11] H. S. Bhat, K. Ranka, and C. M. Isborn, Machine learning a molecular hamiltonian for predicting electron dynamics, International Journal of Dynamics and Control **8**, 1089 (2020).
[12] Z. Wang, S. Ye, H. Wang, J. He, Q. Huang, and S. Chang, Machine learning method for tight-binding hamiltonian parameterization from ab-initio band structure, npj Computational Materials **7**, 1 (2021).
[13] R. C. Sawaya and S. R. White, Constructing hubbard models for the hydrogen chain using sliced-basis density matrix renormalization group, Physical Review B **105**, 045145 (2022).
[14] M. Nakhaee, S. Ketabi, and F. Peeters, Machine learning approach to constructing tight binding models for solids with application to bitecl, Journal of Applied Physics **128**, 215107 (2020).
[15] M. Prüfer, T. V. Zache, P. Kunkel, S. Lannig, A. Bonnin, H. Strobel, J. Berges, and M. K. Oberthaler, Experimental extraction of the quantum effective action for a non-equilibrium many-body system, Nature Physics **16**, 1012 (2020).
[16] S. Yu, Y. Gao, B.-B. Chen, and W. Li, Learning the effective spin hamiltonian of a quantum magnet, Chinese Physics Letters **38**, 097502 (2021).
[17] T. Schuster, M. Niu, J. Cotler, T. O'Brien, J. R. McClean, and M. Mohseni, Learning quantum systems via out-of-time-order correlators, arXiv preprint

arXiv:2208.02254 (2022).
[18] I. Buluta and F. Nori, Quantum simulators, Science **326**, 108 (2009); I. M. Georgescu, S. Ashhab, and F. Nori, Quantum simulation, Reviews of Modern Physics **86**, 153 (2014).
[19] P. Barthelemy and L. M. Vandersypen, Quantum dot systems: a versatile platform for quantum simulations, Annalen der Physik **525**, 808 (2013); T. Hensgens, T. Fujita, L. Janssen, X. Li, C. Van Diepen, C. Reichl, W. Wegscheider, S. Das Sarma, and L. M. Vandersypen, Quantum simulation of a fermi–hubbard model using a semiconductor quantum dot array, Nature **548**, 70 (2017).
[20] W. Pouse, L. Peeters, C. L. Hsueh, U. Gennser, A. Cavanna, M. A. Kastner, A. K. Mitchell, and D. Goldhaber-Gordon, Quantum simulation of an exotic quantum critical point in a two-site charge kondo circuit, Nature Physics **19**, 492 (2023).
[21] A. Anshu, S. Arunachalam, T. Kuwahara, and M. Soleimanifar, Sample-efficient learning of interacting quantum systems, Nature Physics **17**, 931 (2021).
[22] N. Boulant, T. Havel, M. Pravia, and D. Cory, Robust method for estimating the lindblad operators of a dissipative quantum process from measurements of the density operator at multiple time points, Physical Review A **67**, 042322 (2003).
[23] J. Zhang and M. Sarovar, Quantum hamiltonian identification from measurement time traces, Physical review letters **113**, 080401 (2014).
[24] M. Kieferová and N. Wiebe, Tomography and generative training with quantum boltzmann machines, Physical Review A **96**, 062327 (2017).
[25] T. Olsacher, L. Pastori, C. Kokail, L. M. Sieberer, and P. Zoller, Digital quantum simulation, learning of the floquet hamiltonian, and quantum chaos of the kicked top, Journal of Physics A: Mathematical and Theoretical **55**, 334003 (2022).
[26] S. Nandy, M. Schmitt, M. Bukov, and Z. Lenarčič, Reconstructing effective hamiltonians from nonequilibrium (pre-) thermal steady states, arXiv preprint arXiv:2308.08608 (2023).
[27] K. G. Wilson, Renormalization group and critical phenomena. i. renormalization group and the kadanoff scaling picture, Phys. Rev. B **4**, 3174 (1971).
[28] S. D. Głazek and K. G. Wilson, Renormalization of hamiltonians, Physical Review D **48**, 5863 (1993).
[29] P. Kopietz, L. Bartosch, and F. Schütz, *Introduction to the functional renormalization group*, Vol. 798 (Springer Science & Business Media, 2010).
[30] C. Bény and T. J. Osborne, Information-geometric approach to the renormalization group, Phys. Rev. A **92**, 022330 (2015).
[31] J. R. Garrison and T. Grover, Does a single eigenstate encode the full hamiltonian?, Physical Review X **8**, 021026 (2018).
[32] A. C. Hewson, *The Kondo Problem to Heavy Fermions*, Cambridge Studies in Magnetism (Cambridge University Press, 1993).
[33] D. Goldhaber-Gordon, H. Shtrikman, D. Mahalu, D. Abusch-Magder, U. Meirav, and M. Kastner, Kondo effect in a single-electron transistor, Nature **391**, 156 (1998).
[34] A. K. Mitchell, K. G. L. Pedersen, P. Hedegård, and J. Paaske, Kondo blockade due to quantum interference in single-molecule junctions, Nature Communications **8**, 15210 EP (2017), article.
[35] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions, Reviews of Modern Physics **68**, 13 (1996).
[36] J. Park, A. N. Pasupathy, J. I. Goldsmith, C. Chang, Y. Yaish, J. R. Petta, M. Rinkoski, J. P. Sethna, H. D. Abruña, P. L. McEuen, and D. C. Ralph, Coulomb blockade and the kondo effect in single-atom transistors, Nature **417**, 722 (2002).
[37] G. D. Scott and D. Natelson, Kondo resonances in molecular devices, ACS Nano **4**, 3560 (2010).
[38] C. M. Guédon, H. Valkenier, T. Markussen, K. S. Thygesen, J. C. Hummelen, and S. J. Van Der Molen, Observation of quantum interference in molecular charge transport, Nature nanotechnology **7**, 305 (2012).
[39] S. Sen and A. K. Mitchell, Many-body quantum interference route to the two-channel kondo effect: Inverse design for molecular junctions, arXiv preprint arXiv:2310.14775 (2023).
[40] A. Keller, S. Amasha, I. Weymann, C. Moca, I. Rau, J. Katine, H. Shtrikman, G. Zaránd, and D. Goldhaber-Gordon, Emergent su (4) kondo physics in a spin–charge-entangled double quantum dot, Nature Physics **10**, 145 (2014).
[41] Z. Iftikhar, A. Anthore, A. Mitchell, F. Parmentier, U. Gennser, A. Ouerghi, A. Cavanna, C. Mora, P. Simon, and F. Pierre, Tunable quantum criticality and super-ballistic transport in a "charge" kondo circuit, Science **360**, 1315 (2018); A. K. Mitchell, L. Landau, L. Fritz, and E. Sela, Universality and scaling in a charge two-channel kondo device, Physical review letters **116**, 157202 (2016).
[42] A. K. Mitchell, A. Liberman, E. Sela, and I. Affleck, SO(5) non-Fermi liquid in a Coulomb box device, Physical Review Letters **126**, 147702 (2021).
[43] J. B. Rigo and A. K. Mitchell, Machine learning effective models for quantum systems, Physical Review B **101**, 241105 (2020).
[44] S.-i. Amari, Manifold, divergence and dually flat structure, in *Information Geometry and Its Applications* (Springer Japan, Tokyo, 2016) pp. 3–30.
[45] P. Seth, I. Krivenko, M. Ferrero, and O. Parcollet, Triqs/cthyb: A continuous-time quantum monte carlo hybridisation expansion solver for quantum impurity problems, Computer Physics Communications **200**, 274 (2016).
[46] H. Krishna-Murthy, J. Wilkins, and K. Wilson, Renormalization-group approach to the anderson model of dilute magnetic alloys. i. static properties for the symmetric case, Physical Review B **21**, 1003 (1980).
[47] R. Bulla, T. A. Costi, and T. Pruschke, Numerical renormalization group method for quantum impurity systems, Rev. Mod. Phys. **80**, 395 (2008).
[48] A. K. Mitchell, M. R. Galpin, S. Wilson-Fletcher, D. E. Logan, and R. Bulla, Generalized wilson chain for solving multichannel quantum impurity problems, Physical Review B **89**, 121105 (2014); K. M. Stadler, A. K. Mitchell, J. von Delft, and A. Weichselbaum, Interleaved numerical renormalization group as an efficient multiband impurity solver, **93**, 235101 (2016).
[49] Alternative optimization methods in computational physics have employed exact gradients from automati-

cally differentiable algorithms, see e.g. Refs. [73–75].
[50] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Quantifying entanglement, Phys. Rev. Lett. **78**, 2275 (1997).
[51] H. J. Kappen, Learning quantum models from quantum or classical data, Journal of Physics A: Mathematical and Theoretical **53**, 214001 (2020).
[52] K. Haule, Quantum monte carlo impurity solver for cluster dynamical mean-field theory and electronic structure calculations with adjustable cluster base, Physical Review B **75**, 155113 (2007).
[53] J. Li, M. Wallerberger, and E. Gull, Diagrammatic monte carlo method for impurity models with general interactions and hybridizations, arXiv preprint arXiv:2004.00721 (2020).
[54] J. Duchi, E. Hazan, and Y. Singer, Adaptive subgradient methods for online learning and stochastic optimization., Journal of machine learning research **12** (2011).
[55] J. Linderberg and Y. öhrn, Derivation and analysis of the pariser–parr–pople model, The Journal of Chemical Physics **49**, 716 (1968).
[56] B. P. Dolan, A geometrical interpretation of renormalization group flow, International Journal of Modern Physics A **9**, 1261 (1994).
[57] P. W. Anderson, Localized magnetic states in metals, Physical Review **124**, 41 (1961).
[58] F. D. M. Haldane, Theory of the atomic limit of the anderson model. i. perturbation expansions re-examined, Journal of Physics C: Solid State Physics **11**, 5015 (1978).
[59] P. W. Anderson, A poor man s derivation of scaling laws for the kondo problem, Journal of Physics C: Solid State Physics **3**, 2436 (1970).
[60] R. Bulla, T. A. Costi, and T. Pruschke, Numerical renormalization group method for quantum impurity systems, Rev. Mod. Phys. **80**, 395 (2008).
[61] L. Che, C. Wei, Y. Huang, D. Zhao, S. Xue, X. Nie, J. Li, D. Lu, and T. Xin, Learning quantum hamiltonians from single-qubit measurements, Physical Review Research **3**, 023246 (2021).
[62] W. Foulkes, L. Mitas, R. Needs, and G. Rajagopal, Quantum monte carlo simulations of solids, Reviews of Modern Physics **73**, 33 (2001).
[63] R. J. Bartlett and M. Musiał, Coupled-cluster theory in quantum chemistry, Reviews of Modern Physics **79**, 291 (2007).
[64] G. Sugiyama and S. Koonin, Auxiliary field monte-carlo for quantum many-body ground states, Annals of Physics **168**, 1 (1986).
[65] M. Motta and S. Zhang, Ab initio computations of molecular systems by the auxiliary-field quantum monte carlo method, Wiley Interdisciplinary Reviews: Computational Molecular Science **8**, e1364 (2018).
[66] D. A. Mazziotti, Two-electron reduced density matrix as the basic variable in many-electron quantum chemistry and physics, Chemical reviews **112**, 244 (2012).
[67] E. Gull, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner, Continuous-time monte carlo methods for quantum impurity models, Rev. Mod. Phys. **83**, 349 (2011).
[68] A. N. Rubtsov, V. V. Savkin, and A. I. Lichtenstein, Continuous-time quantum monte carlo method for fermions, Phys. Rev. B **72**, 035122 (2005).
[69] P. Werner and A. J. Millis, Hybridization expansion impurity solver: General formulation and application to kondo lattice and two-orbital models, Phys. Rev. B **74**, 155107 (2006).
[70] K. G. Wilson, The renormalization group: Critical phenomena and the kondo problem, Rev. Mod. Phys. **47**, 773 (1975).
[71] F. B. Anders and A. Schiller, Real-time dynamics in quantum-impurity systems: A time-dependent numerical renormalization-group approach, Phys. Rev. Lett. **95**, 196801 (2005).
[72] A. Weichselbaum and J. von Delft, Sum-rule conserving spectral functions from the numerical renormalization group, Phys. Rev. Lett. **99**, 076402 (2007).
[73] H. Xie, J.-G. Liu, and L. Wang, Automatic differentiation of dominant eigensolver and its applications in quantum physics, Physical Review B **101**, 245139 (2020).
[74] L. Coopmans, D. Luo, G. Kells, B. K. Clark, and J. Carrasquilla, Protocol discovery for the quantum control of majoranas by differential programming and natural evolution strategies, arXiv preprint arXiv:2008.09128 (2020).
[75] I. Luchnikov, M. Krechetov, and S. Filippov, Riemannian optimization and automatic differentiation for complex quantum architectures, arXiv preprint arXiv:2007.01287 (2020).