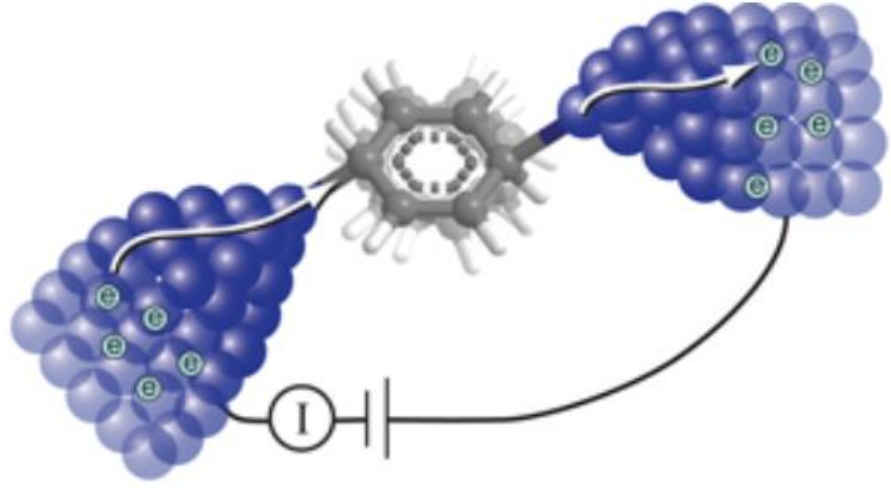
Generative Model Learning for molecular electronics



Andrew Mitchell

with Jonas Rigo and Sudeshna Sen

University College Dublin

IRISH RESEARCH COUNCIL
An Chomhairle um Thaighde in Éirinn

PHYSICAL REVIEW B **101**, 241105(R) (2020)

Rapid Communications

Machine learning effective models for quantum systems

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

(Received 1 November 2019; revised manuscript received 18 May 2020; accepted 26 May 2020; published 5 June 2020)

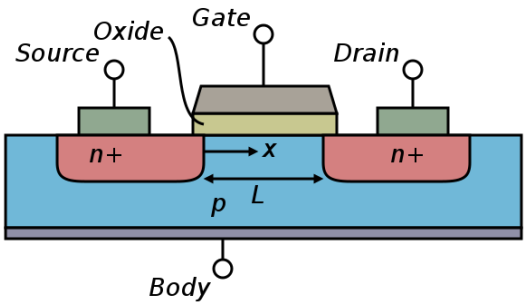
Generative Model Learning For quantum impurity systems

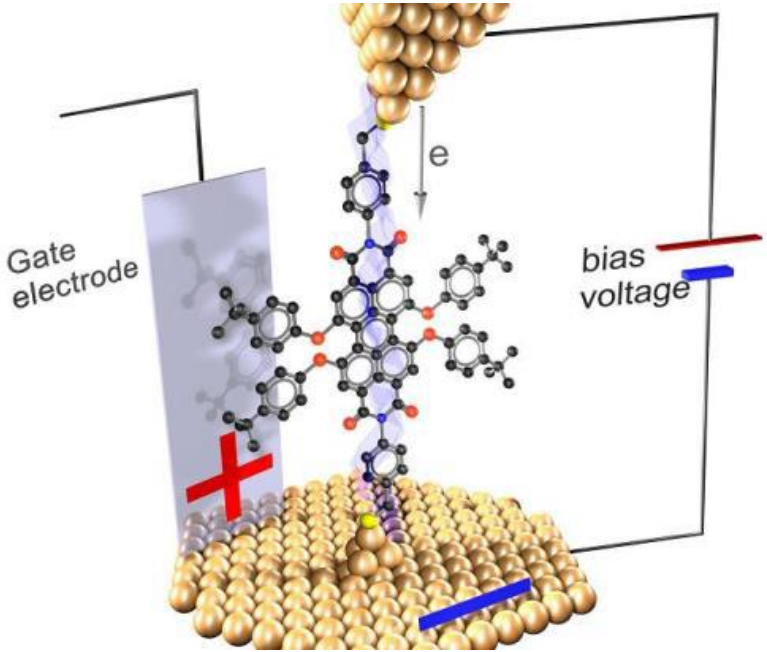
Jonas B. Rigo* and Andrew K. Mitchell†
School of Physics, University College Dublin, Belfield, Dublin 4, Ireland

The use of single-molecule transistors in nanoelectronics devices requires a deep understanding of the generalized 'quantum impurity' models describing them. Microscopic models comprise molecular orbital complexity and strong electron interactions while also treating explicitly conduction electrons in the external circuit. No single theoretical method can treat the low-temperature physics of such systems exactly. To overcome this problem, we use a generative machine learning approach to formulate effective models that are simple enough to be treated exactly by methods such as the numerical renormalization group, but still capture all observables of interest of the physical system.

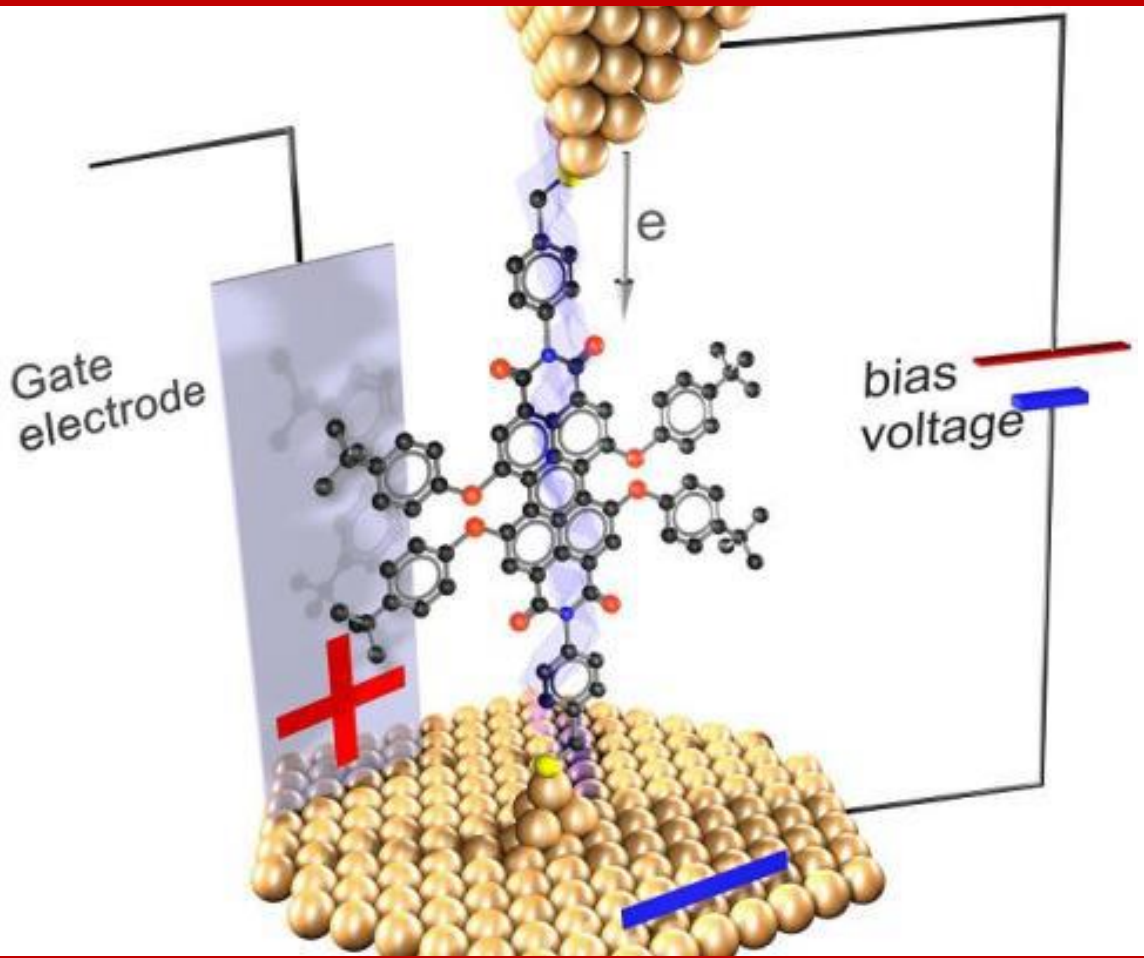
Motivation: molecular electronics

Extreme miniaturization // Quantum advantage

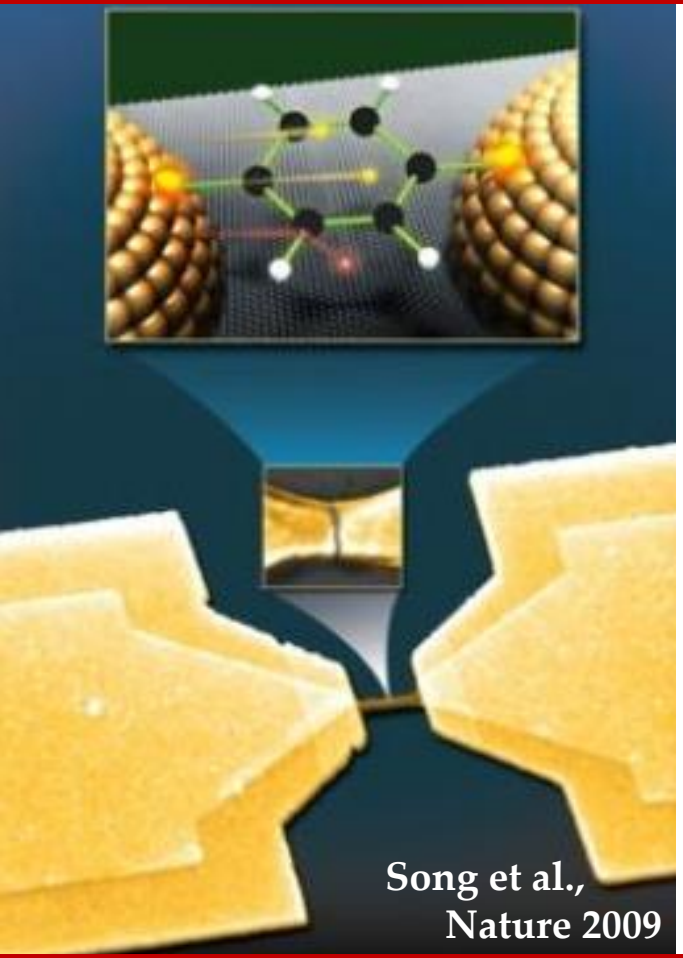




Single-molecule transistors



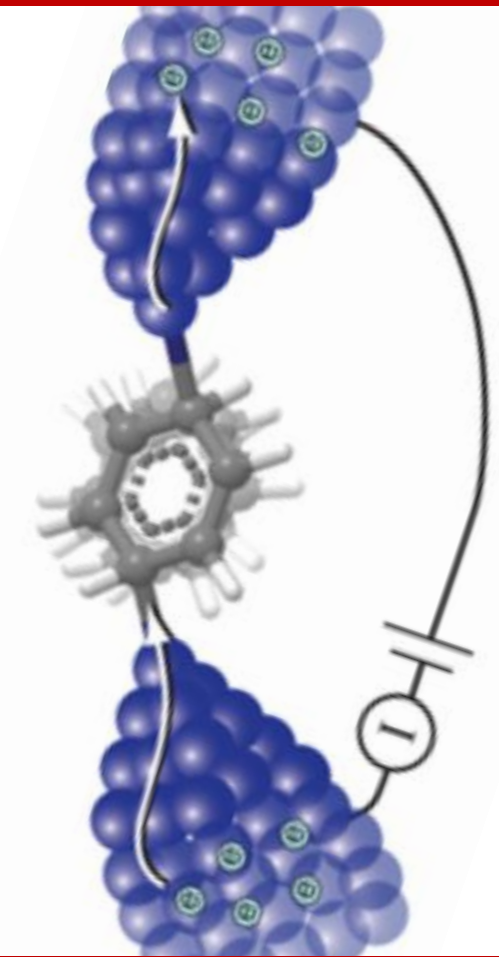
Modelling and simulation



Retain only low energy orbitals

Retain effective spin local moment

Standard methodology



Ab initio methods to treat molecular junction

↓

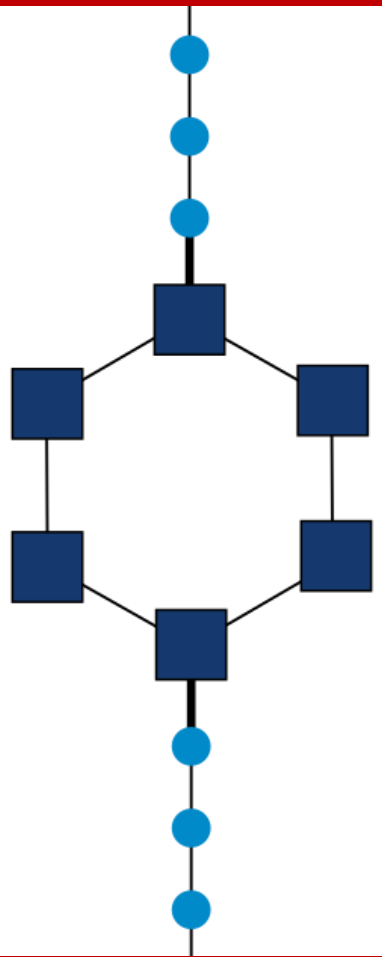
Fit to extended Hubbard model of active orbitals

↓

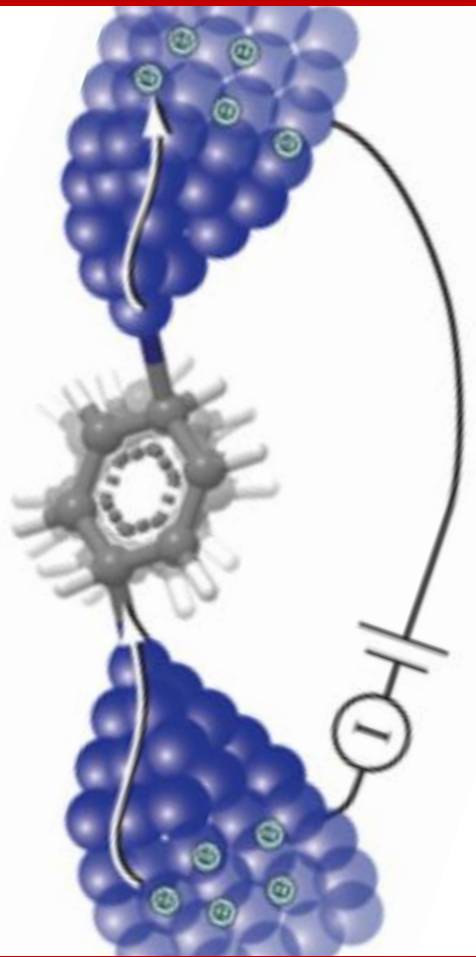
Map to quantum impurity model using Perturbation Theory

↓

Solve using quantum many-body computational techniques



Example: benzene transistor

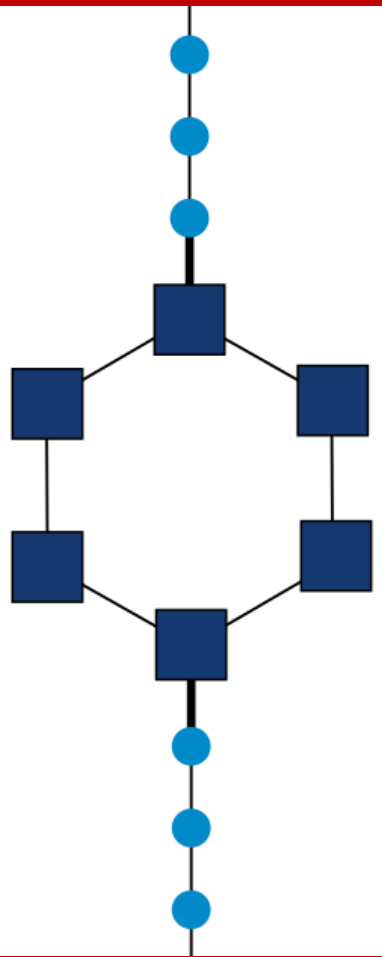


$$H_{\rm SMJ} = H_{\rm leads} + H_{\rm mol} + H_{\rm hyb}$$

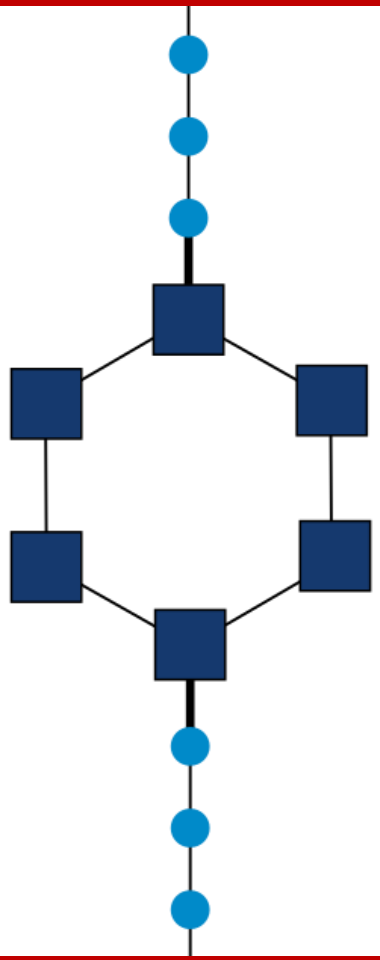
$$H_{\rm leads} = \sum_{\alpha,k,\sigma} \epsilon_k c^\dagger_{\alpha k\sigma} c_{\alpha k\sigma}$$

$$H_{\rm mol} = \sum_{i,j,\sigma} t_{ij} d^\dagger_{i\sigma} d_{j\sigma} + \sum_i U_i \hat{n}_{i\uparrow}\hat{n}_{i\downarrow} + \sum_{i\neq j} U'_{ij}\hat{n}_i\hat{n}_j$$

$$H_{\rm hyb} = \sum_{\alpha,i,\sigma} \left(V_i^\alpha d^\dagger_{i\sigma} c_{\alpha\sigma} + {\rm H.c.} \right)$$



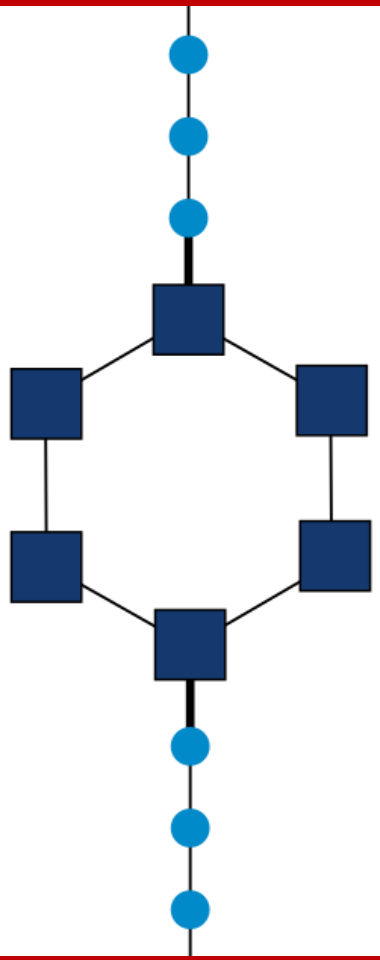
Example: benzene transistor

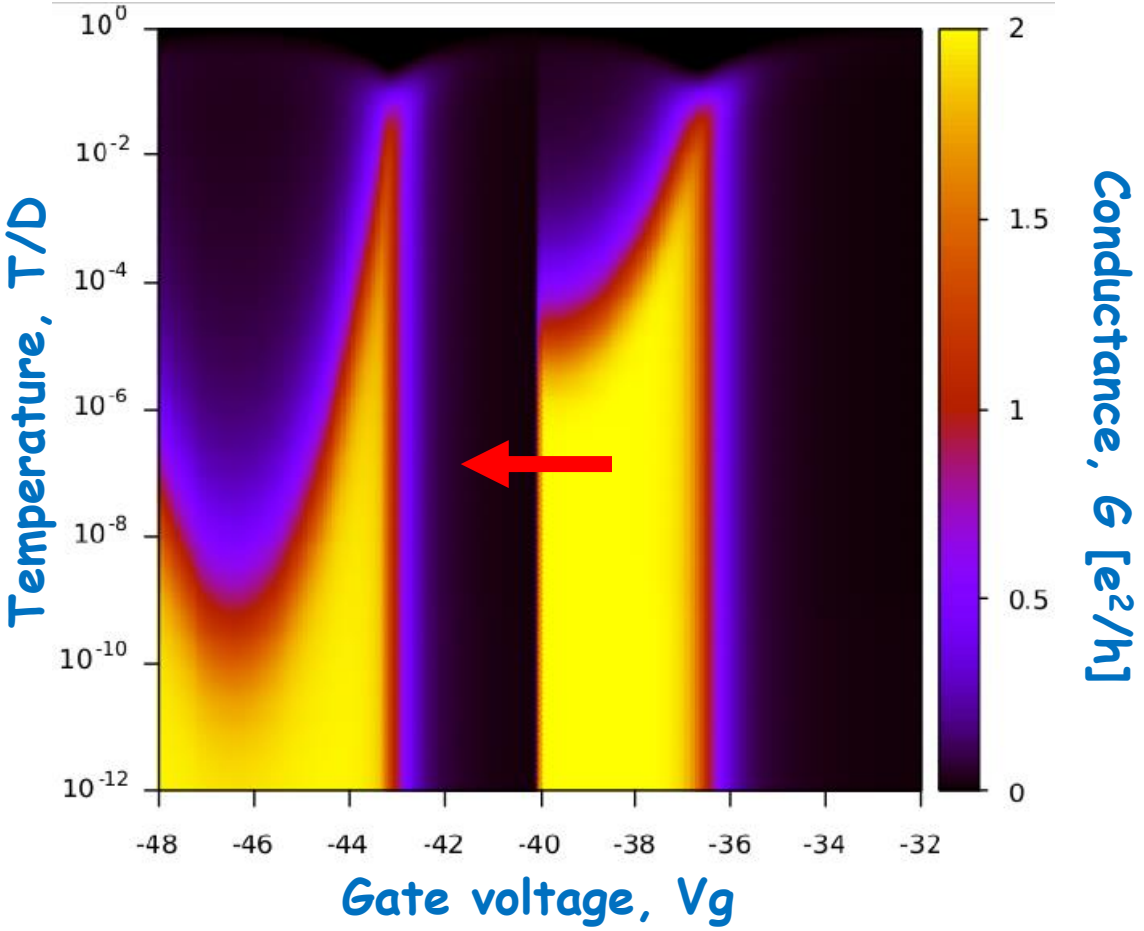


$$H_{\rm eff} = \hat{P} H_{\rm hyb} \hat{Q} \left(E - \hat{Q} H \hat{Q} \right)^{-1} \hat{Q} H_{\rm hyb} \hat{P}$$

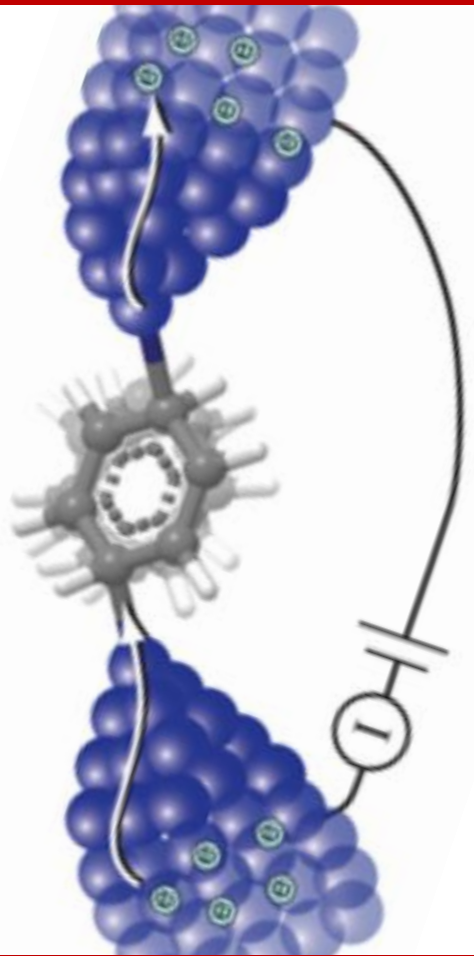
$$H_{\rm eff} = \sum_{\alpha} \hat{\vec{\mathbf{S}}}^{M} \cdot \hat{\vec{\mathbf{s}}}_{\alpha} \left[\hat{\mathcal{T}}^{0} \tilde{\mathbf{J}}^{0} + \hat{\mathcal{T}}^{z} \tilde{\mathbf{J}}^{z} \right]$$

Example: benzene transistor



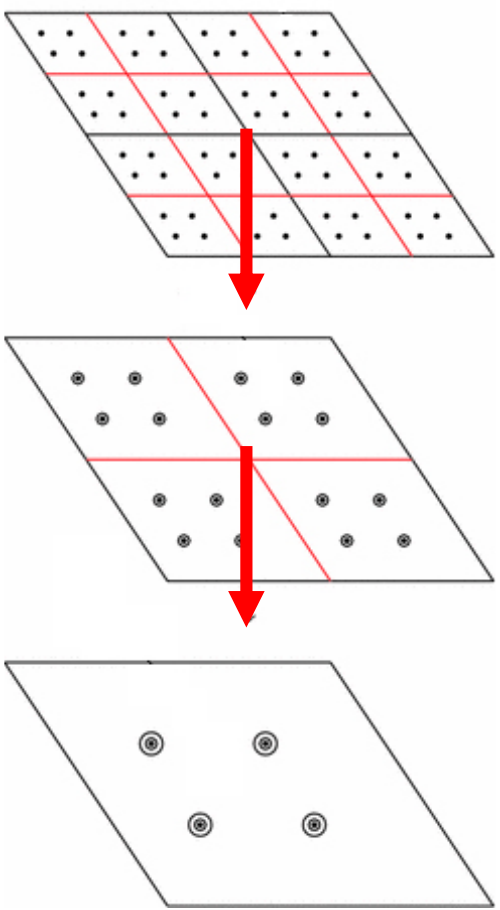


Problems with traditional approach



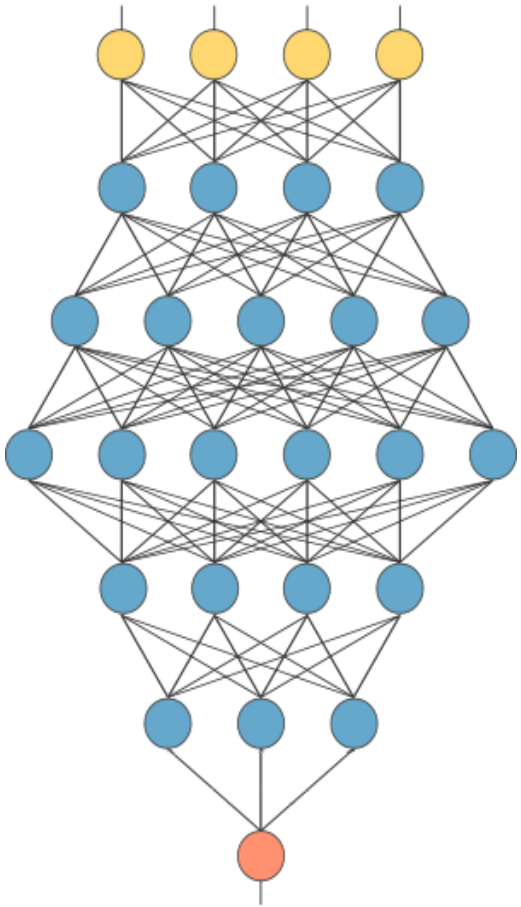
- ✗ Indirect: two-stage mapping assumes separation of scales
- ✗ Perturbative: assumes large-U limit and neglects renormalization
- ✗ Complexity: need to "solve" the molecule exactly
- ✗ Observables: physical quantities might not match

Machine learning effective models



Goal:

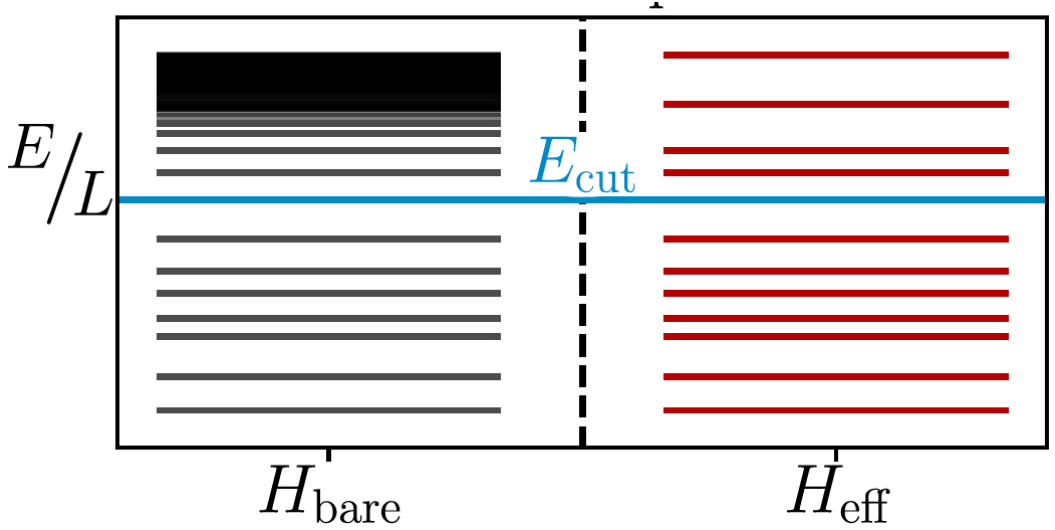
find the simplest model that captures physical phenomena of interest using tools from ML



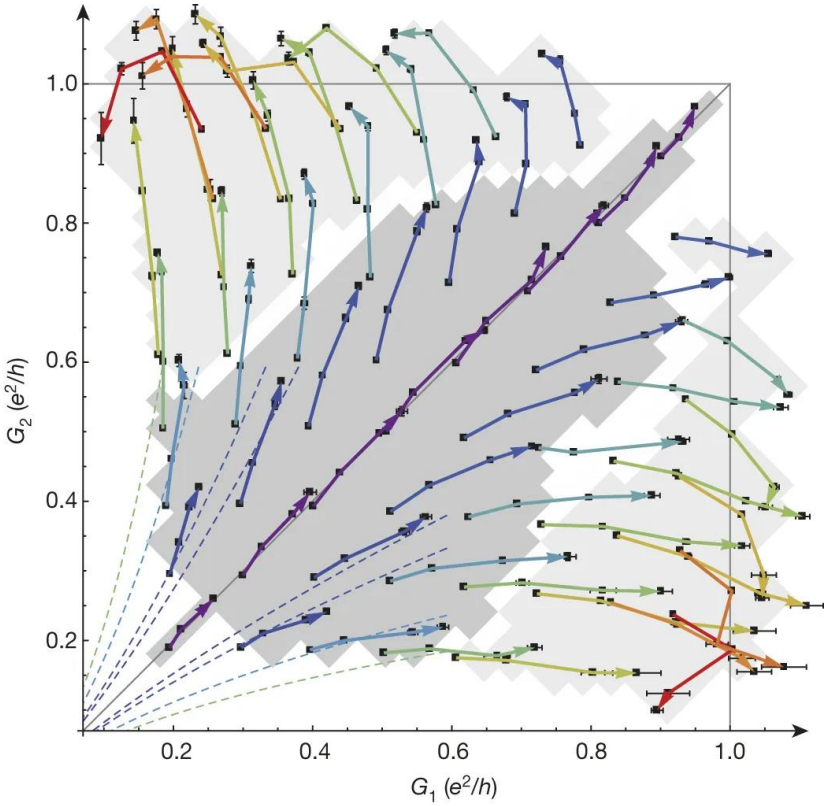
What is a good effective model?

✓ Simpler description

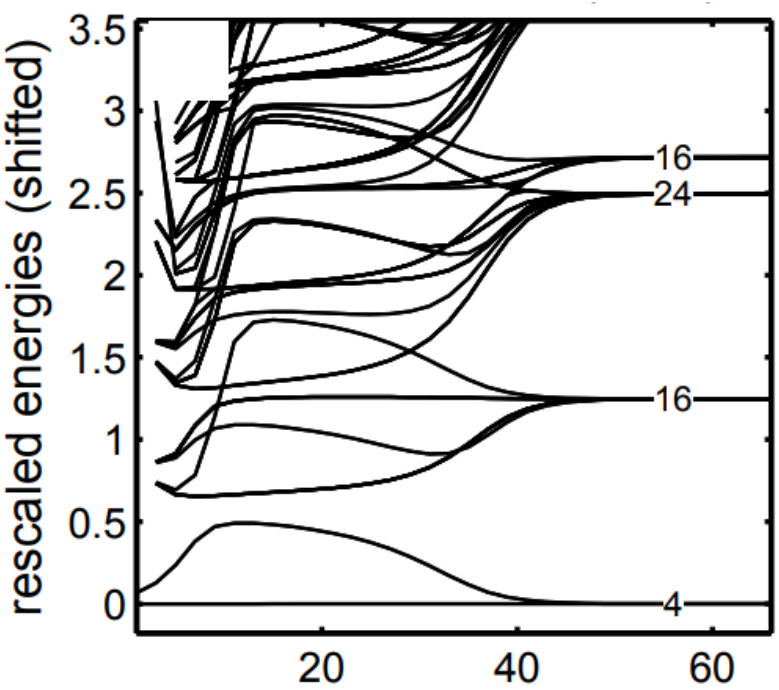
✓ Describes low-energy properties



RG-derivable models



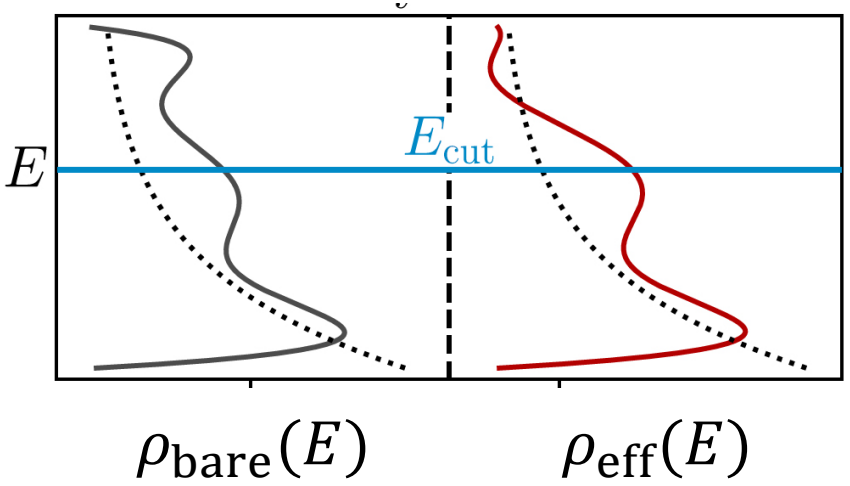
Iftikhar et al, Nature 526, 233 (2015)

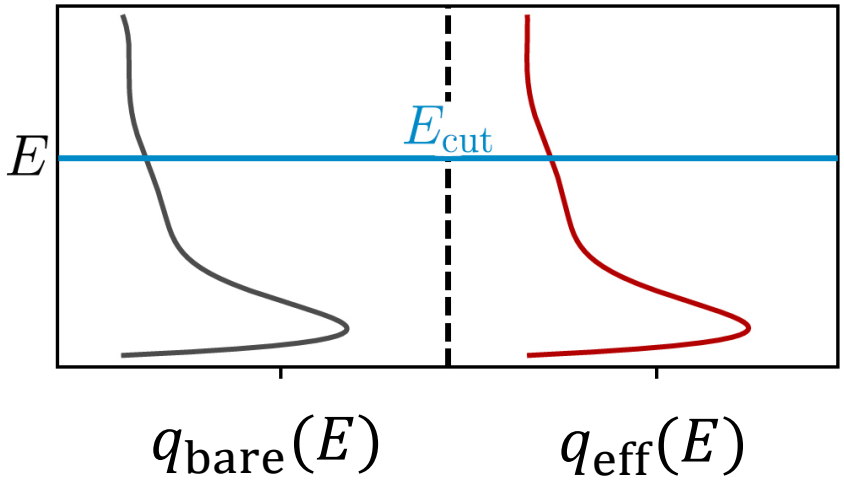


Phys. Rev. B 84, 125130 (2011)

Density Matrix Spectrum

✓ Low-temperature thermodynamics and emergent energy scales guaranteed by matching the spectrum of density matrix: $q(E) = \exp[-\beta E]\,\rho(E)$





Partition Function

Optimized effective models therefore have the same partition function as the bare model

$$Z = \int dE\, q(E)$$

✓ Optimize an effective model on the level of the partition function such that: $Z_{eff} = Z_{bare}$

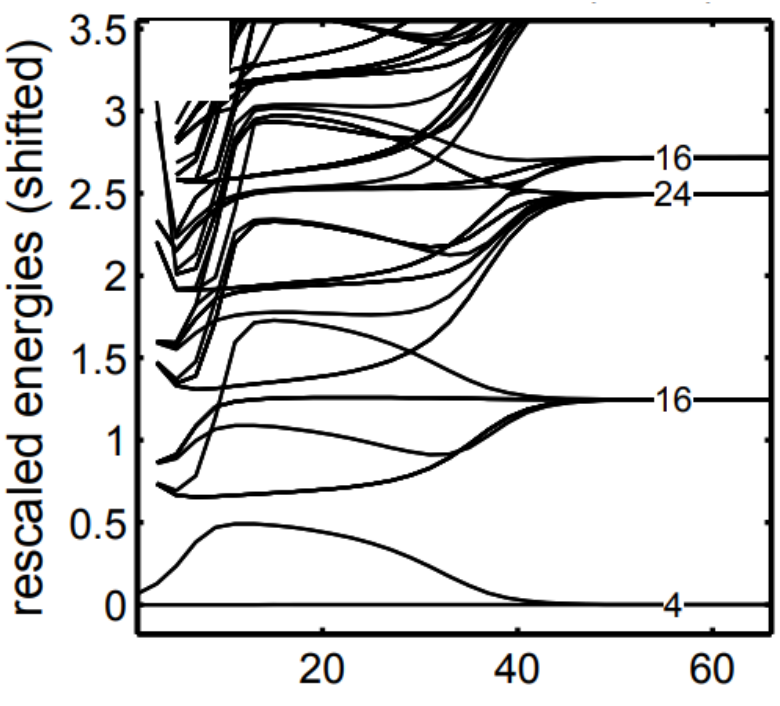
BUT: Two totally different, arbitrary and unrelated, models may "accidentally" have the same Z...

RG-derivable Minimal Models

Partition function is preserved under RG flow!

→ Consider only effective models that are "RG-derivable" from the microscopic model (use prior knowledge to limit search)

→ Consider "Minimal Models" involving RG relevant and marginal terms only (ensures search convexity)



Partition function optimization

Effective model:

$$\hat{H}_{\text{eff}} = \sum_i \theta_i \, \hat{h}_i$$

Optimization loss function:

$$L_Z = [\log(Z_{\text{eff}}) - \log(Z_{\text{bare}})]^2$$

Gradient:

$$\frac{\partial L_Z}{\partial \theta_i} \sim [\log(Z_{\text{eff}}) - \log(Z_{\text{bare}})] \times \langle \hat{h}_i \rangle_{\text{eff}}$$

expectation value of operators appearing in effective Hamiltonian

Anderson to Kondo mapping

Quantum impurity model: Anderson

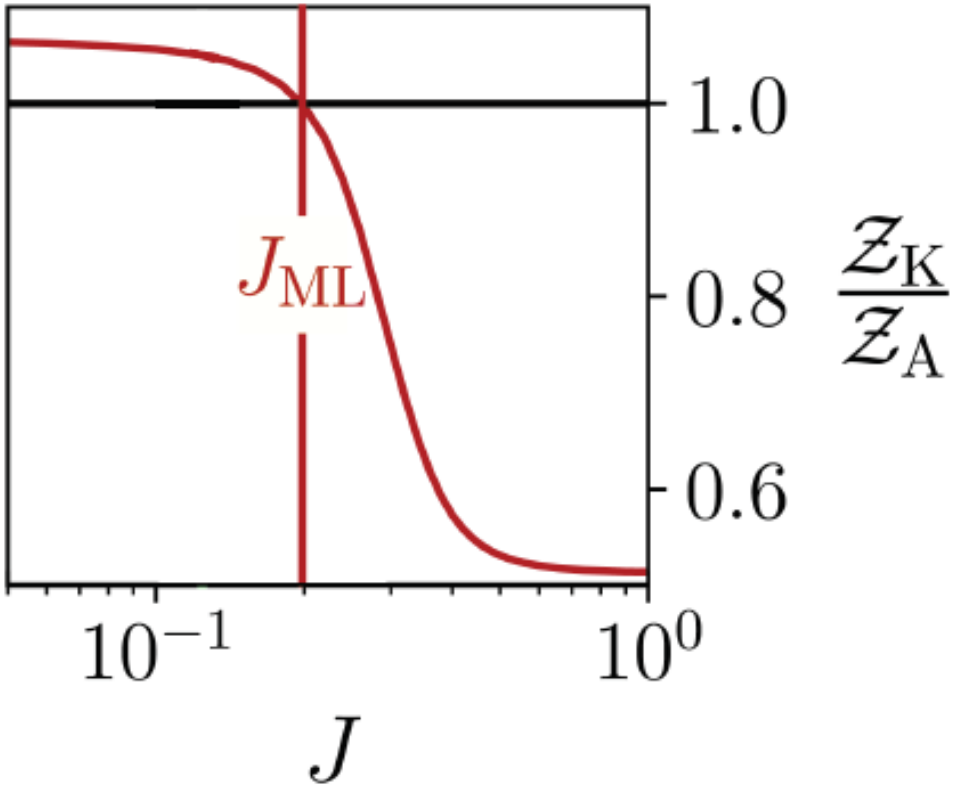
$$\hat{H}_{\text{bare}} = \hat{H}_{\text{bath}} + \epsilon_d \left(d_\uparrow^\dagger d_\uparrow + d_\downarrow^\dagger d_\downarrow \right) + U d_\uparrow^\dagger d_\uparrow \, d_\downarrow^\dagger d_\downarrow + \sum_{k,\sigma} V_k \left(d_\sigma^\dagger c_{k\sigma} + c_{k\sigma}^\dagger d_\sigma \right)$$

Minimal effective model: Kondo

$$\hat{H}_{\text{eff}} = \hat{H}_{\text{bath}} + J \, \hat{\boldsymbol{S}}_d \cdot \hat{\boldsymbol{S}}_c$$

Optimize Z_{eff}=Z_{bare} to find $J \equiv J(\epsilon_d, U, V)$

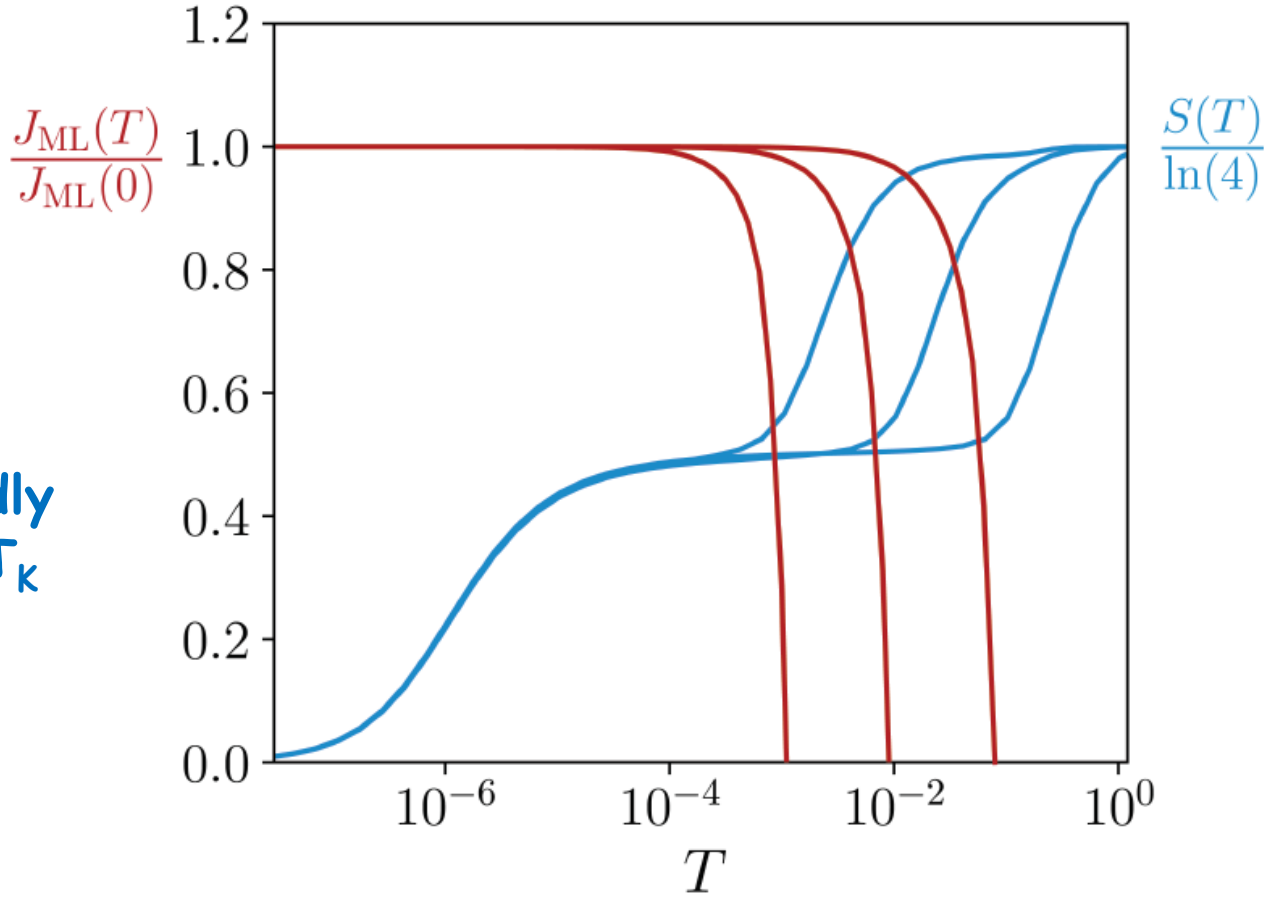
Anderson to Kondo mapping



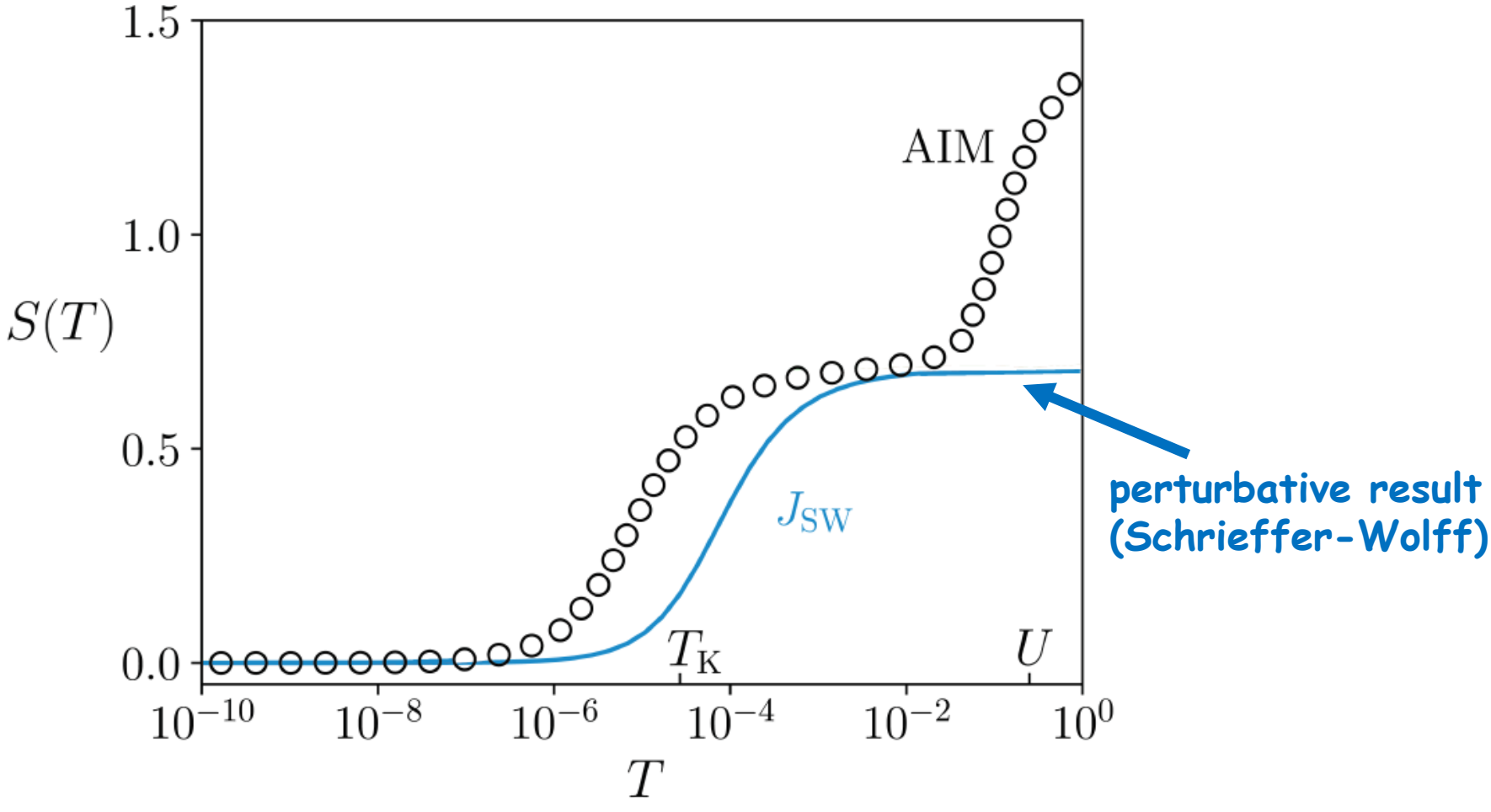
Keeping only RG relevant terms in the effective model ensures a single solution

Anderson to Kondo mapping

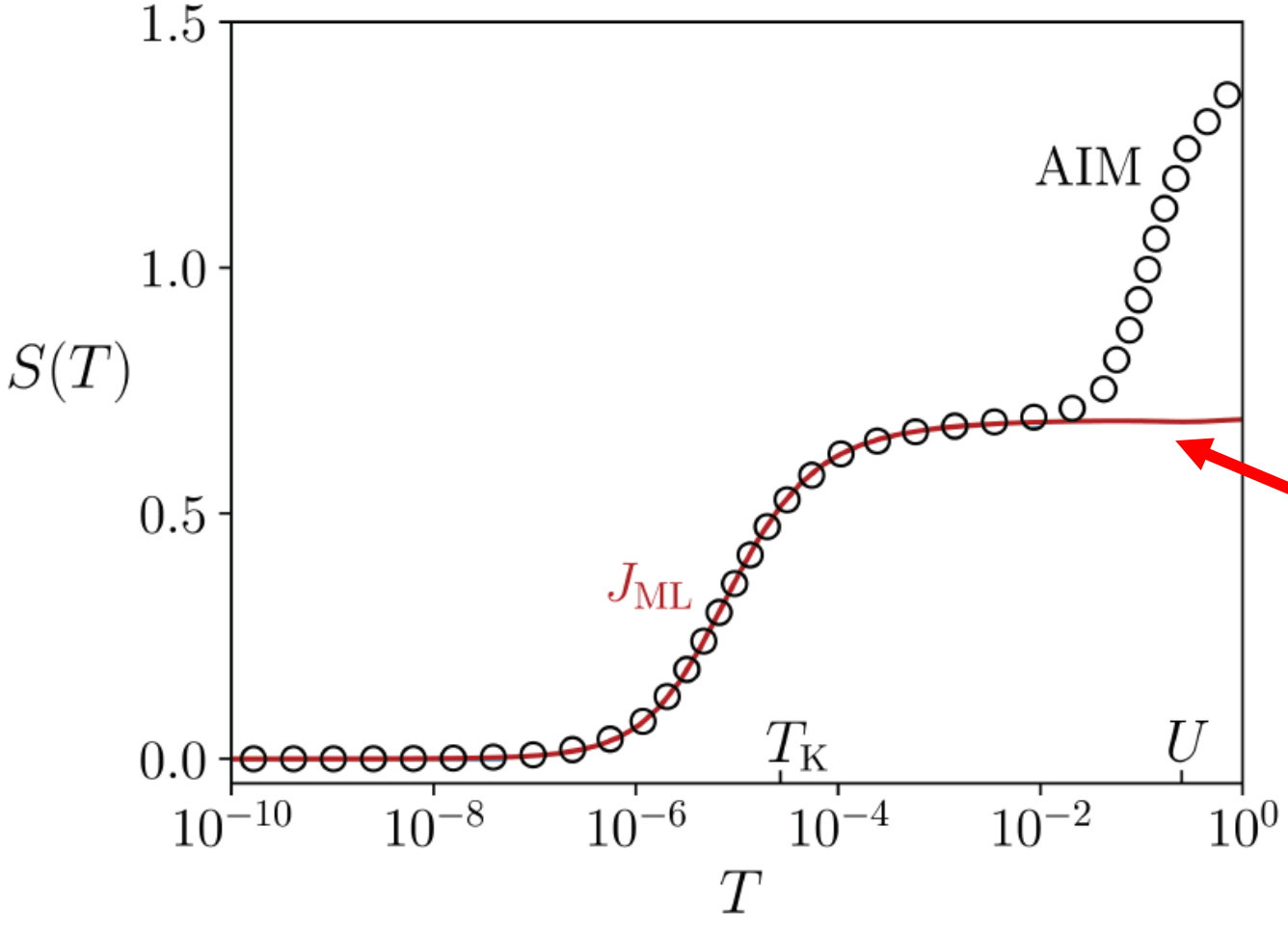
Optimization of effective model can be done at high temperatures (converges rapidly for T<U). Do NOT need $T << T_K$



Anderson to Kondo mapping



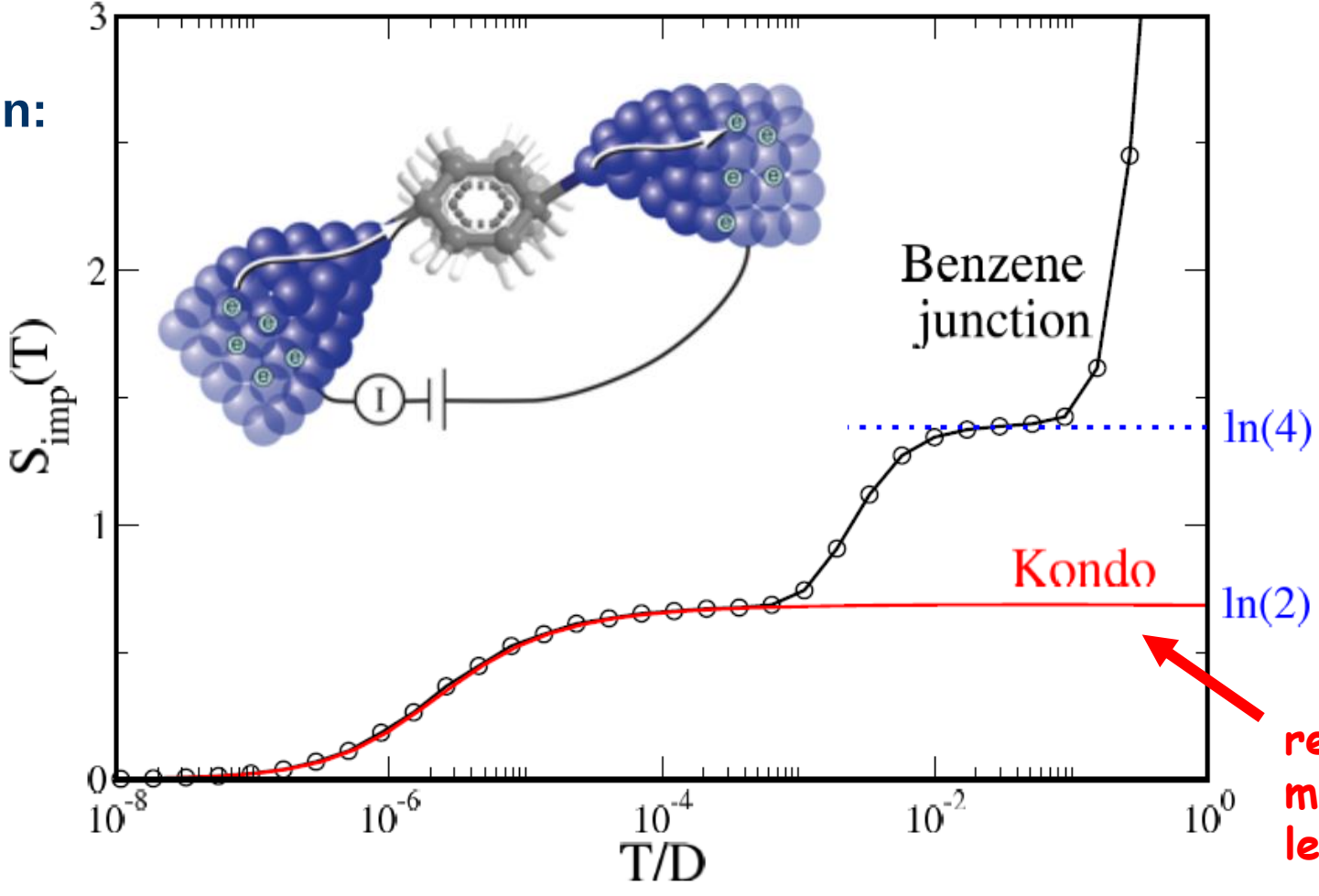
Anderson to Kondo mapping



result from model machine learning

Effective model for Benzene junction

See talk by Sudeshna Sen: session R21



Limitations of Z-optimization

Must use prior knowledge of form of effective model

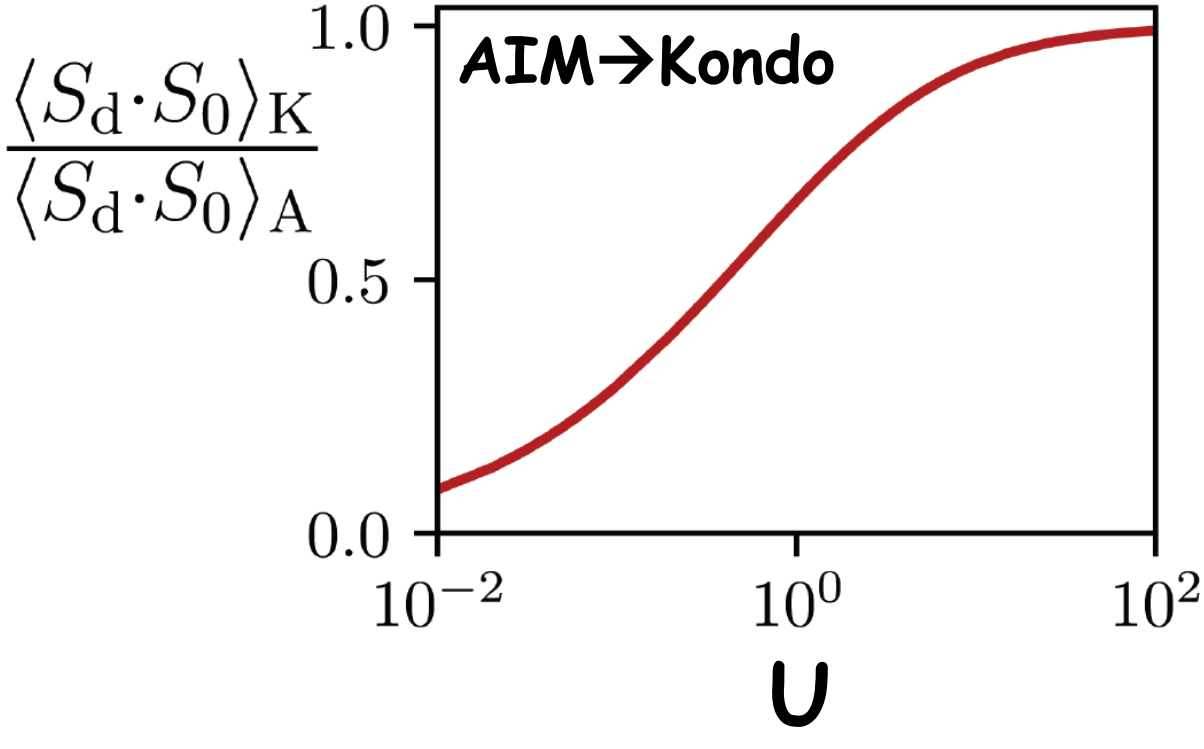
Only works for "minimal models"

Solutions may not be unique for more general models

Observable mismatch?!

Observables in minimal effective models MAY NOT AGREE with those of the bare model!

Must keep higher-order terms (RG marginal and RG irrelevant terms) to reproduce observables!



Generative machine learning

Compare density matrices

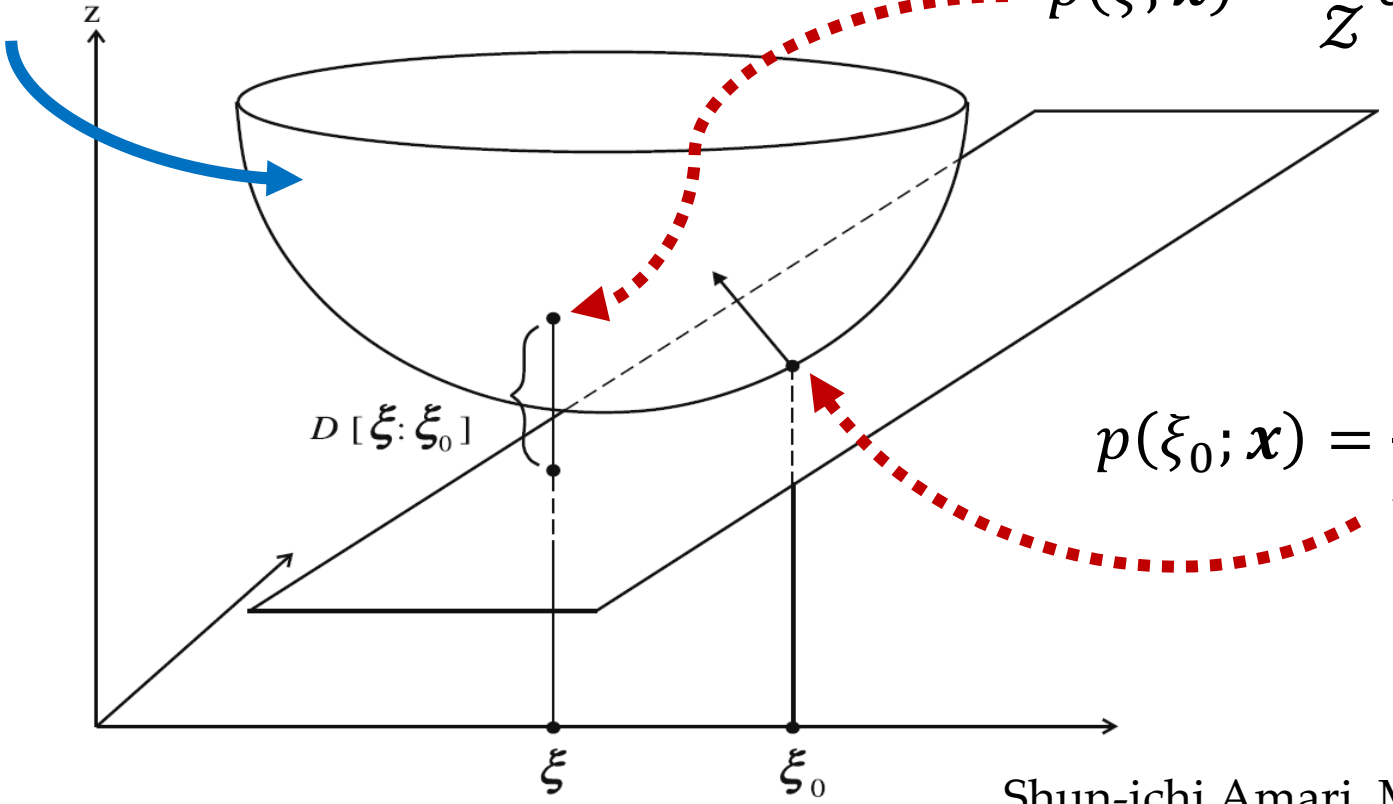
$$\hat{H}_{\text{bare}} \Rightarrow \hat{\rho}_{\text{bare}}(\boldsymbol{x}) = \frac{1}{\mathcal{Z}} e^{-\beta[\hat{H}_{\text{bath}}+\hat{H}_{\text{hyb}}+\hat{\boldsymbol{H}}_{\boldsymbol{imp}}]}$$

$$\hat{H}_{\text{eff}}(\{\theta\}) = \sum_i \theta_i \hat{h}_i \Rightarrow \hat{\rho}_{\text{eff}}(\theta;\boldsymbol{x}) = \frac{1}{\mathcal{Z}} e^{-\beta[\hat{H}_{\text{bath}}+\hat{H}_{\text{hyb}}+\hat{\boldsymbol{H}}_{\boldsymbol{eff}}(\{\boldsymbol{\theta}\})]}$$

Generative machine learning

$$\varphi(\xi) = \int d\boldsymbol{x}\; p(\xi;\boldsymbol{x})\log[p(\xi;\boldsymbol{x})]$$

$$p(\xi;\boldsymbol{x}) = \frac{1}{\mathcal{Z}} e^{-\beta \hat{H}_{\text{eff}}(\{\xi\})}$$



$$p(\xi_0;\boldsymbol{x}) = \frac{1}{\mathcal{Z}} e^{-\beta \hat{H}_{\text{bare}}}$$

Shun-ichi Amari. Manifold, 2016

Generative machine learning

Optimize effective model parameters
by minimizing KL divergence using gradient descent

$$D_{KL}[\xi_0 : \xi] = \int d\boldsymbol{x}\, p(\xi_0; \boldsymbol{x}) \log \left[\frac{p(\xi_0; \boldsymbol{x})}{p(\xi; \boldsymbol{x})} \right]$$

Generative machine learning

Need to represent the thermal density matrix as a classical probability distribution: $P(x) = (\mathcal{Z}_{\rm bath}/\mathcal{Z})w(x)$

Expand the partition function:

$$\mathcal{Z} = \mathcal{Z}_{\rm bath} \int dx\ w(x)$$

$$w(x) = \det(\Delta^{(x)})\Lambda^{(x)} \exp[-\beta\langle \hat{H}_{\rm imp}\rangle_x]$$

K. Haule, PRB 75, 155113 (2007)

Generative machine learning

P(x) is an energy-based distribution!

Gradient descent:

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

$$P(x) = (\mathcal{Z}_{\text{bath}}/\mathcal{Z}) w(x)$$

$$\partial_i D_{KL}[P_{\text{bare}}, : P_{\text{eff}}] \simeq \langle \hat{Q}\hat{h}_i\hat{Q} \rangle_{\text{bare}} - \langle \hat{h}_i \rangle_{\text{eff}}$$

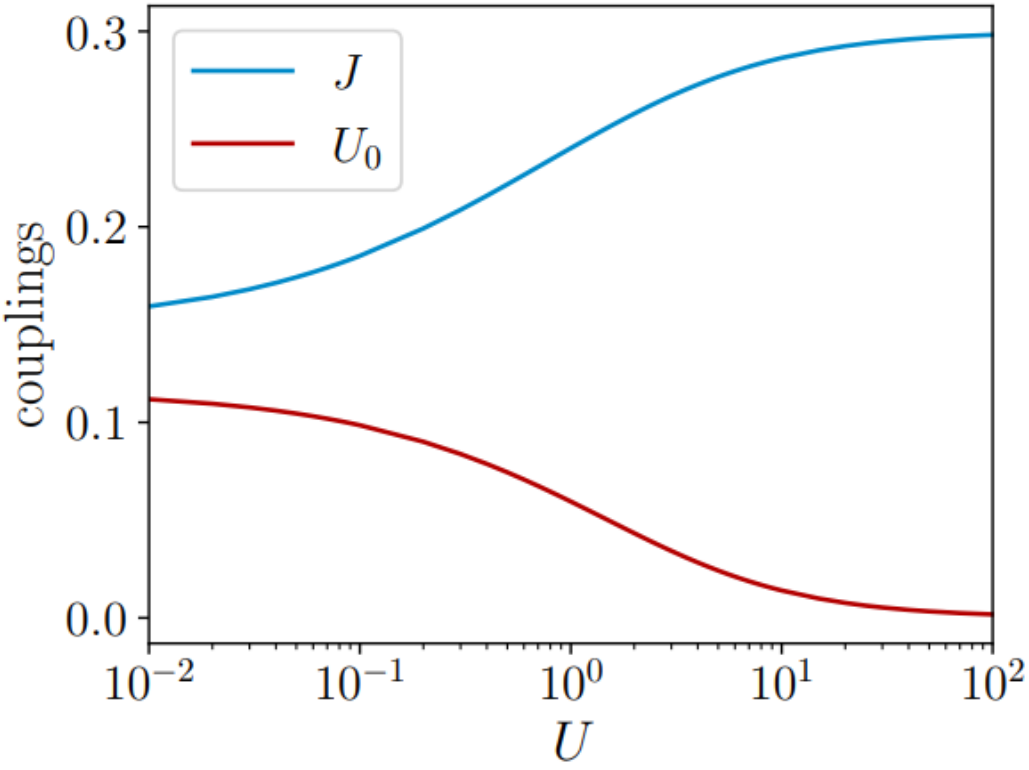
Observable matching optimization!

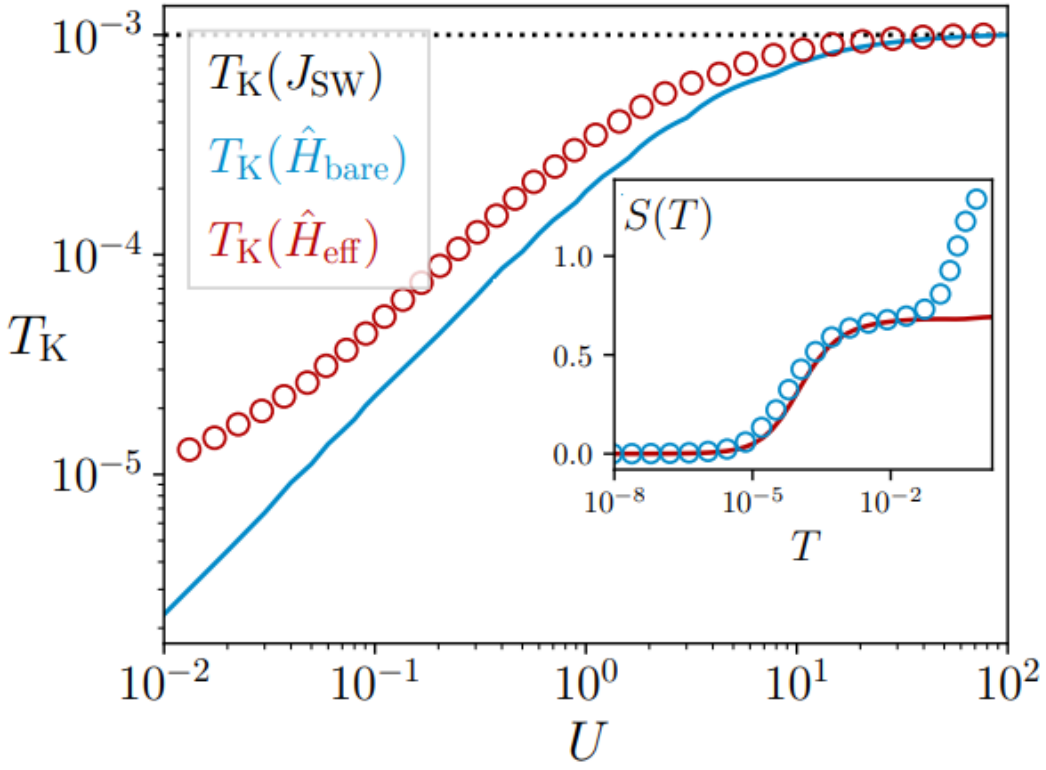
Anderson to Kondo mapping

$$\hat{H}_{\rm A} = \hat{H}_{\rm bath} + \tfrac{1}{2}U_d(\hat{n}_d - 1)^2 + V\sum_\sigma (\hat{d}^\dagger_\sigma \hat{c}_{0\sigma} + \hat{c}^\dagger_{0\sigma}\hat{d}_\sigma)$$

$$\hat{H}_{\rm eff} = \hat{H}_{\rm bath} + J\hat{\vec{S}}_d \cdot \hat{\vec{S}}_0 + \tfrac{1}{2}U_0\,(\hat{n}_0 - 1)^2$$

Anderson to Extended Kondo mapping





Conclusion and outlook

Complex microscopic models of single-molecule junctions can be mapped to simplified quantum impurity models using machine learning techniques

Classical energy-based distribution obtained by expanding the partition function

Minimizing the KL divergence gives a stringent condition on matching observables; strictly convex optimization!