

# Parametric Matrix Models

Dean Lee  
Facility for Rare Isotope Beams  
Michigan State University

Theory Seminar  
Technical University Darmstadt  
July 4, 2024

Patrick Cook, Danny Jammooa, Hjorth-Jensen, Da. Lee, D.L., arXiv:2401.11694  
will be updated on arXiv this weekend



MICHIGAN STATE  
UNIVERSITY



**NUCLEI**  
Nuclear Computational Low-Energy Initiative  
A SciDAC-5 Project

**JÜLICH**  
FORSCHUNGSZENTRUM

**OAK RIDGE**  
National Laboratory | LEADERSHIP COMPUTING FACILITY

## Outline

Eigenvector continuation

Parametric matrix models (PMMs)

Affine eigenvalue PMMs

Universal approximation theorem

Eigenvector continuation and PMMs

Applications to function interpolation

Applications to quantum computing

Applications to quantum many-body systems

Applications to general machine learning

Summary and outlook

## Eigenvector continuation

Consider a quantum Hamiltonian known as the Bose-Hubbard model that describes a system of identical bosons on a three-dimensional cubic lattice

$$H = -t \sum_{\langle \mathbf{n}', \mathbf{n} \rangle} a^\dagger(\mathbf{n}') a(\mathbf{n}) + \frac{U}{2} \sum_{\mathbf{n}} \rho(\mathbf{n}) [\rho(\mathbf{n}) - 1] - \mu \sum_{\mathbf{n}} \rho(\mathbf{n})$$
$$\rho(\mathbf{n}) = a^\dagger(\mathbf{n}) a(\mathbf{n})$$

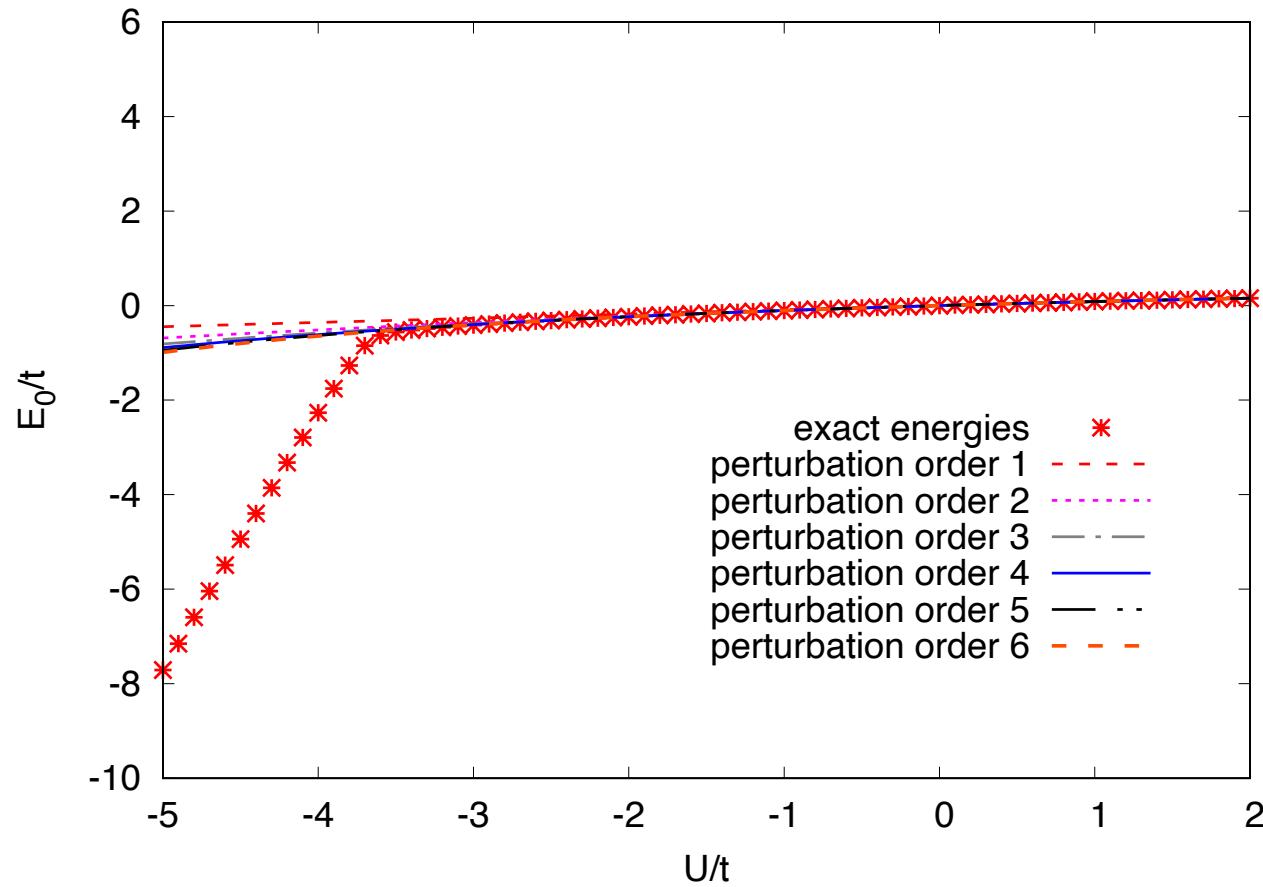
The parameter  $t$  controls the hopping of the bosons on the lattice, and  $U$  is the single-site pairwise interaction. We set the chemical potential to be

$$\mu = -6t$$

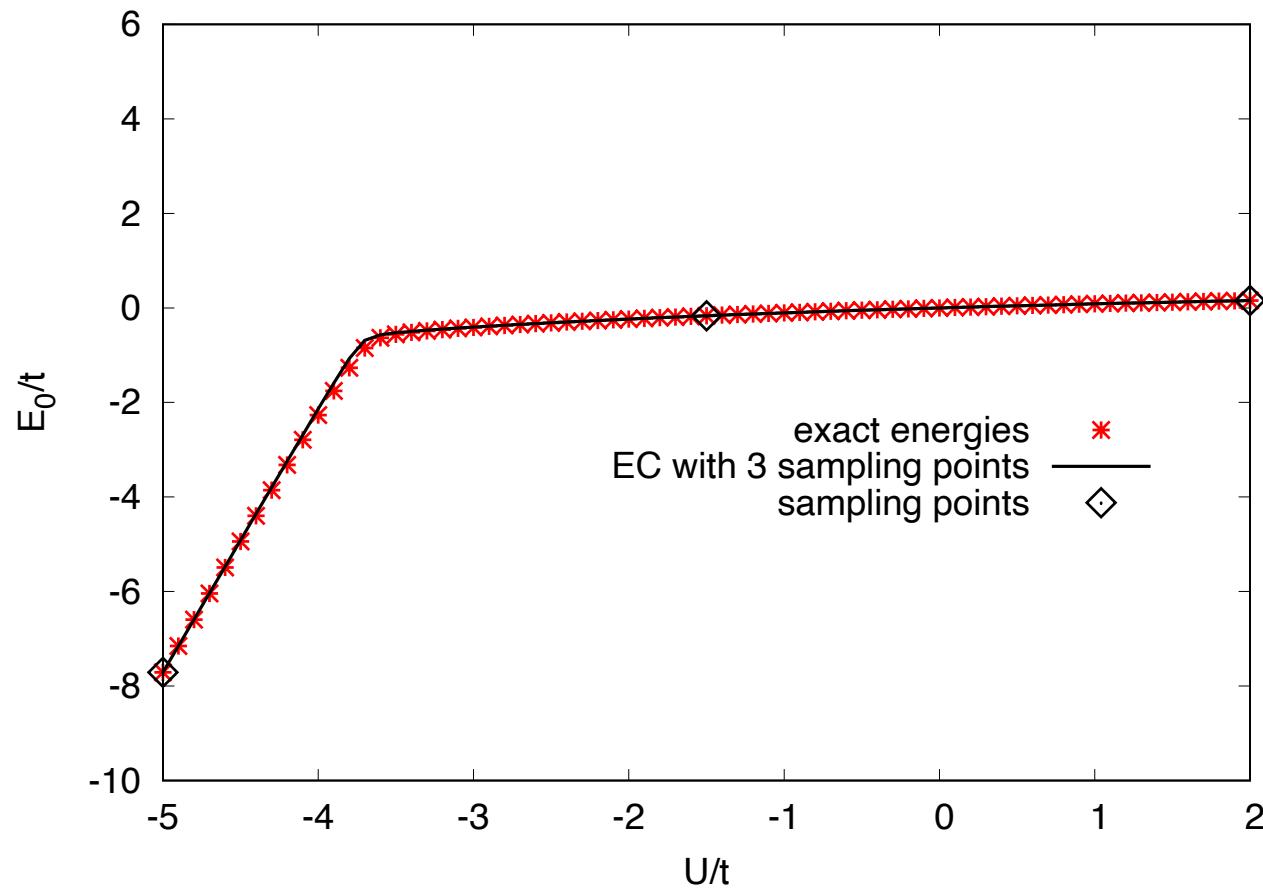
Frame, He, Ipsen, Da. Lee, D.L., Rrapaj, PRL 121 (2018) 032501

Duguet, Ekström, Furnstahl, König, D.L., arXiv:2310.19419, RMP in press

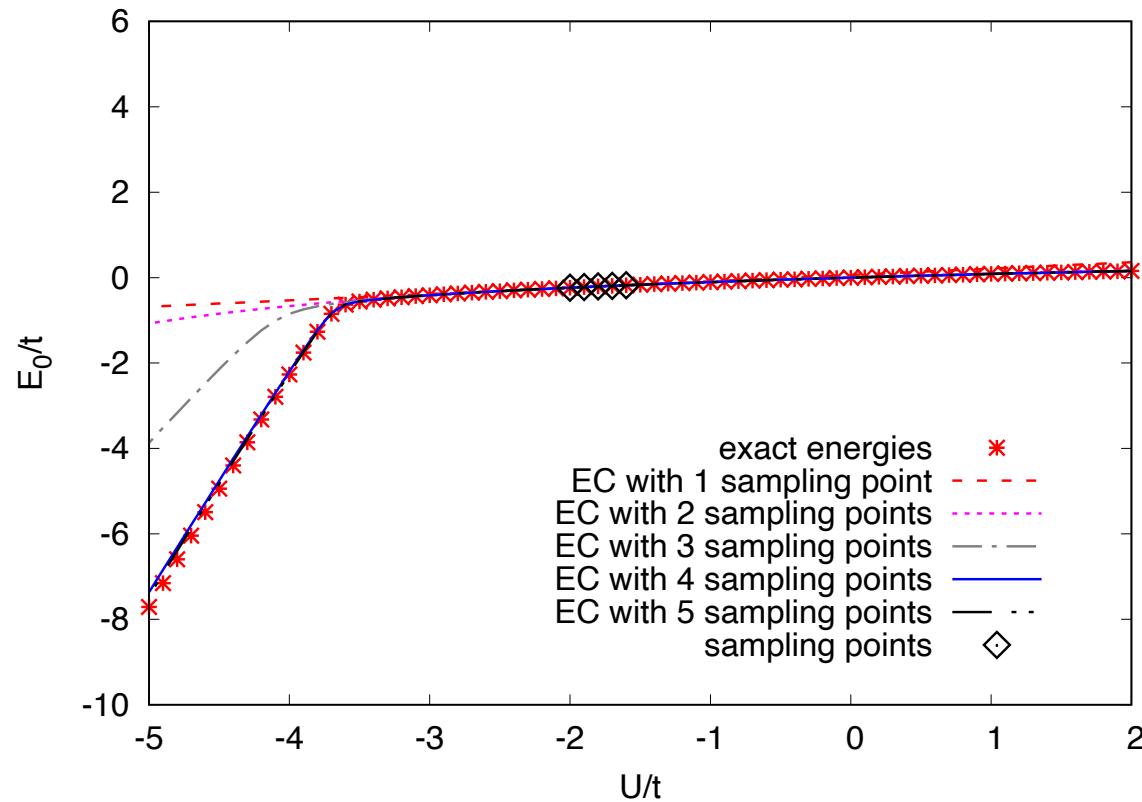
## Perturbation theory fails at strong attractive coupling



Restrict the linear space to the span of three vectors



We can “learn” the eigenvector trajectory in one region and perform eigenvector continuation to another region



# Improved many-body expansions from eigenvector continuation

P. Demol, T. Duguet, A. Ekström, M. Frosini, K. Hebeler, S. König, D. Lee, A. Schwenk, V. Somà, and A. Tichai  
Phys. Rev. C **101**, 041302(R) – Published 9 April 2020

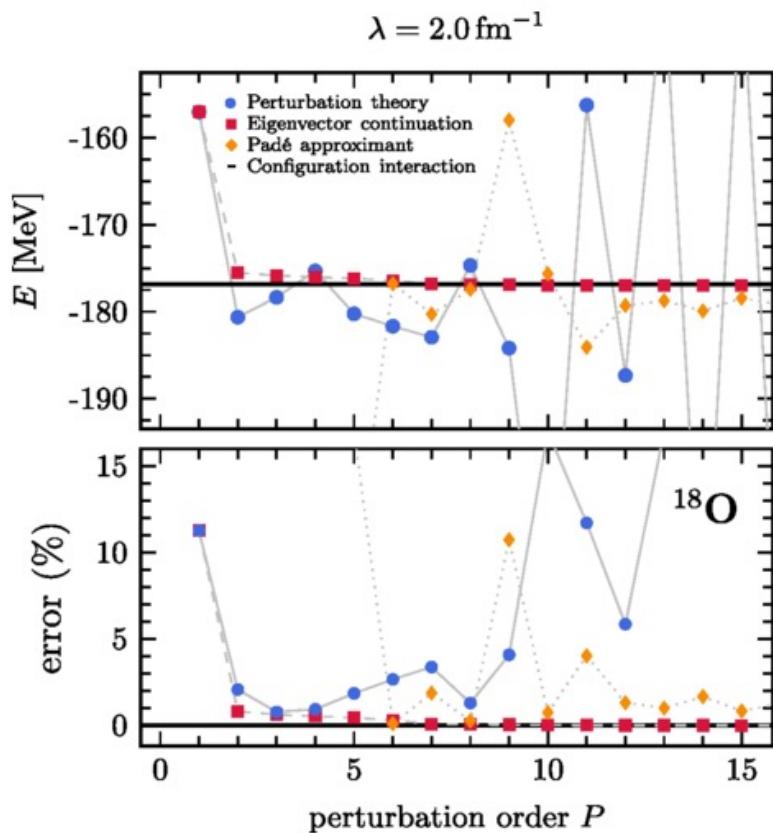


Figure 2

Ground-state energy of  $^{18}\text{O}$  for BMBPT (blue circles), BMBPT-based EC (red squares), and BMBPT-based Padé (yellow diamonds) as a function of the perturbative order  $P$  against exact CI diagonalization (full line) for  $\lambda = 2.0 \text{ fm}^{-1}$ . Top panel: absolute energies. Bottom panel: relative error to the CI result.



# Eigenvector continuation as an efficient and accurate emulator for uncertainty quantification

S. König<sup>a b c</sup>  , A. Ekström<sup>d</sup>  , K. Hebeler<sup>a b</sup>  , D. Lee<sup>e</sup>  , A. Schwenk<sup>a b f</sup> 

2

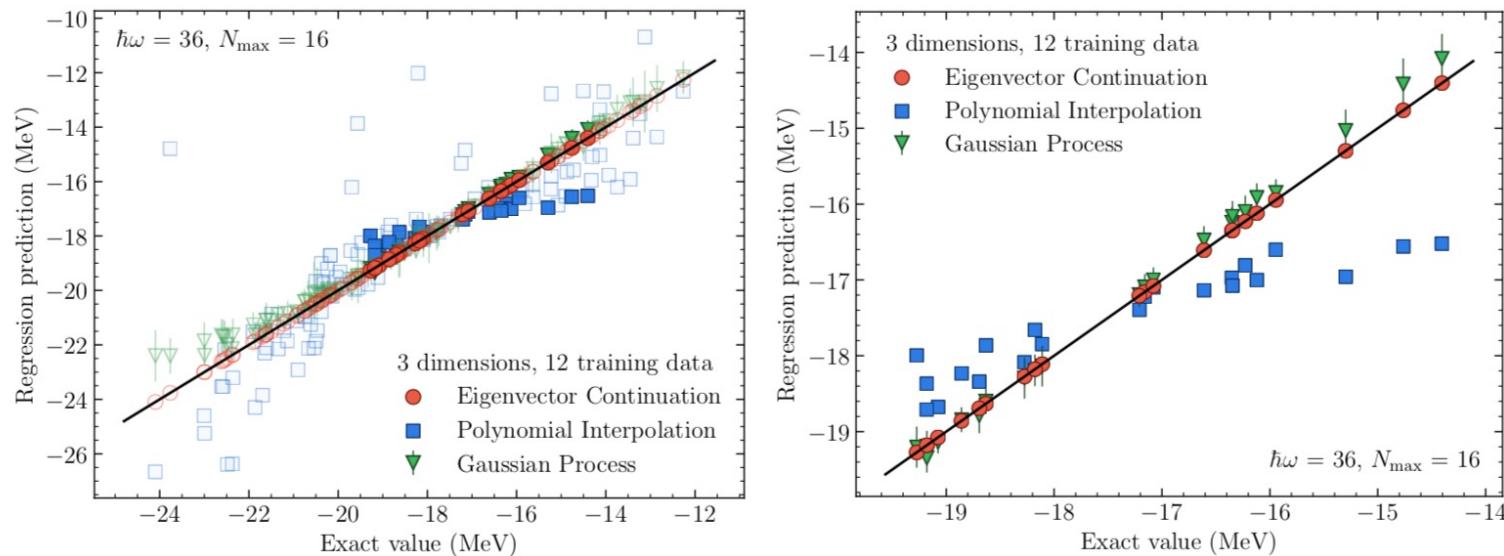
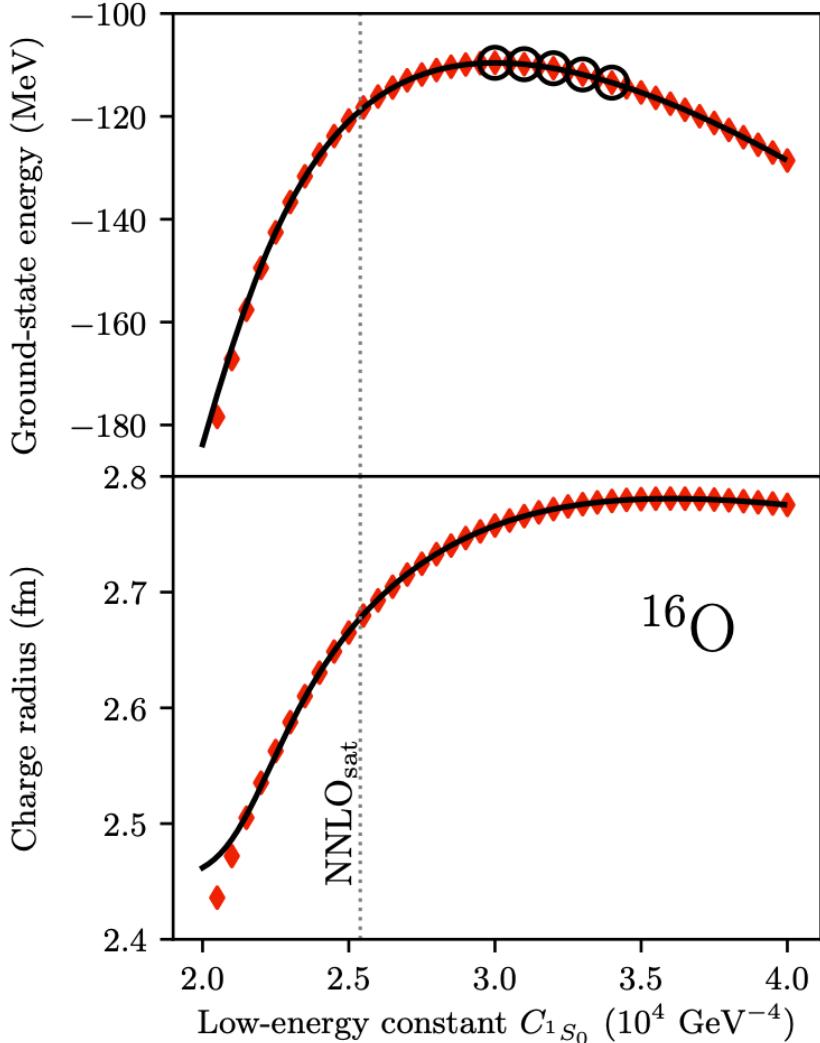


Figure 1. Comparison of different emulators for the  ${}^4\text{He}$  ground-state energy using 12 training data points to explore a space where three LECs are varied. The left panel includes samples for both interpolation (solid symbols) and extrapolation (semi-transparent symbols). See main text on how these are defined. The right panel shows the same data restricted to interpolation samples (note the smaller axis range).

# Sub-space projected coupled-cluster

Andreas Ekström, Gaute Hagen PRL 123, 252501 (2019)



- Generalization of the eigenvector continuation method [Frame D. et al., Phys. Rev. Lett. 121, 032501 (2018)]
- Write the Hamiltonian in a linearized form

$$H(\vec{\alpha}) = \sum_{i=0}^{N_{\text{LECs}}=16} \alpha_i h_i$$

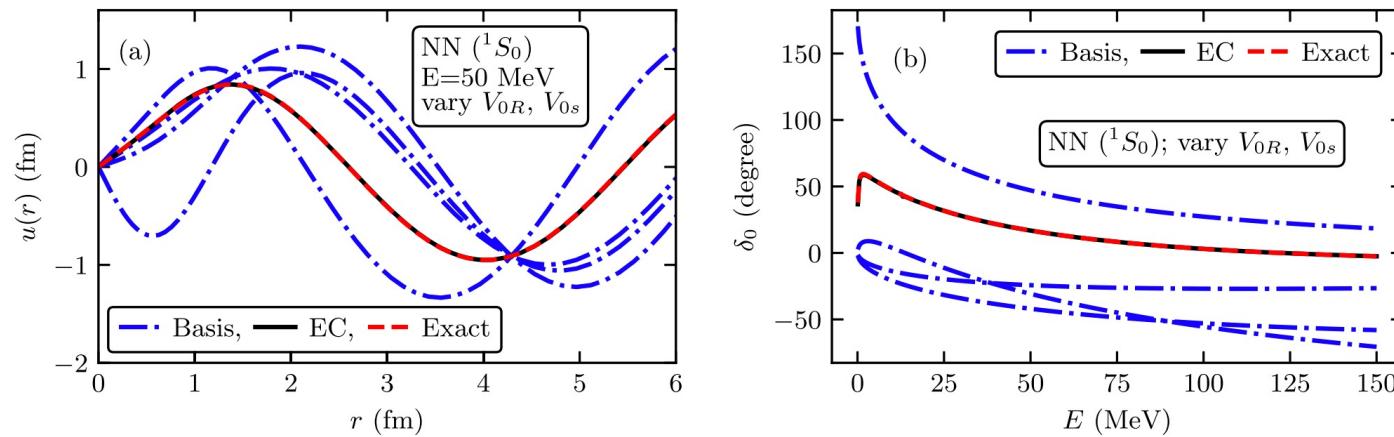
- Select “training points” where we solve exact CCSD
- Project the target Hamiltonian onto sub-space of training vectors and diagonalize the generalized eigen value problem

$$\mathbf{H}(\vec{\alpha}_\odot) \vec{c} = E(\vec{\alpha}_\odot) \mathbf{N} \vec{c}$$



# Efficient emulators for scattering using eigenvector continuation

R.J. Furnstahl  , A.J. Garcia  , P.J. Millican  , Xilin Zhang  



## Convergence of Eigenvector Continuation

Avik Sarkar<sup>\*</sup> and Dean Lee<sup>†</sup>

*Facility for Rare Isotope Beams and Department of Physics and Astronomy, Michigan State University,  
East Lansing, Michigan 48824, USA*

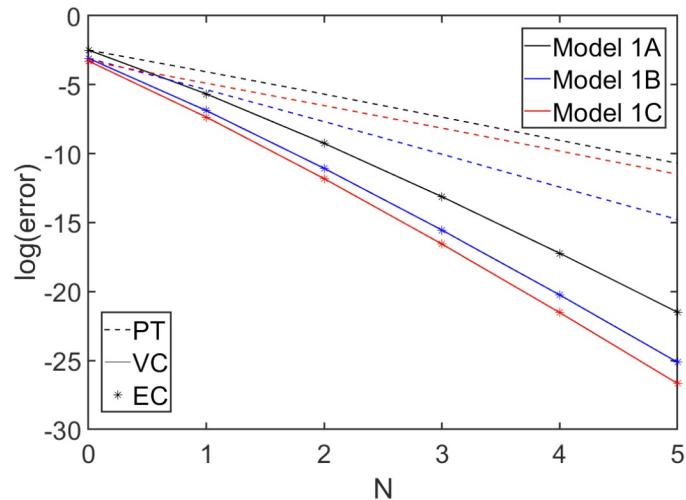
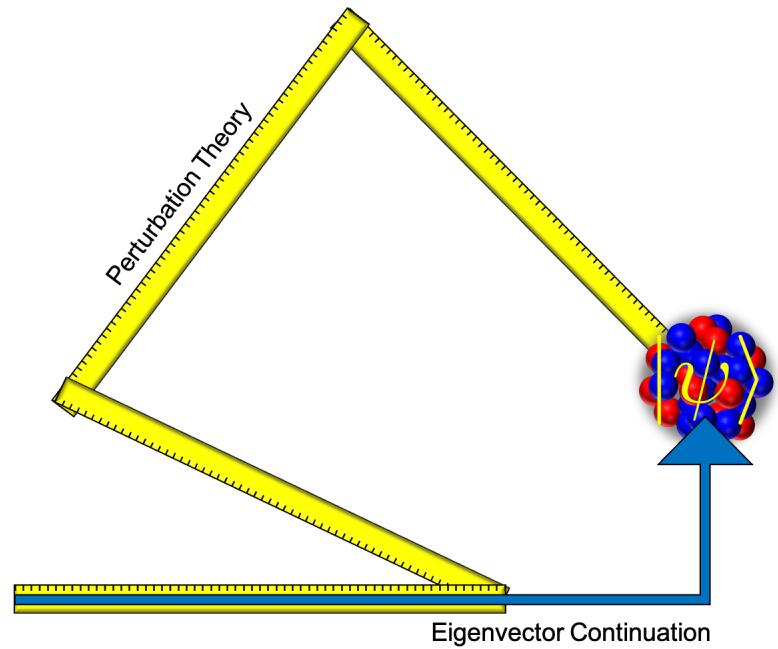


FIG. 1: (Color online) Logarithm of the error versus order  $N$  for eigenvector continuation (asterisks), vector continuation (solid lines), and perturbation theory (dashed lines). The three different colors (black, blue and red) correspond with Models 1A, 1B, and 1C respectively.



## Eigenvector continuation for the pairing Hamiltonian

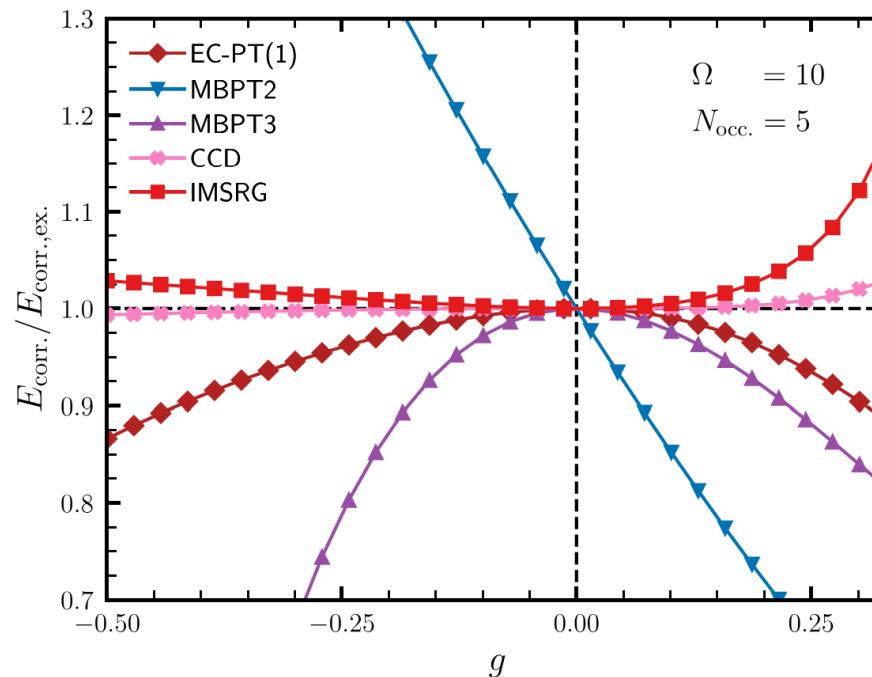
M. Companys Franzke<sup>1,\*</sup>, A. Tichai<sup>1,2,3,†</sup>, K. Hebeler<sup>1,2,3,‡</sup> and A. Schwenk<sup>1,2,3,§</sup>

<sup>1</sup>Technische Universität Darmstadt, Department of Physics, 64289 Darmstadt, Germany

<sup>2</sup>ExtreMe Matter Institute EMMI, GSI Helmholtzzentrum für Schwerionenforschung GmbH, 64291 Darmstadt, Germany

<sup>3</sup>Max-Planck-Institut für Kernphysik, Saupfercheckweg 1, 69117 Heidelberg, Germany

$$H(g) = \sum_p^{\Omega} \epsilon_p (c_p^\dagger c_p + c_{\bar{p}}^\dagger c_{\bar{p}}) - g \sum_{pq} c_p^\dagger c_{\bar{p}}^\dagger c_{\bar{q}} c_q$$



## Projection-based emulators

### Model reduction methods for nuclear emulators

J A Melendez<sup>1</sup> , C Drischler<sup>2</sup> , R J Furnstahl<sup>3,1</sup> , A J Garcia<sup>1</sup>  and Xilin Zhang<sup>2</sup> 

Published 1 September 2022 • © 2022 IOP Publishing Ltd

[Journal of Physics G: Nuclear and Particle Physics, Volume 49, Number 10](#)

**Citation** J A Melendez *et al* 2022 *J. Phys. G: Nucl. Part. Phys.* **49** 102001

**DOI** [10.1088/1361-6471/ac83dd](https://doi.org/10.1088/1361-6471/ac83dd)

Training and projecting: A reduced basis method emulator for many-body physics

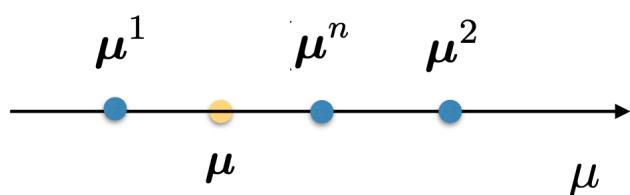
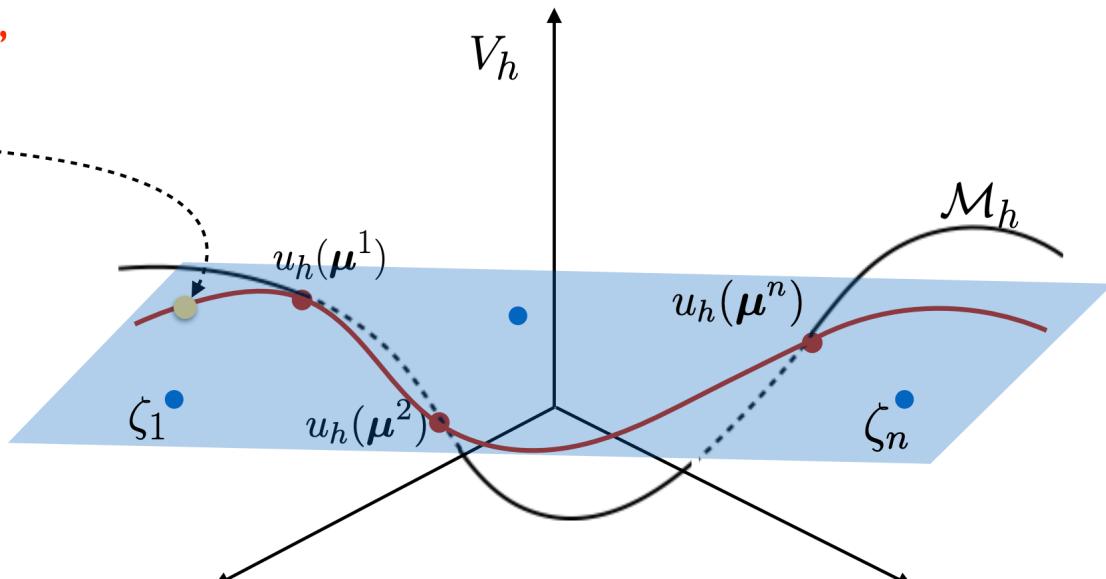
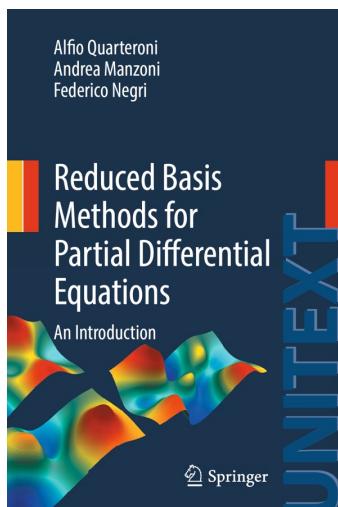
Edgard Bonilla, Pablo Giuliani, Kyle Godbey, and Dean Lee  
Phys. Rev. C **106**, 054322 – Published 17 November 2022

Reduced basis methods have been well-studied in the field of partial differential equations for several decades. Part of a larger class of methods called model order reduction. Use of Galerkin methods, extensions to nonlinear systems, etc.

# RB Approximation

RB Approximation  
(new parameter value,  
PROJECTION)

$$\{ u_n(\mu) : \mu \in \mathcal{P} \}$$



- **Snapshots computed offline**  
(properly selected parameters)
- **RB space:**  $V_n = \text{span}\{u_h(\mu^1), \dots, u_h(\mu^n)\}$   
 $= \text{span}\{\zeta_1, \dots, \zeta_n\}$
- **RB problem**  $P_n(\mu)$  solved **online**

## Parametric matrix models



Patrick Cook



Danny Jammooa

Patrick Cook, Danny Jammooa, Hjorth-Jensen, Da. Lee, D.L., arXiv:2401.11694

## Parametric matrix models

Consider the one-parameter affine eigenvalue problem

$$H(c) = H_0 + cH_1$$

We do eigenvector continuation with  $d$  training vectors. This corresponds to solving the eigenvalue problem for the new one-parameter affine system

$$M(c) = M_0 + cM_1$$

where the matrices are  $d$  by  $d$ .

Note that the eigenvalues will be roots of the characteristic polynomial

$$P[E(c)] = \det [E(c)I - M(c)] = \det [E(c)I - M_0 - cM_1]$$

The polynomial  $P$  will be degree  $d$  in with respect to  $E(c)$  and degree  $d$  with respect to  $c$ .

$$\begin{aligned} P[E(c)] &= [E(c)]^d c^0 \\ &\quad + [E(c)]^{d-1} [a_{d-1,1} c^1 + a_{d-1,0} c^0] \\ &\quad \dots \\ &\quad + [E(c)]^1 [a_{1,d-1} c^{d-1} + a_{1,d-2} c^{d-2} + \dots + a_{1,0}] \\ &\quad + [E(c)]^0 [a_{0,d} c^d + a_{0,d-1} c^{d-1} + a_{0,d-2} c^{d-2} + \dots + a_{0,0}] \end{aligned}$$

In contrast with polynomial interpolation or rational interpolation (Padé approximants), eigenvector continuation is performing algebraic interpolation using roots of polynomials.

Suppose that we don't have access to the training vectors. We can still make a matrix model

$$M(c) = M_0 + cM_1$$

We learn the unknown elements of the matrices using training data for the eigenvalues  $E(c)$

This is a very simple example of a general approach that we call parametric matrix models (PMM)

In contrast with most existing machine learning models that imitate the biology of neurons, parametric matrix models use matrix equations that emulate the physics of quantum systems

Instead trying to directly fit some output functions, we try to learn the set of equations whose solutions produce the output functions

This is similar to how we solve physics problems

While the set of possible solutions to the equations may be very complex, the equations themselves have a simple and logical structure

## Basic form for PMMs

Input features

$$\{c_1, c_2, \dots, c_{N_c}\}$$

Hermitian or unitary primary matrices that are analytic functions of the input features

$$\{P_1, P_2, \dots, P_{N_P}\}$$

Normalized eigenvectors of primary matrices

$$\{v_1^{(1)}, v_1^{(2)}, \dots, v_2^{(1)}, \dots, v_{N_P}^{(1)}, \dots\}$$

Secondary matrices that are analytic functions of the input features

$$\{S_1, S_2, \dots, S_{N_P}\}$$

Scalar outputs

$$\{v_j^{(i)\dagger} S_k v_{j'}^{(i')}, \dots\}$$

## Affine eigenvalue PMMs

Input features

$$\{c_1, c_2, \dots, c_{N_c}\}$$

Only one primary matrix  $P$  which is an affine function (degree 1 polynomial) of the input features

$$\{P_1, P_2, \dots, P_{N_P}\} \rightarrow P$$

Only one eigenvector of the primary matrix  $P$

$$\{v_1^{(1)}, v_1^{(2)}, \dots, v_2^{(1)}, \dots, v_{N_P}^{(1)}, \dots\} \rightarrow v^{(i)}$$

Only one secondary matrix  $S$  that equals  $P$

$$\{S_1, S_2, \dots, S_{N_P}\} \rightarrow S = P$$

The scalar output is the eigenvalue

$$v^{(i)\dagger} S v^{(i)} = v^{(i)\dagger} P v^{(i)} = \lambda_P^{(i)}$$

## Addition theorem for affine eigenvalue PMMs

Suppose that we have two affine eigenvalue PMMs with output functions

$$f(\{c_l\}) \text{ and } g(\{c_l\})$$

Claim: There exists another affine eigenvalue PMM with output function

$$f(\{c_l\}) + g(\{c_l\})$$

Proof: Let

$$P_f(\{c_l\}) \text{ and } P_g(\{c_l\})$$

be primary matrices, with some selected eigenvalues corresponding to

$$f(\{c_l\}) \text{ and } g(\{c_l\})$$

respectively.

Using the tensor product space, we define a new primary matrix as

$$P(\{c_l\}) = P_f(\{c_l\}) \otimes \mathbb{I} + \mathbb{I} \otimes P_g(\{c_l\})$$

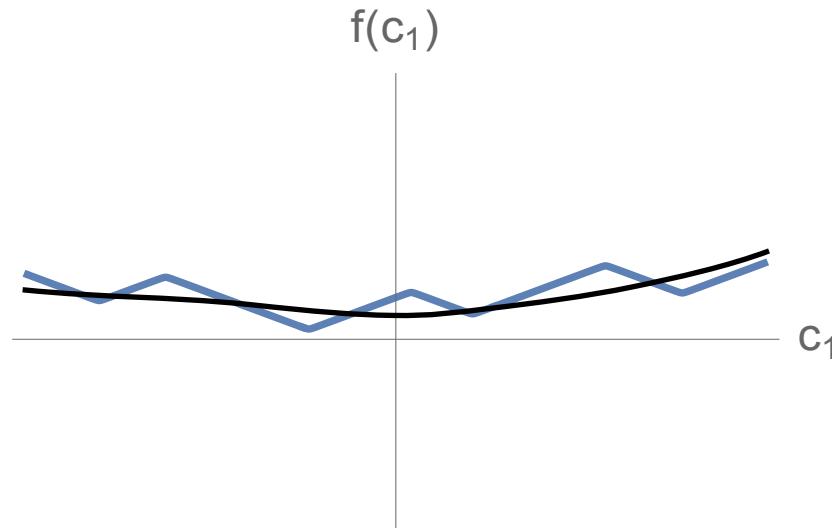
One of the eigenvalues will then give

$$f(\{c_l\}) + g(\{c_l\}) \quad [\text{end of proof}]$$

## Universal approximation theorem for affine eigenvalue PMMs *with one input feature*

Claim: Any continuous function  $f(c_1)$  over a compact domain can be uniformly approximated to arbitrary accuracy using an affine eigenvalue PMM

Proof: Any continuous function  $f(c_1)$  over a compact domain can be uniformly approximated to arbitrary accuracy using a concatenation of line segments with finite slope

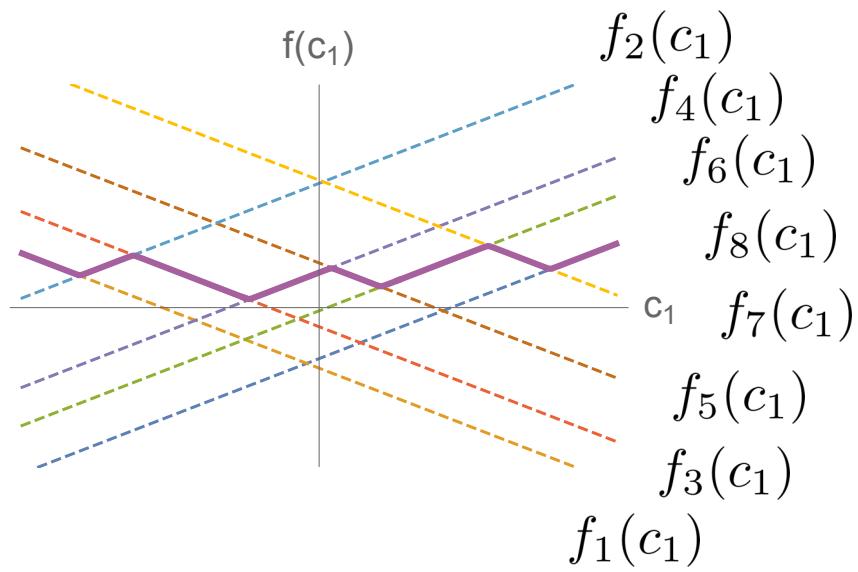


It suffices to prove that we can produce any such concatenation of line segments using an affine eigenvalue PMM

Consider the affine functions that pass through the line segments

$$\begin{aligned}f_1(c_1) &= a_1 c_1 + b_1 \\f_2(c_1) &= a_2 c_1 + b_2 \\f_3(c_1) &= a_3 c_1 + b_3 \\f_4(c_1) &= a_4 c_1 + b_4\end{aligned}$$

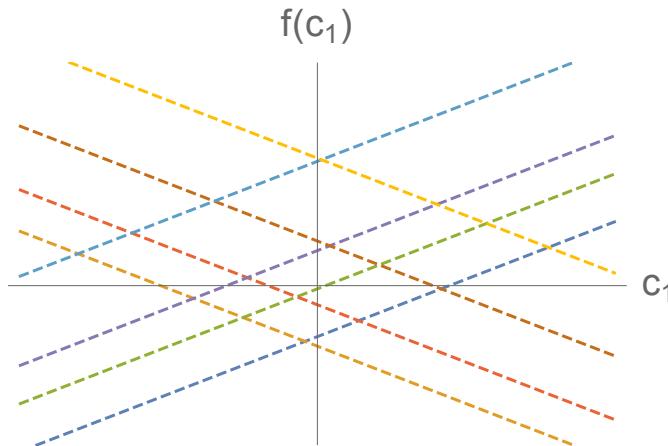
$$\begin{aligned}f_5(c_1) &= a_5 c_1 + b_5 \\f_6(c_1) &= a_6 c_1 + b_6 \\f_7(c_1) &= a_7 c_1 + b_7 \\f_8(c_1) &= a_8 c_1 + b_8\end{aligned}$$



Consider a diagonal matrix with elements given by the affine functions

$$\begin{bmatrix} a_1c_1 + b_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_2c_1 + b_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_3c_1 + b_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_4c_1 + b_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a_5c_1 + b_5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_6c_1 + b_6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & a_7c_1 + b_7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_8c_1 + b_8 \end{bmatrix}$$

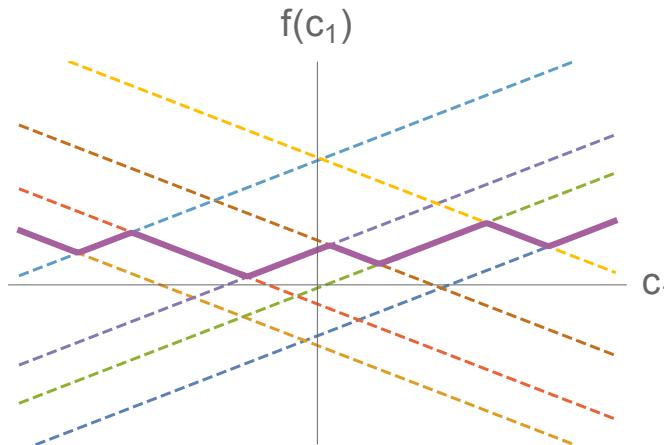
The eigenvalues are then just the affine functions



We now include nonzero infinitesimal off-diagonal elements

$$\begin{bmatrix} a_1c_1 + b_1 & \epsilon \\ \epsilon & a_2c_1 + b_2 & \epsilon & \epsilon & \epsilon & \epsilon & \epsilon & \epsilon \\ \epsilon & \epsilon & a_3c_1 + b_3 & \epsilon & \epsilon & \epsilon & \epsilon & \epsilon \\ \epsilon & \epsilon & \epsilon & a_4c_1 + b_4 & \epsilon & \epsilon & \epsilon & \epsilon \\ \epsilon & \epsilon & \epsilon & \epsilon & a_5c_1 + b_5 & \epsilon & \epsilon & \epsilon \\ \epsilon & \epsilon & \epsilon & \epsilon & \epsilon & a_6c_1 + b_6 & \epsilon & \epsilon \\ \epsilon & \epsilon & \epsilon & \epsilon & \epsilon & \epsilon & a_7c_1 + b_7 & \epsilon \\ \epsilon & a_8c_1 + b_8 \end{bmatrix}$$

The new eigenvalues exhibit sharp avoided level crossings and produce the desired concatenation of line segments



[end of proof]

## Universal approximation theorem for affine eigenvalue PMMs

Claim: Any continuous multivariate function of the input features

$$\{c_1, c_2, \dots, c_{N_c}\}$$

over any compact domain can be uniformly approximated to arbitrary accuracy using an affine eigenvalue PMM

The Stone-Weierstrass theorem says that we can uniformly approximate any continuous function to arbitrary accuracy over a compact domain using polynomials

So, it suffices to uniformly approximate any polynomial using an affine eigenvalue PMM.

By the addition theorem, it suffices to uniformly approximate any monomial using an affine eigenvalue PMM

Consider any perfect power of a linear form of the input features

$$\left[ \sum_{l=1}^{N_c} a_{j,l} c_l \right]^k = \sum_{k_1+k_2+\dots+k_{N_c}=k} \binom{k}{k_1, k_2, \dots, k_{N_c}} (a_{j,1}^{k_1} a_{j,2}^{k_2} \cdots a_{j,N_c}^{k_{N_c}}) (c_1^{k_1} c_2^{k_2} \cdots c_{N_c}^{k_{N_c}})$$

The coefficients of the monomials

$$c_1^{k_1} c_2^{k_2} \cdots c_{N_c}^{k_{N_c}}$$

are given by

$$\binom{k}{k_1, k_2, \dots, k_{N_c}} (a_{j,1}^{k_1} a_{j,2}^{k_2} \cdots a_{j,N_c}^{k_{N_c}})$$

Each of these coefficients are linearly independent functions of

$$\{a_{j,1}, a_{j,2}, \dots, a_{j,N_c}\}$$

Therefore, we can write any monomial as a linear combination of perfect powers of linear forms

$$c_1^{k_1} c_2^{k_2} \cdots c_{N_c}^{k_{N_c}} = \sum_{j=1}^J b_j \left[ \sum_{l=1}^{N_c} a_{j,l} c_l \right]^k$$

Since each

$$b_j \left[ \sum_{l=1}^{N_c} a_{j,l} c_l \right]^k$$

is a function of only one variable, it can be uniformly approximated using an affine eigenvalue PMM. By the addition theorem, the universal approximation theorem is proven.

[end of proof]

## Eigenvector continuation and PMMs

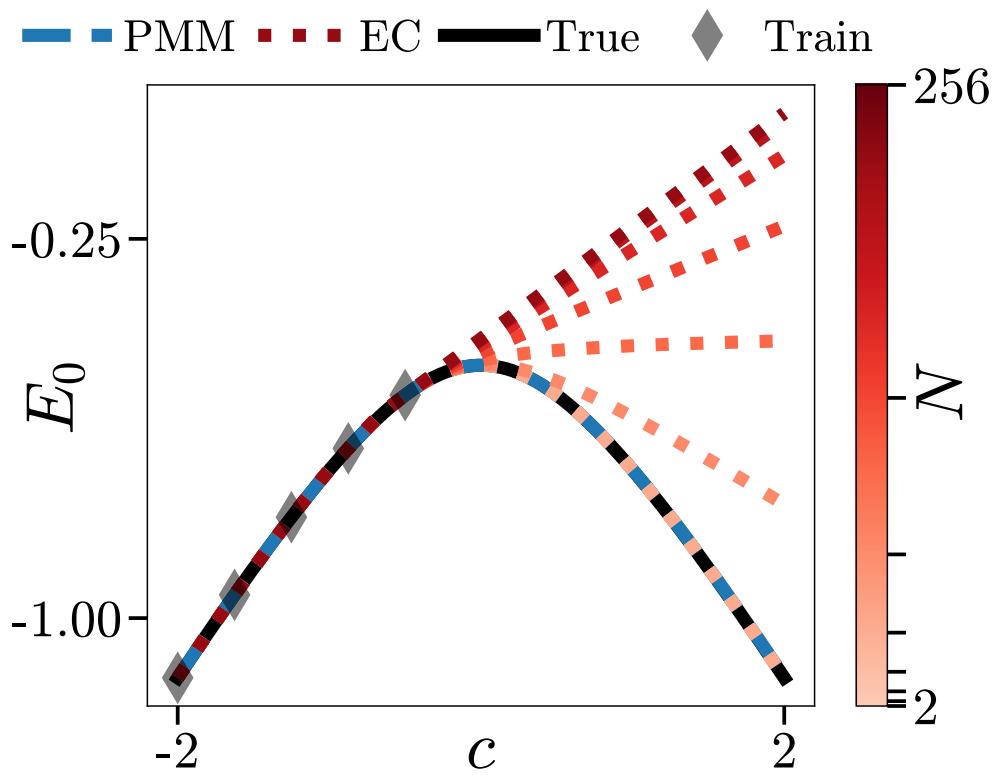
Eigenvector continuation (EC) can be viewed as an example of a PMM, with an efficient means to construct the elements of the matrix model using high-fidelity snapshots.

With enough training data, a PMM can learn the matrix elements corresponding to any EC calculation, up to an arbitrary unitary transformation that is not relevant for the output

Interestingly, for extremely large vector spaces there are systems where a general PMM does better than EC

Consider  $N = 256$  non-interacting spin-1/2 fermions with Hamiltonian

$$H(c) = \frac{1}{2N} \sum_i^N (\sigma_i^z + c\sigma_i^x)$$

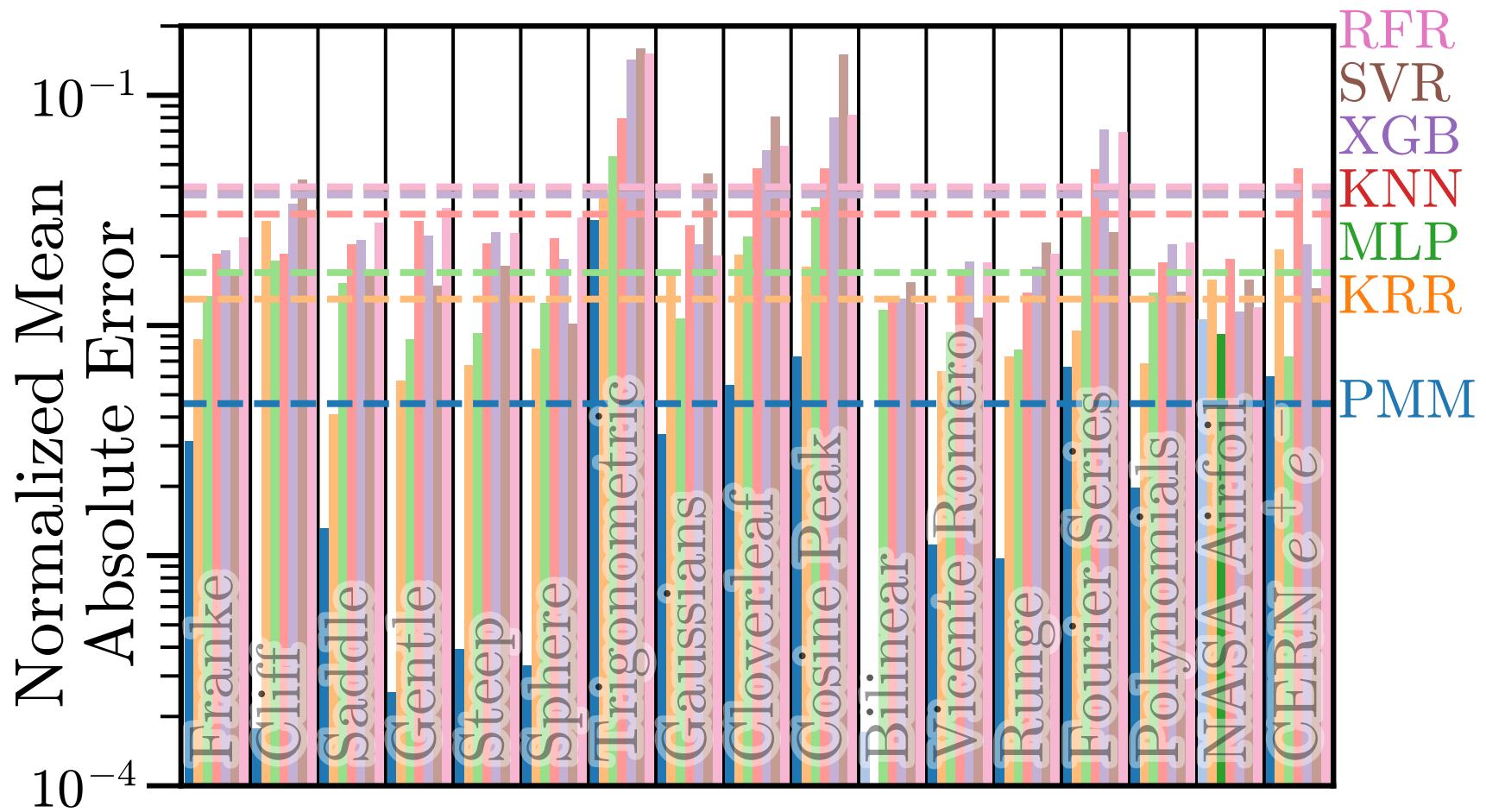


# Application to function interpolation

Name	Equation
Franken <sup>†36, 38</sup>	$\frac{3}{4} \exp \left\{ -\frac{(9x-2)^2 + (9y-2)^2}{4} \right\} + \frac{3}{4} \exp \left\{ -\frac{(9x+1)^2 - 9y+1}{49} \right\} \\ + \frac{1}{2} \exp \left\{ -\frac{(9x-7)^2 + (9y-3)^2}{4} \right\} - \frac{1}{5} \exp \{ -(9x-4)^2 - (9y-7)^2 \}$
Cliff <sup>†38</sup>	$\frac{1}{9} \tanh[9(y-x)] + \frac{1}{9}$
Saddle <sup>†38</sup>	$\frac{5/4 + \cos(27y/5)}{6 + 6(3x-1)^2}$
Gentle <sup>†38</sup>	$\frac{1}{3} \exp \left\{ -\alpha \left[ (x - \frac{1}{2})^2 + (y - \frac{1}{2})^2 \right] \right\}, \quad \alpha = 81/16$
Steep <sup>†38</sup>	$\frac{1}{3} \exp \left\{ -\alpha \left[ (x - \frac{1}{2})^2 + (y - \frac{1}{2})^2 \right] \right\}, \quad \alpha = 81/4$
Sphere <sup>†38</sup>	$-\frac{1}{2} + \sqrt{\left(\frac{8}{9}\right)^2 - (x - \frac{1}{2})^2 - (y - \frac{1}{2})^2}$
Trigonometric <sup>†38</sup>	$2 \cos(10x) \sin(10y) + \sin(10xy)$
Gaussians <sup>†38</sup>	$\exp \{ -u^2/2 \} + \frac{3}{4} \exp \{ -v^2/2 \} [1 + \exp \{ -u^2/2 \}], \quad \begin{cases} u = 5 - 10x \\ v = 5 - 10y \end{cases}$
Cloverleaf <sup>†38</sup>	$\left[ \left( \frac{20}{3} \right)^3 uv \right]^2 \left[ \left( \frac{1}{1+u} \right) \left( \frac{1}{1+v} \right) \right]^5 \left[ u - \frac{2}{1+u} \right] \left[ v - \frac{2}{1+v} \right], \quad \begin{cases} u = \exp \left\{ \frac{10 - 20x}{3} \right\} \\ v = \exp \left\{ \frac{10 - 20y}{3} \right\} \end{cases}$
Cosine Peak <sup>†38</sup>	$\exp \left\{ -\frac{2}{5} r \right\} \cos \left( \frac{3}{2} r \right), \quad r = \sqrt{(8x-4)^2 + (9y-\frac{9}{2})^2}$
Bilinear <sup>†38</sup>	$xy + x$
Vicente Romero <sup>†37, 38</sup>	$\frac{6}{5} r + \frac{21}{40} \sin \left( \frac{12\pi}{5\sqrt{2}} r \right) \sin \left[ \frac{13}{10} \text{atan2}(y, x) \right], \quad r = \sqrt{x^2 + y^2}$
Runge <sup>†35, 38</sup>	$[(10x-5)^2 + (10y-5)^2 + 1]^{-1}$
Fourier series <sup>‡</sup>	$\sum_{n=1}^N \sum_{m=1}^M \left[ a_{nm} \sin \left( \frac{n\pi x}{3} \right) \sin \left( \frac{m\pi y}{3} \right) + b_{nm} \cos \left( \frac{n\pi x}{3} \right) \cos \left( \frac{m\pi y}{3} \right) \right], \quad a_{nm}, b_{nm} \sim \mathcal{N}(0, 1)$
Polynomials <sup>‡</sup>	$\sum_{n=0}^N \sum_{m=0}^M a_{nm} x^n y^m, \quad a_{nm} \sim \mathcal{N}(0, 1)$

†  $(x, y) \in [0, 1] \times [0, 1]$

‡  $(x, y) \in [-1, 1] \times [-1, 1]$



Parametric Matrix Model (PMM), Kernel Ridge Regression (KRR), Multilayer Perceptron (MLP), k-Nearest Neighbors (KNN), Extreme Gradient Boosting (XGB), Support Vector Regression (SVR), and Random Forest Regression (RFR)

## Applications to quantum computing

In quantum computing, we often need to compute the time evolution of a quantum state

$$|\psi(t)\rangle = \exp(-iHt) |\psi(0)\rangle$$

The Baker-Campbell-Hausdorff expansion gives

$$\exp(A) \exp(B) = \exp(C)$$

$$C = A + B + \frac{1}{2}[A, B] + \frac{1}{12}[A, [A, B]] - \frac{1}{12}[B, [A, B]] \dots$$

We can use this to exponentiate a Hamiltonian with pieces that do not commute

If our Hamiltonian has two non-commuting pieces

$$H = H_A + H_B$$

then we can use either of the first-order Trotter-Suzuki approximations

$$\exp(-iH\Delta t) = \exp(-iH_A\Delta t) \exp(-iH_B\Delta t) + O(\Delta t^2)$$

$$\exp(-iH\Delta t) = \exp(-iH_B\Delta t) \exp(-iH_A\Delta t) + O(\Delta t^2)$$

If our Hamiltonian has three non-commuting pieces

$$H = H_A + H_B + H_C$$

Then we have the first-order Trotter-Suzuki expressions

$$\exp(-iH\Delta t) = \exp(-iH_A\Delta t) \exp(-iH_B\Delta t) \exp(-iH_C\Delta t) + O(\Delta t^2)$$

$$\exp(-iH\Delta t) = \exp(-iH_B\Delta t) \exp(-iH_A\Delta t) \exp(-iH_C\Delta t) + O(\Delta t^2)$$

(also other orderings)

The second-order Trotter-Suzuki approximation has the form

$$\exp(-iH\Delta t) =$$

$$\begin{aligned} & \exp(-iH_C \frac{\Delta t}{2}) \exp(-iH_B \frac{\Delta t}{2}) \exp(-iH_A \Delta t) \exp(-iH_B \frac{\Delta t}{2}) \exp(-iH_C \frac{\Delta t}{2}) \\ & \quad + O(\Delta t^3) \end{aligned}$$

(also other orderings)

## Example: Heisenberg spin chain

Let us consider a one-dimension spin chain with an external magnetic field and couplings between nearest neighbor sites

$$H = -J \sum_j \sigma_{j+1}^x \sigma_j^x - J \sum_j \sigma_{j+1}^y \sigma_j^y + U \sum_j \sigma_{j+1}^z \sigma_j^z + \sum_j h_j \sigma_j^z$$

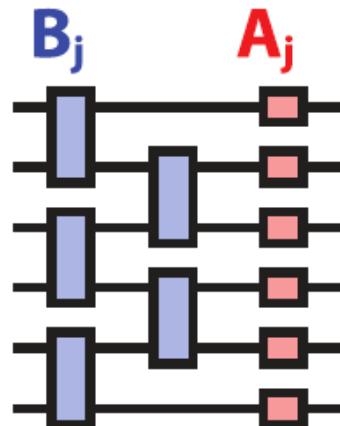
Let us define

$$A_j = \exp(-ih_j \sigma_j^z \Delta t)$$

$$B_j = \exp[-i(-J\sigma_j^x \sigma_{j+1}^x - J\sigma_j^y \sigma_{j+1}^y + U\sigma_j^z \sigma_{j+1}^z) \Delta t]$$

We can use the first-order Trotter-Suzuki approximation

$$\exp(-iH\Delta t) = \left(\prod_j A_j\right) \left(\prod_{j \text{ even}} B_j\right) \left(\prod_{j \text{ odd}} B_j\right)$$



Smith *et al.*, npj Quant. Info. 5 106 (2019)

We need to take the limit  $\Delta t \rightarrow 0$

But this is computationally difficult due to qubit decoherence and gate fidelity

For the PMM, we can make a low-dimensional matrix model using the same Trotter-Suzuki structure

$$\exp(-iM\Delta t) = \exp(-iM^A\Delta t) \exp(-iM_{\text{even}}^B\Delta t) \exp(-iM_{\text{odd}}^B\Delta t)$$

## Heisenberg spin chain with Dzyaloshinskii-Moriya interaction

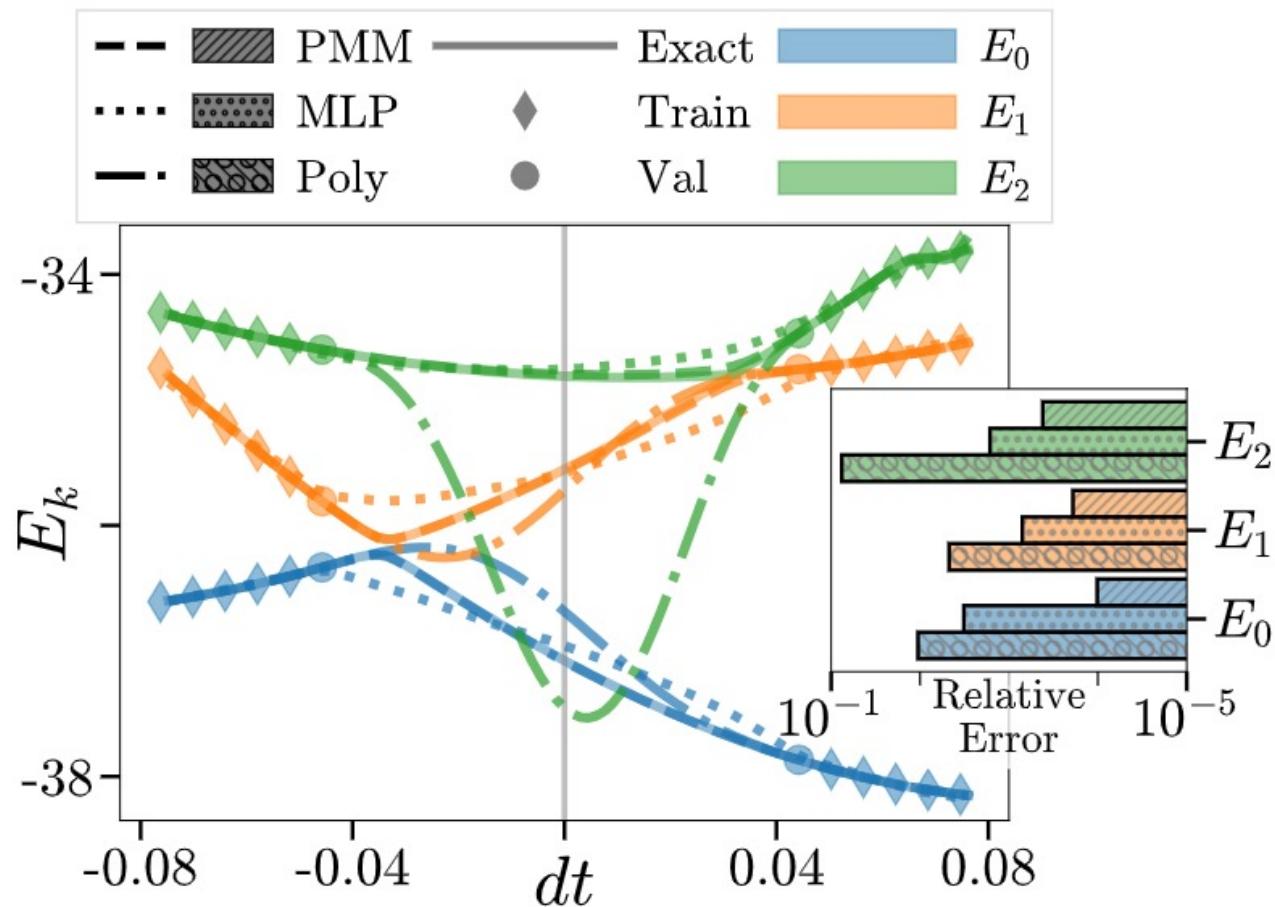
$$H = H_B + H_J^0 + H_J^1 + H_D^0 + H_D^1$$

$$H_B = B \sum_i r_i \sigma_i^z$$

$$H_J^{0|1} = J \sum_{i \text{ even}| \text{odd}} (\sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y + \sigma_i^z \sigma_{i+1}^z)$$

$$H_D^{0|1} = D \sum_{i \text{ even}| \text{odd}} (\sigma_i^x \sigma_{i+1}^y - \sigma_i^y \sigma_{i+1}^x)$$

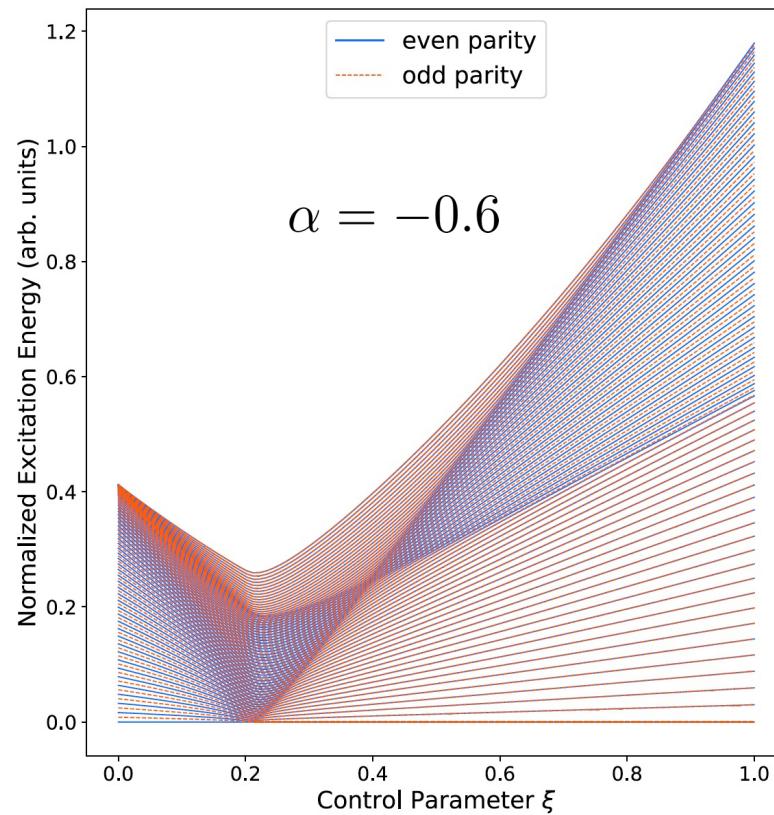
## Heisenberg spin chain with Dzyaloshinskii-Moriya interaction



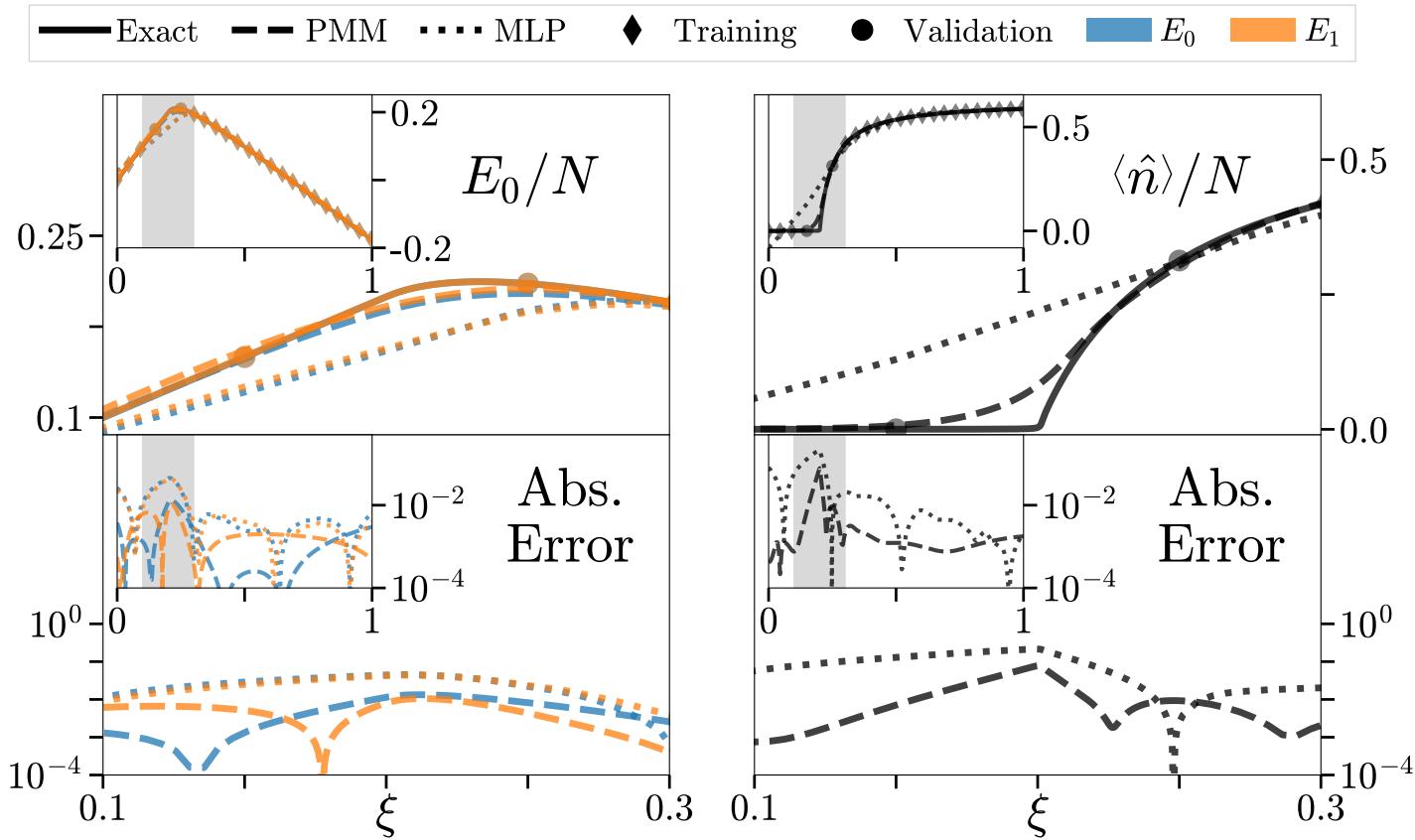
## Applications to quantum many-body systems

Anharmonic Lipkin model

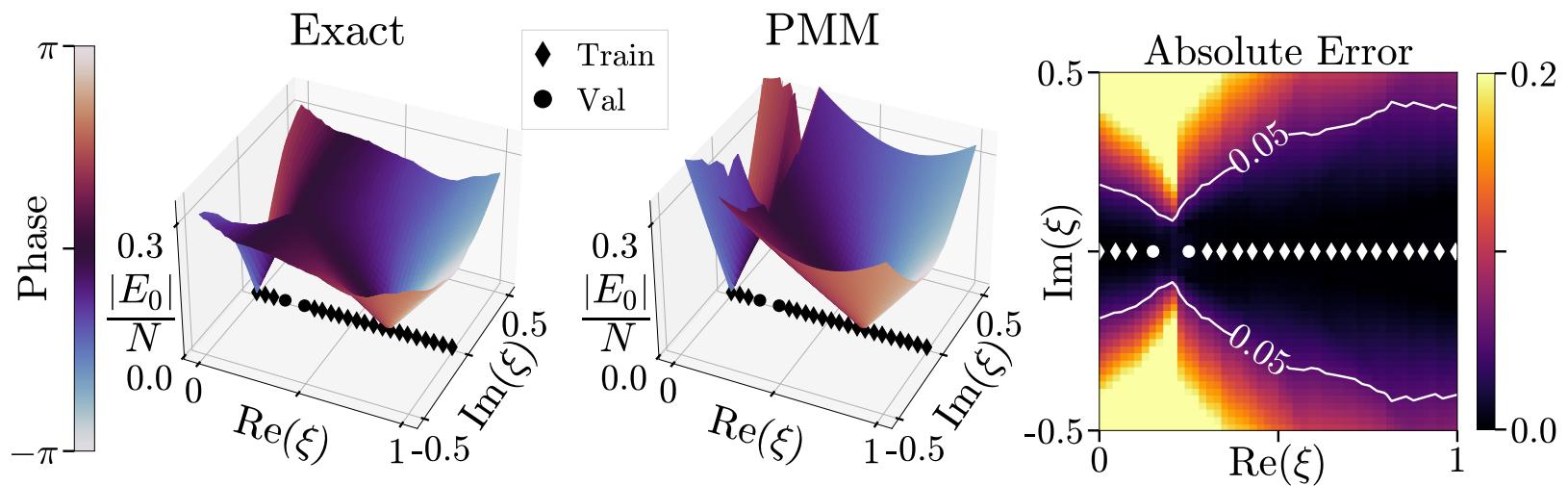
$$H = (1 - \xi)(S + \hat{S}_z) + 2\xi/S(S^2 - \hat{S}_x^2) + \alpha/2S(S + \hat{S}_z + 1)$$



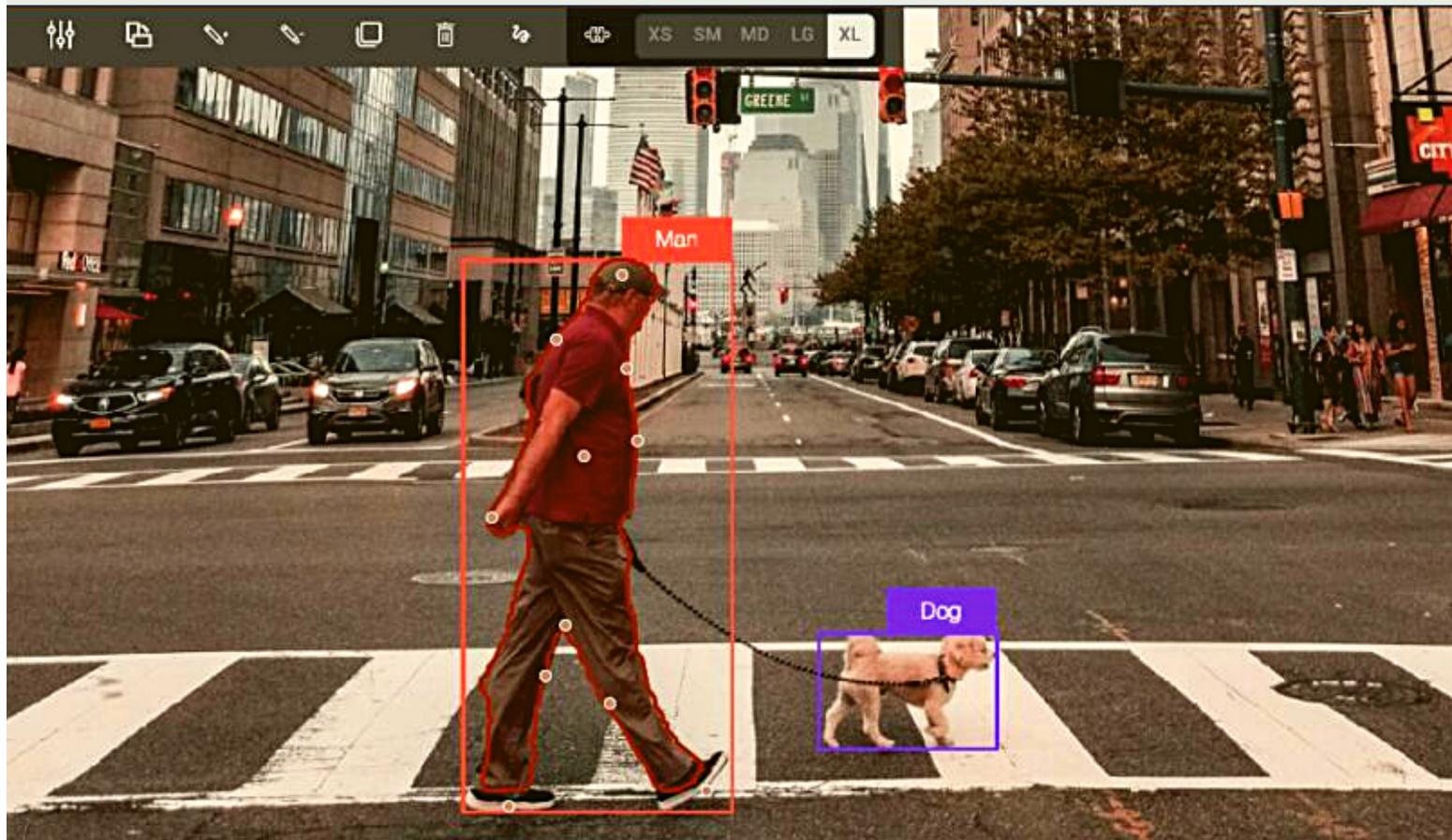
## Results for the anharmonic Lipkin model for $N = 1000$



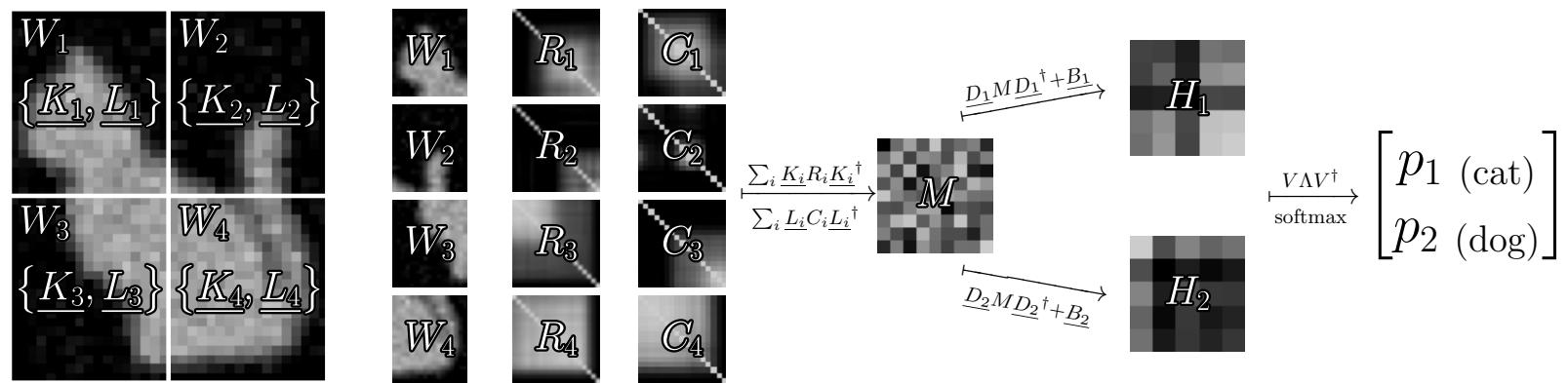
## Extrapolation to the complex plane



## Applications to general machine learning



Kili Technology

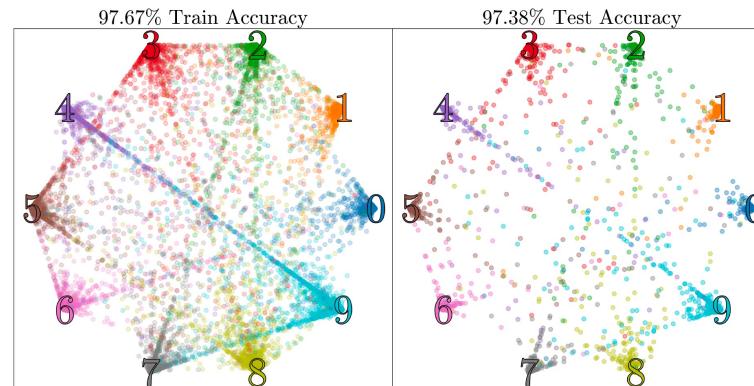


# MNIST Digits

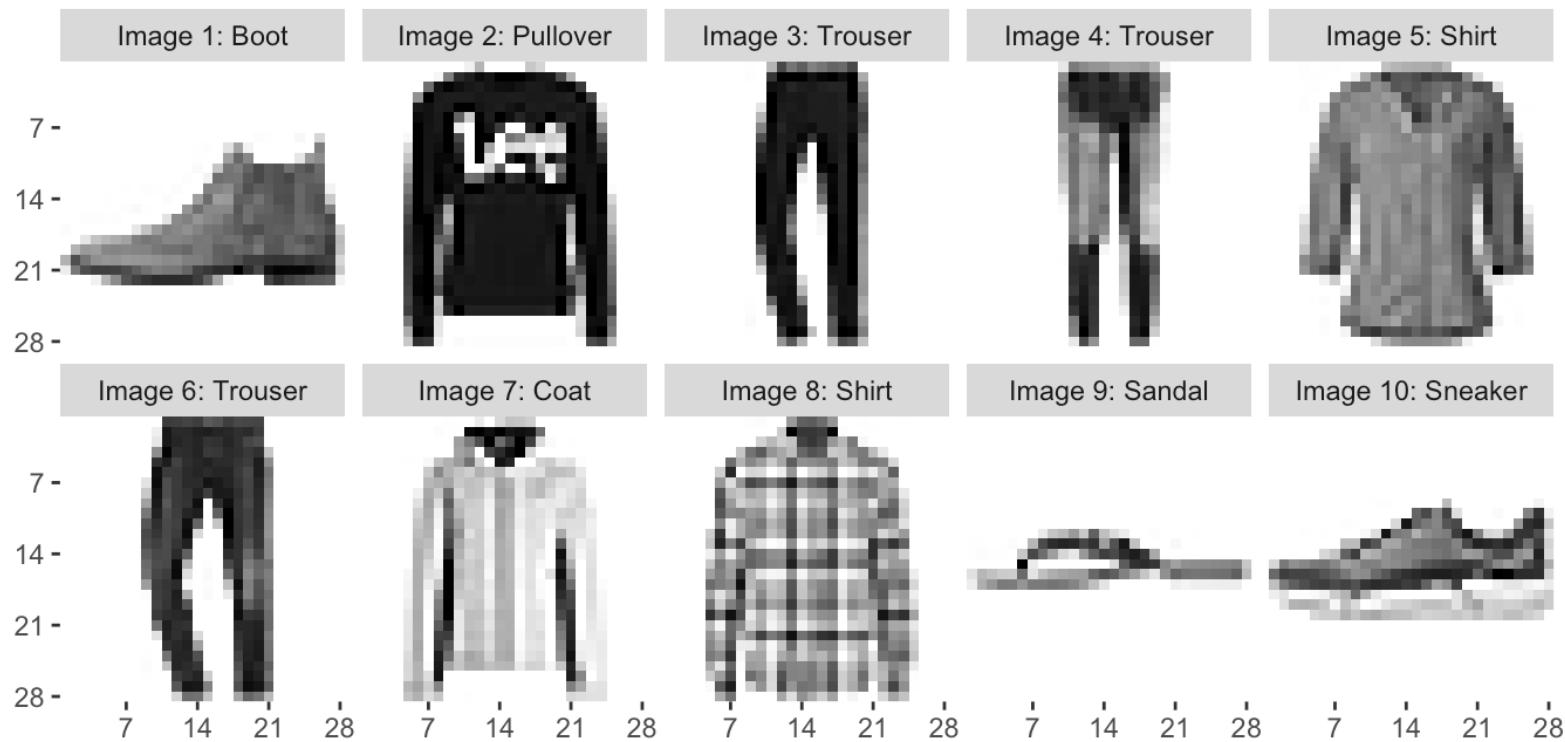


**DNN-2** 96.4% with 5500 parameters

**PMM** 97.4% with 4990 parameters



## Fashion MNIST



## Extended MNIST



<b>Dataset</b>	<b>Model</b>	<b>Accuracy</b>	<b>Trainable float</b>
MNIST Digits <sup>53</sup>	PMM <sup>†</sup>	97.38	4990
	DNN-2 <sup>54</sup>	96.5	5500
	DNN-3 <sup>54</sup>	97.0	80 598
	DNN-5 <sup>54</sup>	97.2	175 180
	GECCO <sup>55</sup>	98.04	19 000
	CTM-250 <sup>56</sup>	98.82	31 750
	CTM-8000 <sup>56</sup>	99.33	527 250
Fashion MNIST <sup>58</sup>	Efficient-CapsNet <sup>57</sup>	99.84	161 824
	PMM <sup>†</sup>	88.58	16 744
	GECCO <sup>55</sup>	88.09	19 000
	CTM-250 <sup>56</sup>	88.25	31 750
	CTM-8000 <sup>56</sup>	91.18	527 250
	MLP <sup>†</sup> <sup>59</sup>	91.63	$2.9 \times 10^6$
	VGG8B(2x) <sup>59</sup>	95.45	$28 \times 10^6$
EMNIST Balanced <sup>61</sup>	Fine-Tuning DARTS <sup>60</sup>	96.91	$3.2 \times 10^6$
	PMM <sup>†</sup>	81.57	13 792
	OPIUM <sup>†</sup> <sup>61</sup>	78.02	$8.32 \times 10^6$
	HM2-BP <sup>62</sup>	85.57	$6.7 \times 10^5$
	CNN <sup>63</sup>	79.61	21 840
	CNN (Spinal FC) <sup>63</sup>	82.77	13 820
	CNN (Spinal FC) <sup>63</sup>	83.21	16 050

## Summary and outlook

The first part of the talk introduced some concepts of eigenvector continuation and recent applications. We then discussed a new machine learning approach called parameter matrix models that uses the matrix equations of quantum physics. We proved a universal approximation theorem and showed applications to scientific computing as well as general machine learning applications such as image recognition. There are numerous possible applications across many fields.