Finite-volume pionless effective field theory for few-nucleon systems with differentiable programming

Xiangkai Sun, William Detmold, Di Luo, and Phiala E. Shanahan

Finite-Volume Pionless Effective Field Theory for Few-Nucleon Systems with Differentiable Programming

Xiangkai Sun, William Detmold, Di Luo, and Phiala E. Shanahan

Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA and
The NSF AI Institute for Artificial Intelligence and Fundamental Interactions

Finite-volume pionless effective field theory provides an efficient framework for the extrapolation of nuclear spectra and matrix elements calculated at finite volume in lattice QCD to infinite volume, and to nuclei with larger atomic number. In this work, it is demonstrated how this framework may be implemented via a set of correlated Gaussian wavefunctions optimised using differentiable programming and via solution of a generalised eigenvalue problem. This approach is shown to be significantly more efficient than a stochastic implementation of the variational method based on the same form of correlated Gaussian wavefunctions, yielding comparably accurate representations of the ground-state wavefunctions with an order of magnitude fewer terms. The efficiency of representation allows such calculations to be extended to larger systems than in previous work. The method is demonstrated through calculations of the binding energies of nuclei with atomic number $A \in \{2, 3, 4\}$ in finite volume, matched to lattice QCD calculations at quark masses corresponding to $m_\pi = 806$ MeV, and infinite-volume effective field theory calculations of $A \in \{2, 3, 4, 5, 6\}$ systems based on this matching.

I. INTRODUCTION

A central goal of nuclear physics is to make predictions for the spectra and properties of nuclear systems based on the underlying degrees of freedom of the Standard Model, most pertinently quarks and gluons. Since nuclei and other relevant systems exhibit dynamics at energy scales where the interactions between quarks and gluons, governed in the Standard Model by the theory of Quantum Chromodynamics (QCD), are nonperturbative, this goal can be addressed directly only by numerical calculations in the framework of lattice QCD (LQCD). However, due to computational limitations, LQCD studies of nuclei have so far been restricted to systems with atomic number $A \leq 4$, with unphysically large values of the quark masses. Moreover, to date only proof-of-principle LQCD calculations of nuclei have been performed [1–16], in which systematic uncertainties such as those from the lattice discretisation are estimated but not fully quantified.

While fully-controlled LQCD calculations of light nuclei will likely be achieved in the near future, the computational costs of such studies scale exponentially with A in current approaches, and as such, the restriction to small nuclear systems is likely to persist until novel algorithms or other computational breakthroughs render calculations of larger nuclei tractable. Pionless nuclear effective field theory (EFT) [17–24] provides a bridge between tractable LQCD calculations of light nuclei and the broader scope of low-energy nuclear phenomenology. In nuclear physics, it is apparent that there is a hierarchy of interactions, in that two-body interactions are more important in governing nuclear structure and reactions than three-body interactions, which are in turn more important than four-body interactions, and so on. Consequently, LQCD calculations of $A \leq 4$ systems can be used to constrain the most relevant couplings in nuclear

EFT which can subsequently be used to make predictions for larger nuclear systems and for matrix elements which may not have been directly computed in LQCD. In addition, since the finite volume in which LQCD calculations are performed produces effects which are long-distance in nature, they can be captured in nuclear EFT calculations in appropriately matched finite volumes (finite volume nuclear EFT (FVEFT)). With the couplings of the EFT determined by this matching, the EFT provides a method to extract infinite volume physics from finite-volume LQCD spectra and matrix elements.

Existing applications of FVEFT to the matching and extrapolation of LQCD results for nuclear spectra [25] and matrix elements [26] have used the stochastic variational method (SVM) [27] with trial wavefunctions composed of shifted correlated Gaussian functions [28]. Because of the stochastic nature of this approach, a large number of terms are required to approximate the ground state of each nuclear system. In this work, a new differentiable programming (DP) approach is introduced that implements an optimisation of the parameters defining each Gaussian term that is included in the trial wavefunction, as opposed to the stochastic selection of terms, resulting in much more efficient representations. Further improvement through the combination of multiple sets of optimised trial states can be achieved through solution of a generalised eigenvalue problem (GEVP). The compactness of the resulting wavefunction representations makes it feasible to extend previous calculations to systems of larger A. In this work, FVEFT predictions are made for the ^{4}He ground state as a function of volume, and the FVEFT matching of two and three-body interactions enables predictions for the infinite-volume energies of $A \in \{5, 6\}$ systems.

The following section outlines important aspects of nuclear FVEFT and the differential programming method used to determine optimal wavefunctions in the approach proposed here. Section III presents results of the op-

timisation procedure for finite-volume systems with $A \in \{2, 3, 4\}$, and their matching to LQCD energy determinations. Infinite-volume binding energies are also presented for $A \leq 6$. Section IV provides a summary and outlook for this approach.

II. METHODOLOGY

A. Hamiltonian for pionless effective field theory

The low-energy interactions of nucleons are described in pionless EFT (EFT$_{\not\pi}$) by the Lagrangian [17–24]

$$
\mathcal{L} = N^\dagger \left(iD_0 + \frac{\mathbf{D}^2}{2M_N} \right) N
$$
$$
- \frac{1}{2} \left[C_0 \left(N^\dagger N \right)^2 + C_1 \left(N^\dagger \vec{\sigma} N \right)^2 \right] \quad (1)
$$
$$
- \frac{D_0}{6} (N^\dagger N)^3 + \dots.
$$

The first, second, and third lines present the leading-order single-nucleon kinetic operator expanded in the non-relativistic (NR) limit, the two-body interaction, and the three-body interaction, respectively (the latter is promoted to leading order to define a valid power-counting scheme). N denotes the nucleon field, M_N the nucleon mass, $\vec{\sigma}$ the vector of Pauli matrices acting in spin space of a given nucleon, and $\{C_0, C_1\}$ and D_0 denote the relevant two and three-body low-energy constants (LECs). A common alternate basis for the two-nucleon interactions yields related LECs

$$
C_T = C_0 - 3C_1 \quad \text{and} \quad C_S = C_0 + C_1. \quad (2)
$$

The corresponding n-particle non-relativistic Hamiltonian can be expressed as

$$
H = -\frac{1}{2M_N} \sum_i \nabla_i^2 + \sum_{i<j} V_2 \left(\mathbf{r}_{ij} \right) + \sum_{i<j<k} V_3 \left(\mathbf{r}_{ij}, \mathbf{r}_{jk} \right),
$$
$$(3)$$

where the n particles are labelled by indices $i, j, k \in \{1, \dots, n\}$ and the Laplacian for particle i is expressed as ∇_i^2. $V_2 \left(\mathbf{r}_{ij} \right)$ and $V_3 \left(\mathbf{r}_{ij}, \mathbf{r}_{jk} \right)$ denote the two and three-particle potentials, which are regulated using Gaussian smearing, and are functions of the displacements between particles, defined for particles i and j as $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$, where $\mathbf{r}_i = (r_i^{(x)}, r_i^{(y)}, r_i^{(z)})$. In particular,

$$
V_2 \left(\mathbf{r}_{ij} \right) = \left(C_0 + C_1 \sigma^{(i)} \cdot \sigma^{(j)} \right) g_\Lambda \left(\mathbf{r}_{ij} \right), \quad (4)
$$

and

$$
V_3 \left(\mathbf{r}_{ij}, \mathbf{r}_{jk} \right) = D_0 \sum_{\text{cyc}} g_\Lambda \left(\mathbf{r}_{ij} \right) g_\Lambda \left(\mathbf{r}_{jk} \right), \quad (5)
$$

where $\sum_{\text{cyc}}$ denotes the sum over all cyclic permutations of $\{i, j, k\}$, and the Gaussian regulator in infinite spatial volume is defined as

$$
g_\Lambda(\mathbf{r}) = \frac{\Lambda^3}{8\pi^{3/2}} \exp \left(-\Lambda^2 |\mathbf{r}|^2 / 4 \right)
$$
$$
= \frac{\Lambda^3}{8\pi^{3/2}} \prod_{\alpha \in \{x,y,z\}} \exp \left(-\Lambda^2 r^{(\alpha)2} / 4 \right). \quad (6)
$$

The regulator parameter Λ can be expressed in terms of a length-scale r_0 as $\Lambda = \sqrt{2}/r_0$. Physical quantities are independent of this cutoff [26].

In a finite cubic spatial volume with side-length L, the regulator can be constructed to be periodic by summing $g_\Lambda(\mathbf{r})$ over copies translated by multiples of L in each spatial direction:

$$
g_\Lambda(\mathbf{r}, L) = \frac{\Lambda^3}{8\pi^{3/2}} \prod_{\alpha \in \{x,y,z\}}
$$
$$
\times \sum_{q^{(\alpha)}=-\infty}^{\infty} \exp \left(-\Lambda^2 (r^{(\alpha)} - Lq^{(\alpha)})^2 / 4 \right). \quad (7)
$$

B. Variational method framework

The variational method provides a systematically-improvable approach to bounding the ground (and excited) state energies of quantum systems; given any wavefunction ansatz $\Psi_h (\mathbf{x})$ for a state h defined over coordinates $\mathbf{x}$, the ground-state energy E_h is bounded as

$$
E_h \leq \mathcal{E}[\Psi_h] = \frac{\int \Psi_h(\mathbf{x})^* H \Psi_h(\mathbf{x}) \, d\mathbf{x}}{\int \Psi_h(\mathbf{x})^* \Psi_h(\mathbf{x}) \, d\mathbf{x}}. \quad (8)
$$

A wavefunction ansatz that depends on some number of free parameters may be varied over those parameters to determine an optimal bound within that ansatz class.

One approach to the variational method that has been successfully applied to the study of nuclear systems in a finite volume within the framework of pionless effective field theory is the stochastic variational method (SVM) [27]. In this approach, a wavefunction is generated constructively through the iterative addition of stochastically-proposed terms, and a generalised eigenvalue problem is solved to optimise the linear combination of the proposed terms. In particular, this approach has been applied in a finite volume in Refs. [25, 26] using a basis of correlated Gaussian terms. Here, the same wavefunction ansatz is considered, but is optimised using an alternative to the stochastic optimisation procedure that is based on differentiable programming (detailed in Sec. II C).

The Gaussian wavefunction ansatz used in this work is based on the approximation that the spatial and spin-isospin wavefunctions for nuclear states can be factorised, with spatial wavefunctions constructed as linear combinations of appropriately symmetrised Gaussians.[1] As

---

[1] Although the factorisation of the spatial and spin wavefunctions

also used in Refs. [25, 26], a trial wavefunction of this form satisfying the periodic boundary conditions of a finite spatial volume can be expressed for some nucleus h as

$$\Psi_h^{(N)}(\mathbf{x}) = \sum_{j=1}^{N} c_j \Psi_L^{\text{sym}}(A_j, B_j, \mathbf{d}_j; \mathbf{x}) |\chi_h\rangle, \qquad (9)$$

where the sum runs over the N terms included in the trial wavefunction, the c_j, $j \in \{1, \ldots, N\}$, are numerical coefficients, $|\chi_h\rangle$ denotes an appropriately normalised antisymmetric spin-flavour wavefunction for the n-body state h, $\mathbf{x} = (\mathbf{r}_1, \ldots, \mathbf{r}_n)$ denotes the collected spatial coordinates of the n nucleons, and the A_j, B_j and $\mathbf{d}_j$ denote collected parameters of the jth spatial wavefunction Ψ_L^{sym} included in the sum (whose dependence on h is suppressed). To obtain an optimal representation of the wavefunction with a given number of terms, the values of the parameters c_j and those encoded in A_j, B_j and $\mathbf{d}_j$ are optimised as described further in Sec. II C.

Explicitly, the symmetrised spatial wavefunction Ψ_L^{sym} is constructed from Gaussian components for each Cartesian direction α:

$$\Psi_\infty^{(\alpha)}(A^{(\alpha)}, B^{(\alpha)}, \mathbf{d}^{(\alpha)}; \mathbf{x}^{(\alpha)}) = \exp\left[-\frac{1}{2}\mathbf{x}^{(\alpha)T} A^{(\alpha)} \mathbf{x}^{(\alpha)} \right.$$
$$\left. -\frac{1}{2}(\mathbf{x}^{(\alpha)} - \mathbf{d}^{(\alpha)})^T B^{(\alpha)}(\mathbf{x}^{(\alpha)} - \mathbf{d}^{(\alpha)}) \right], \qquad (10)$$

where the αth Cartesian components of the spatial coordinates of each particle are collected in the n-component vector $\mathbf{x}^{(\alpha)}$. The $n \times n$ matrices $A^{(\alpha)}$ and $B^{(\alpha)}$ are symmetric, containing $n(n-1)/2$ real parameters, and diagonal, with n real parameters, respectively, and $\mathbf{d}^{(\alpha)}$ is an n-component real-valued vector. This wavefunction can be made periodic in a cubic volume of finite spatial extent L by implementing a sum over copies shifted in each Cartesian direction by integer multiples of L [28]:

$$\Psi_L^{(\alpha)}\left(A^{(\alpha)}, B^{(\alpha)}, \mathbf{d}^{(\alpha)}; \mathbf{x}^{(\alpha)}\right) =$$
$$\sum_{\mathbf{b}^{(\alpha)}} \Psi_\infty^{(\alpha)}(A^{(\alpha)}, B^{(\alpha)}, \mathbf{d}^{(\alpha)}; \mathbf{x}^{(\alpha)} - \mathbf{b}^{(\alpha)} L), \quad (11)$$

where $\mathbf{b}^{(\alpha)}$ is an n-component vector with components $b_k^{(\alpha)} \in \mathbb{Z}$. The finite-volume wavefunctions for each Cartesian direction α can be combined to define the complete three-dimensional finite-volume wavefunction

$$\Psi_L(A, B, \mathbf{d}; \mathbf{x}) = \prod_{\alpha \in \{x,y,z\}} \Psi_L^{(\alpha)}\left(A^{(\alpha)}, B^{(\alpha)}, \mathbf{d}^{(\alpha)}; \mathbf{x}^{(\alpha)}\right), \qquad (12)$$

is a crude approximation for larger nuclei, the goal of the present work is to explore the effectiveness of the differentiable programming approach in representing nuclear states in comparison to the stochastic variational method. As such, the same approximation is used as in Ref. [26].

where the parameters $A^{(\alpha)}$, $B^{(\alpha)}$, and $\mathbf{d}^{(\alpha)}$ for each Cartesian direction are combined into the quantities A, B and $\mathbf{d}$. Finally, a finite-volume wavefunction that is also symmetric under particle exchange can be constructed by explicitly symmetrising with respect to permutations of the rows and columns of $A^{(\alpha)}$ and $B^{(\alpha)}$ and of the rows of $\mathbf{d}^{(\alpha)}$, for all Cartesian components α. Denoting the set of all such permutations as $\mathcal{P}$, a symmetric wavefunction ansatz can thus be expressed as

$$\Psi_L^{\text{sym}}(A, B, \mathbf{d}; \mathbf{x}) = \sum_{\mathcal{P}} \Psi_L(A_{\mathcal{P}}, B_{\mathcal{P}}, \mathbf{d}_{\mathcal{P}}; \mathbf{x}), \qquad (13)$$

where $A_{\mathcal{P}}$, $B_{\mathcal{P}}$ and $\mathbf{d}_{\mathcal{P}}$ are the permuted forms of the relevant matrices and vectors.

A particular advantage of this class of trial wavefunctions is that the integrals needed to compute the normalisation and Hamiltonian matrix elements that appear in the ground-state energy bound of Eq. (8) can be performed analytically, as detailed in Ref. [26]. As also discussed in Ref. [26], these Gaussian-based wavefunctions are able to represent finite-volume "scattering states", i.e., eigenstates above the two-particle threshold, for $N = 2$ systems, and the method does not rely on deeply-bound infinite volume states. The only restrictions on its applicability are that states that are integrated out of the pionless EFT, such as those involving pions, Δ-resonances and particle–anti-particle excitations, are not representable. These restrictions are similar to those in the Lüscher quantisation condition approach [29] where the partial-wave expansion of scattering amplitudes must be truncated and the presence of inelastic thresholds limits applicability.

C. Variational optimisation by differentiable programming

To achieve effective bounds on the ground-state energies E_h of various nuclear systems, h, the trial wavefunction $\Psi_h^{(N)}(\mathbf{x})$ defined in Eq. (9) is optimised using a differentiable programming approach combined with solution of a GEVP. Differentiable programming is a programming paradigm in which the computational flow of a program can be explicitly differentiated with respect to its parameters, thereby allowing gradient-descent optimisation of those parameters. The approach is widely used as a backbone of machine learning tools [30] and has been applied to variational problems in quantum many-body physics [31–34] and quantum technology [35–37].

In the current application, the differentiable programming approach is applied to ground-state energy minimisation through a two-stage procedure:

1. Differentiable program (DP) block: optimisation of N'-element wavefunctions with fixed LECs:

   - Values of the LECs C_0, C_1, and (for $n \geq 3$ body states) D_0 are chosen;

(a) Differentiable programming (DP) block

(b) Generalised eigenvalue problem (GEVP) block.
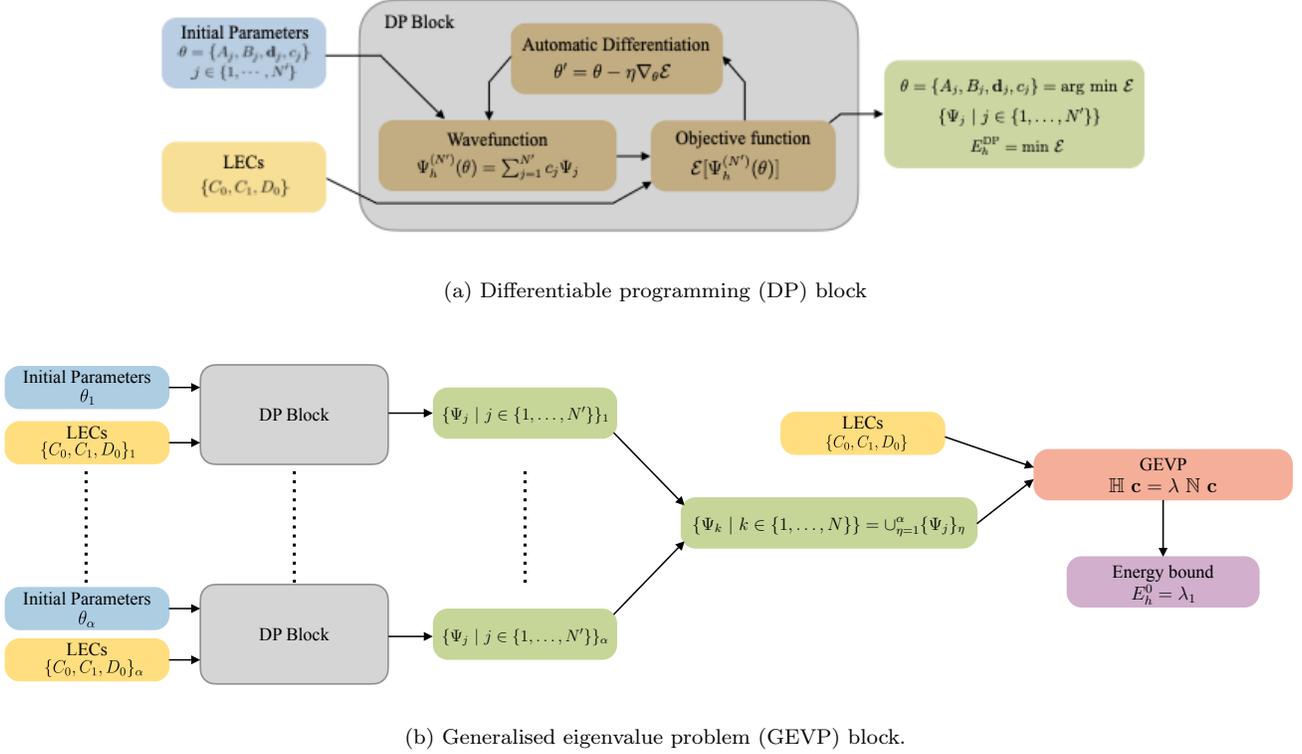
Figure 1. Diagrammatic representation of the wavefunction optimisation procedure used in this work. (a) Differentiable programming block: for fixed LECs and a random initialisation, automatic differentiation (defined in the figure, where η denotes the self-adaptive learning rate) is used to optimise the parameters of an N'-term Gaussian wavefunction ansatz. (b) GEVP block: basis elements obtained from differentiable programming blocks constructed with different initialisations and/or LECs are combined to form a larger basis; the GEVP as defined in Eq. (14) is solved to determine the optimal energy bound E_h^0, the smallest eigenvalue, for a given set of LECs.

- The free parameters c_j, and those encoded in A_j, B_j and $\mathbf{d}_j$, for $j \in \{1, \ldots N'\}$ for a N'-term Gaussian wavefunction $\Psi_h^{(N')}(\mathbf{x})$ are initialised randomly (details of the choice of initialisation for the numerical study detailed in Sec. III are provided in Appendix B);

- The gradient of the ground-state energy bound provided by the trial wavefunction with respect to the free parameters is computed, and the minimum is approached via gradient descent to optimise c_j and the parameters in A_j, B_j and $\mathbf{d}_j$ (details of the gradient computation and descent method are presented in Appendices A and B).

2. GEVP block: N-element wavefunction construction:

- A set of α N'-element wavefunctions, possibly optimised using different LECs C_0, C_1, D_0, and different initialisations, but defined for the same quantum numbers (number of particles and spin-flavour structure) and finite spatial extent, L, are constructed through α in-

dependent DP blocks;

- For a fixed choice of LECs, the linear combination of the optimised Gaussian wavefunction components (i.e., the $N = \alpha N'$ Gaussians with each optimised choice of A, B, $\mathbf{d}$) is optimised by solving the GEVP

$$\mathbb{H}\,\mathbf{c} = \lambda \mathbb{N}\,\mathbf{c}, \qquad (14)$$

for the eigenvalues $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_N$ and eigenvectors $\mathbf{c} = (c_1, \ldots, c_N)^T$ which contain the coefficients in Eq. (9). The matrices $\mathbb{H}$ and $\mathbb{N}$ have matrix elements

$$[\mathbb{N}]_{ij} \equiv \int \Psi_i(\mathbf{x})^* \Psi_j(\mathbf{x})\,d\mathbf{x}, \qquad (15)$$

$$[\mathbb{H}]_{ij} \equiv \int \Psi_i(\mathbf{x})^* \langle \chi_h | H | \chi_h \rangle \Psi_j(\mathbf{x})\,d\mathbf{x}, \qquad (16)$$

using the compressed notation $\Psi_j(\mathbf{x}) \equiv \Psi_L^{\mathrm{sym}}(A_j, B_j, \mathbf{d}_j; \mathbf{x})$.

- The lowest eigenvalue, λ_1, of the GEVP solved for a given set of LECs, C_0, C_1, D_0, corresponds to an upper bound E_h^0 on the ground

state energy E_h for the given system with quantum numbers defined by h.

This approach, illustrated graphically in Fig. 1, has several advantages. First, the use of direct optimisation as opposed to stochastic selection of Gaussian basis elements enables wavefunction representations of comparable quality to be obtained with far fewer terms, as demonstrated in numerical experiments detailed in Sec. III A. Second, this particular (sequential) optimisation approach enables the efficient construction of N-term wavefunctions by combining the Gaussian basis elements obtained by optimising systems with fewer terms; this is computationally efficient since the cost of directly optimising an N-term wavefunction grows quadratically with N.[2] Simultaneously, by combining wavefunctions optimised for different choices of LECs, this approach enables the construction of a combined set of Gaussian terms that can provide efficient wavefunction representations across a range of values of the LECs. With such a basis defined, constraining the LECs to match the FVEFT to LQCD calculations of nuclear states in the same finite volume is straightforward; computing the energy bound as a function of the LECs simply amounts to repeating the GEVP for choices of the LECs within a range of interest (involving no additional differentiable programming optimisation).

III. RESULTS

The differentiable programming approach described in Sec. II C is applied to the determination of ground-state energies of $A \in \{2, 3, 4, 5, 6\}$ nuclear systems, via optimisation of spatial nuclear wavefunctions with the relevant LECs tuned to match the results of LQCD calculations for $A \in \{2, 3\}$. As was previously investigated in the SVM in Refs. [25, 26], the differentiable programming method can be used to extrapolate existing LQCD results for light nuclei to infinite volume. However, the more efficient representation provided by the DP wavefunctions also allows extrapolation of the LQCD results to systems with larger A.

A. Illustration of differentiable programming optimisation

This section provides a numerical illustration of the differentiable programming approach of Sec. II C. The following examples demonstrate each step of the method,

while discussion of several more technical aspects of the approach such as the initialisation of the free parameters, the schedule of optimisation ('training'), and the convergence criteria used in the applications in the following sections, are left to Appendix B.

As discussed in Sec. II C, the differentiable programming optimisation procedure proceeds via DP blocks and GEVP blocks. The DP block step yields an optimised N'-term wavefunction at fixed LECs; Fig. 2 provides an example of the convergence of this optimisation with N'. In particular, the figure illustrates the bound on the binding energy of the $A = 2$ deuteron ($h = d$) system achieved through a DP block optimisation (i.e., $\Delta E_d \equiv E_d - 2E_p$, with $E_d = E_d^{\mathrm{DP}}$) for a fixed choice of the relevant LEC $C_S = -132$ MeV $\cdot$ fm^3, and spatial volume $L = 4.5$ fm (these parameters are approximately in the centre of the ranges that are used in the application of the method in the following sections). Clearly, an improved bound on the binding energy is achieved with increasing N', although this improvement need not be monotonic since the optimisation is performed from a new initialisation for each N'.[3] Different initialisation seeds typically yield consistent results for $N' \gtrsim 4$. The figure also shows the result of the SVM optimisation method from Ref. [26], demonstrating that the DP optimisation procedure provides a far more efficient description of the ground state in terms of the number of parameters that are required; for most initialisation seeds, the DP wavefunctions with $N' \gtrsim 4$ outperform the $N' = 100$ term wavefunction of Ref. [26].

The second step of optimisation combines α sets of N' Gaussian functions determined in independent DP blocks through a GEVP block to determine the optimal linear combination of all $N = \alpha N'$ Gaussian functions. As α increases, the bound on the ground-state energy of the system necessarily improves. Figure 3 shows the binding energy of the deuteron with $C_S = -132$ MeV $\cdot$ fm^3 and $L = 4.5$ fm obtained from GEVP-optimised combinations of $\alpha \in \{1, \ldots, 16\}$ sets of $N' = 4$ Gaussian functions. Results are shown for 25 different initialisations. For $\alpha \geq 8$, all seeds yield values within 0.1% of the minimum energy achieved with $\alpha = 16$ groups of Gaussians (with any seed), and at $\alpha = 8$, more than half of the seeds yield results within 0.05% of that minimum.

While Fig. 3 illustrates the results of a GEVP block combining wavefunctions from several DP blocks optimised at the same set of LECs, a potentially useful alternative is to combine sets of Gaussian functions from DP blocks optimised at different choices of LECs. This produces a set of functions that should be better able to represent the eigenstates of the Hamiltonian across

---

[2] In particular, the cost of optimisation of an n-body state with N terms scales as $O(N^2 n! n^3)$. While this complete optimisation would in principle outperform the sequential approach used here for a fixed number of terms, the sequential approach is superior for a fixed computational budget, scaling as $O(\alpha N'^2 n! n^3)$.

[3] An alternate approach in which an optimised $N' - M$ term wavefunction is used to build an N' term wavefunction by only optimising the parameters associated with the M new Gaussian functions (through DP) and the linear coefficients (through GEVP) could also be applied, and would be monotonic by construction.
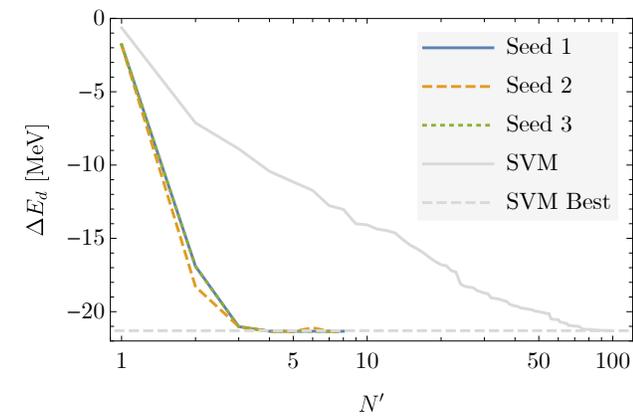
(a)
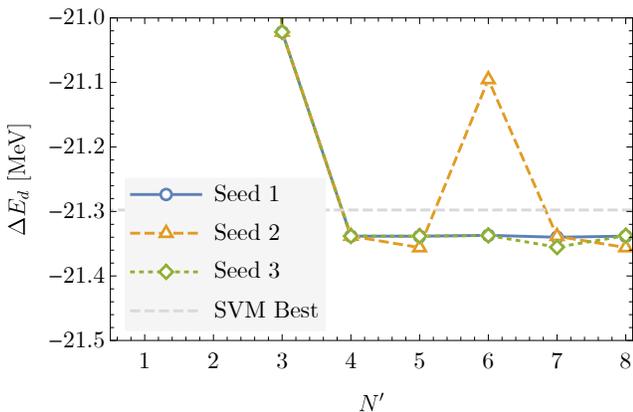
(b)

Figure 2. The bounds on the binding energy of the deuteron obtained from DP blocks as a function of the number of Gaussian functions, N', included in the optimisation. Three different random initialisations are used (corresponding to the solid blue, dashed orange, and dotted green curves) for fixed values of the relevant two-body LEC $C_S = -132$ MeV $\cdot$ fm^3, and for $L = 4.5$ fm. In the upper panel (a), the results are compared with the SVM results of Ref. [26] evaluated at the same C_S and L, shown as a function of the number of Gaussian functions (solid grey line). The lower panel (b) shows the DP results with a different scale. As detailed in the text, an independent optimisation is performed for each value of N', so the behaviour need not be monotonic. The dashed horizontal line in both panels shows the best result obtained with the SVM method with $N' = 100$.

a range of values of the LECs, allowing energy bounds to be evaluated as a function of the LECs without additional DP optimisations. For α sets of N' Gaussian wavefunctions optimised in DP blocks, the quantity

$$\delta_h^{\alpha, N'} = \frac{\Delta E_h^{[N']} - \Delta E_h^{[\alpha \times N']}}{\left| \Delta E_h^{[\alpha \times N']} \right|} \quad (17)$$

can be defined to quantify the relative improvement of the combined $\alpha \times N'$-term wavefunction (yielding a
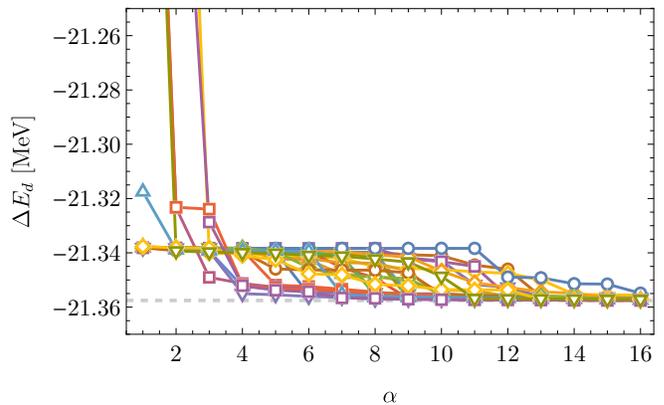
Figure 3. Bounds on the binding energy of the deuteron at $C_S = -132$ MeV $\cdot$ fm^3 and $L = 4.5$ fm as a function of the number of groups of $N' = 4$ Gaussian wavefunctions that are combined through a GEVP block. Each of the 25 curves shows bounds obtained using a different random sampling of groups of 4 Gaussians from a total of 36 groups, each optimised from a different random initialisation. The dashed grey line shows the tightest bound on the binding energy achieved by any of the optimisations.

bound $\Delta E_h^{[\alpha \times N']}$ on the binding energy) over the N' term wavefunction (yielding the bound $\Delta E_h^{[N']}$) at a given LEC value. Fig. 4 shows this quantity for the deuteron at $L = 4.5$ fm, where sets of $N' = 4$ Gaussian functions optimised at four choices of C_S are combined in a GEVP block. By construction, GEVP-optimisation of the superset of 16 Gaussians provides a tighter bound on the binding energy across all LECs in the relevant range, improving the bound by $\lesssim 0.1\%$ even at the LEC values where the individual DP blocks were optimised.

B. Finite-Volume Calculation of Two-body and Three-body Systems

In order to determine the two and three-body LECs in the FVEFT Hamiltonian, C_0 and C_1 (or equivalently, C_S and C_T) and D_0, wavefunction optimisations are performed using the approach of Sec. II C for the deuteron, dineutron, and ^{3}He systems in each of three spatial volumes where LQCD calculations have been performed [4]. In this work, a single EFT cutoff corresponding to $r_0 = 0.2$ fm is used, as Refs. [25, 26] have previously demonstrated the cutoff-independence of the ground-state energies.

Fig. 5 shows the binding energy of the deuteron as a function of the LEC C_S for $L \in \{3.4, 4.5, 6.7\}$ fm. The dependence of the binding energy in each volume on C_S is obtained by solving the GEVP using a 32-dimensional basis of Gaussians, with $\alpha = 8$ sets of $N' = 4$ Gaussians, two optimised from different initialisations at each $C_S \in \{-134, -129, -124, -119\}$ MeV $\cdot$ fm^3. These choices of N' and α achieve a balance between represen-
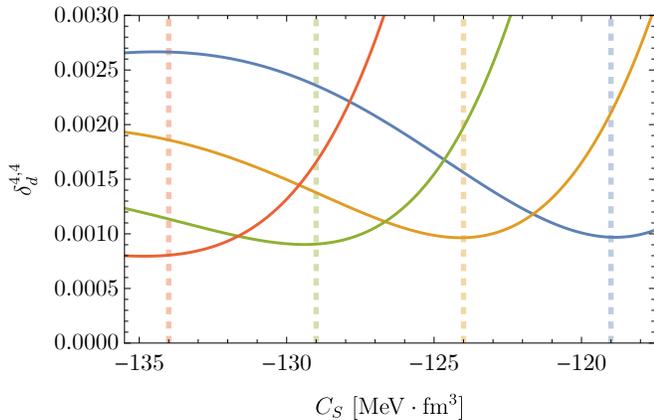
Figure 4. Fractional difference between the binding energy of the deuteron at $L = 4.5$ fm computed via GEVP from $\alpha = 4$ DP blocks optimised at $C_S \in \{-134, -129, -124, -119\}$ MeV $\cdot$ fm^3 with $N' = 4$, and the results of GEVP using each block separately. Each curve corresponds to $\delta_d^{4,4}$ (Eq. (17)) computed based on the DP block optimised at the value of C_S indicated by the colour-matched vertical dotted line).

Figure 5. Binding energy of the deuteron as a function of the two-body LEC C_S. Each curve is obtained by solving the GEVP with various C_S values using a 32-dimensional set of Gaussian wavefunctions optimised for each volume. Each set is composed of $\alpha = 8$ sets of $N' = 4$ DP blocks, with two blocks optimised from different initalisations for each $C_S \in \{-134, -129, -124, -119\}$ MeV $\cdot$ fm^3. The horizontal bands show the binding energies determined in each volume in the LQCD calculations of Ref. [4]. The intersection of each curve with the horizontal band of the same colour constrains the allowed values of C_S through χ^2-minimisation.

	Ref. [26] (SVM)	This work (DP)
C_0	$-131(2)$	$-131.0(21)$
C_1	$-2(1)$	$-1.7(8)$
C_S	$-133(2)$	$-132.7(27)$
C_T	$-126(2)$	$-125.8(20)$
D_0	$17(2)$	$20.0(24)$

TABLE I. LECs in the EFT Lagrangian for a cutoff $r_0 = 0.2$ fm. $C_{S,T,0,1}$ and D_0 are quoted in units of MeV $\cdot$ fm^3 and MeV $\cdot$ fm^6, respectively. The second column provides a comparison with the results obtained in Ref. [26] using the SVM, while the third column presents the results of this work obtained via the DP approach.

tational flexibility and computational cost and are motivated by the observations illustrated in Sec. III A. In particular, taking $N' > 4$ typically does not improve the bound achieved by a single DP block, and the combination of $\alpha = 8$ DP blocks with $N' = 4$ yields results within a fraction of a percent of the best result obtained by continuing to increase the number of blocks included; this difference is negligible in comparison with the uncertainties of the LQCD results used to match the LECs. The optimisation procedure and convergence criteria are detailed in Appendix B. The same optimisations give the dependence of the dineutron binding energy on the LEC C_T; C_S and C_T can thus be obtained by χ^2-minimisation of the difference between the optimised binding energies and the LQCD results of Ref. [4] for the deuteron and dineutron, respectively. Fit results are shown in Table I and are consistent within uncertainties with those obtained in Ref. [26] using the SVM approach to wavefunction optimisation, matched to the same LQCD results.

Having determined the two-body couplings, the analogous procedure can be applied to determine the three-body interaction coefficient, D_0. The GEVP is solved using a 32-dimensional basis of three-body Gaussians with $\alpha = 8$ sets of $N' = 4$ Gaussians, two optimised from different initialisations at each $D_0 \in \{17.8, 18.8, 19.7, 20.6\}$ MeV $\cdot$ fm^6, with the optimised value of $C_{^3\text{He}} = C_0 - C_1$ fixed. The three-body binding energy is shown as a function of D_0 in Fig. 6 and the value of the coupling determined by χ^2-minimisation of results at all three volumes is presented in Table. I.

C. Finite-Volume Calculation of ^{4}He

With all of the leading-order couplings in the EFT Lagrangian determined, the DP approach can be used to compute the ground-state energies of larger systems. In particular, an upper bound on the ground-state energy of ^{4}He is computed; with the LECs fixed, $\alpha = 8$ $N' = 4$ DP blocks optimised from different initialisations are combined via a GEVP block in each of the three spatial volumes in which LQCD calculations have been performed. After optimisation, the uncertainties in the two and three-body LECs are propagated into the estimate of the ^{4}He binding energy by solving the GEVP for each optimised set of Gaussian wavefunctions with the LECs varied within their uncertainty ranges. Table. II shows a comparison between the resulting ^{4}He binding energy and LQCD results for the binding in each finite
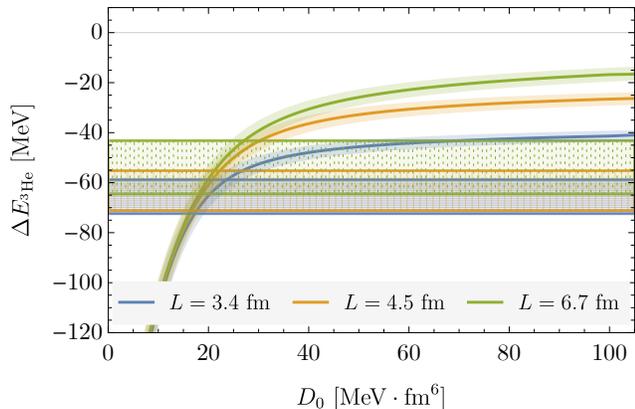
Figure 6. Binding energy of ^{3}He as a function of the three-body LEC D_0, with the relevant two-body coupling fixed to $C_{^3\text{He}} = C_0 - C_1 = 129(2)$ MeV $\cdot$ fm^3. The curves are obtained by solving the GEVP for a 32-dimensional set of Gaussian wavefuntions optimised for each volume. Each set is composed of $\alpha = 8$ $N' = 4$ DP blocks, with two blocks optimised from different initalisations for each $D_0 \in \{16.9, 18.1, 19.4, 20.6\}$ MeV $\cdot$ fm^6. The shaded region for each curve is propagated from the uncertainty in $C_{^3\text{He}}$. The horizontal bands show the binding energies determined in each volume in the LQCD calculations of Ref. [4].

L [fm]	LQCD [4] [MeV]	This work [MeV]
3.4	115(23)	114(13)
4.5	107(25)	109(15)
6.7	107(24)	108(15)

TABLE II. Finite-volume binding energy for ^{4}He obtained for three different volumes as described in the text, with LECs C_0, C_1 and D_0 determined by matching to the two and three-body finite-volume LQCD results of Ref. [4]. The second column lists the LQCD results for $\Delta E_{^4\text{He}}$ computed in the same reference for comparison.

volume. Clearly, the EFT with fixed coefficients produces estimates of the ^{4}He binding energy that are consistent with the LQCD calculations. Having verified the consistency, an alternative strategy is to use the full set of $h \in \{d, nn, ^3\text{He}, ^4\text{He}\}$ binding energies from LQCD to further constrain the two and three-body LECs. However, given the large uncertainties on the LQCD determinations of the ^{4}He energy, χ^2-optimisation leads to values of the two and three-body LECs that are identical to those determined from the two and three-body systems alone.

In principle, the DP method can be used to determine finite-volume energies of still larger nuclei. However, there are currently no LQCD results to compare to for larger systems, and the scaling of the approach with A at finite volume is sufficiently poor that such calculations are numerically demanding. Instead, it is natural to consider the infinite-volume binding energies of larger nuclei, as detailed in the next subsection.
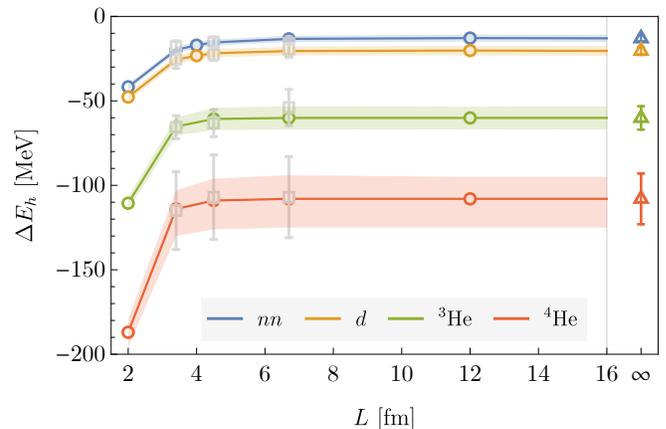
Figure 7. Binding energies of $h \in \{nn, d, ^3\text{He}, ^4\text{He}\}$ states in different volumes (open circles) and in infinite volume (open triangles). For each state at each of the volumes indicated by the open circles, $\alpha = 8$ sets of $N' = 4$ Gaussian wavefunctions are generated and optimised from different initialisations at the best fit values of the LECs $C_{0,1}$ and D_0 (the coloured lines are linear interpolations to guide the eye). The bands result from propagation of the uncertainties in the LECs. The gray squares show the results of the LQCD calculations of Ref. [4].

D. Volume Dependence and Infinite-Volume Calculation

Having performed the finite-volume matching to determine the LECs of the EFT, the DP approach can be used to study the volume-dependence of the binding energies for $h \in \{d, nn, ^3\text{He}, ^4\text{He}\}$, as well as to determine the infinite-volume binding energies of these and other nuclear states. Fig. 7 shows the volume-dependence of the binding energies of the four systems, with optimisations based on $\alpha = 8$ sets of $N' = 4$ Gaussian wavefunctions performed for $L \in \{2, 3.4, 4, 4.5, 6.7, 12\}$ fm, and also at infinite volume. The infinite-volume results are compiled in Table III.

Since the DP approach is more efficiently able to represent ground-state wavefunctions than the SVM method, it is feasible to extend calculations to larger nuclei in infinite volume. In particular, calculations have been performed for $^5_\Lambda$He ($J = 1/2$) and $^6_{\Lambda\Lambda}$He ($J = 0$) in which the spin-flavour structure is such that the simplest configuration has spatial and spin-flavour wavefunctions factorising as in Eq. (9). Note that at the $SU(3)_f$-symmetric quark masses used in the LQCD calculations of Ref. [4] to which the FVEFT calculation is matched, the Λ baryon is degenerate with the proton and neutron, but is not Pauli-blocked from being at zero orbital angular momentum. For this proof-of-principle study, it is assumed that the two and three-body interactions between nucleons and Λ-baryons are the same. Figure 8 and Table III summarise the results of this work as well as previous EFT matching calculations from Refs. [25, 26, 38]. The $A = 4$ EFT calculations are post-dictions of the infinite-

System	I	S	J	Ref. [4]	Ref. [25]	Ref. [26]	Ref. [38]	This work
nn	1	0	0	16(4)	14(2)	12(2)	16(2)	12(2)
d	0	0	1	20(5)	20(2)	20(3)	20(5)	20(3)
^{3}He	1/2	0	1/2	54(11)	58(5)	60(7)	54(11)	60(7)
^{4}He	0	0	0	107(24)	113(10)	–	89(36)	108(15)
^{5}He	1/2	0	3/2	–	–	–	98(39)	–
$^5_\Lambda$He	0	-1	1/2	–	–	–	–	162(24)
^{6}He	1	0	0	–	–	–	122(50)	–
$^6_{\Lambda\Lambda}$He	0	-2	0	–	–	–	–	215(32)

TABLE III. Infinite-volume binding energies for various nuclear systems obtained in this work, compared with the extrapolations in Refs. [4] (LQCD), [25] (SVM), and [26] (SVM), as well as the infinite-volume EFT calculations of Ref. [38]. The isospin, I, strangeness, S, and spin, J, of each state is also listed.
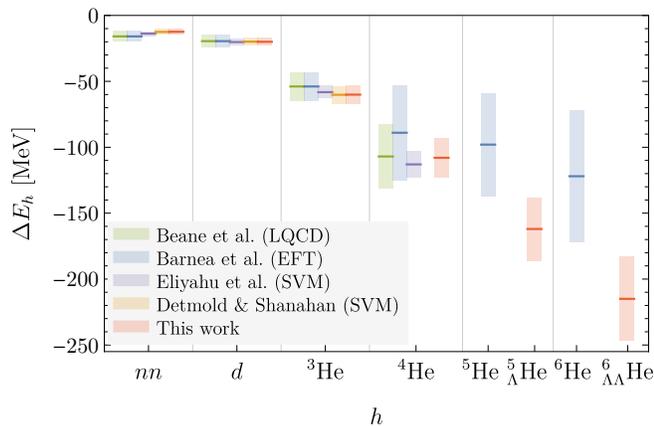
Figure 8. Comparisons of extrapolated infinite-volume binding energies obtained in different approaches: LQCD from Ref. [4], infinite-volume EFT matching from Ref. [38], and finite-volume EFT matching from Refs. [25, 26] and the current work. In each case, the EFT matching is performed to the same LQCD calculations of $h \in \{nn, d, ^3\text{He}\}$ systems.

volume-extrapolated LQCD results of Ref. [4], while the $A \in \{5, 6\}$ results are predictions that could be tested by future LQCD calculations.

IV. SUMMARY AND OUTLOOK

In this work, differentiable programming and a generalised eigenvalue problem have been used to optimise the ground state wavefunctions of nuclei in FVEFT matched to LQCD binding energies. Using sets of correlated Gaussian wavefunctions representing A-nucleon states for $2 \leq A \leq 6$ in both finite and infinite volumes, it was shown that this approach provides a considerably more efficient representation of these states than that obtained in previous work using the stochastic variational method and is able to scale to larger system size for fixed computational resources.

Ongoing work to extend this approach by coupling spatial and spin wavefunctions used for the nuclear states will allow more physical systems to be addressed including p-shell nuclei and hypernuclei. Since these approaches can also provide accurate representations of finite-volume excited states, a more detailed matching to the low energy excitation spectra of two-nucleon systems in LQCD, for example those presented in Ref. [39], will also allow more precise constraints on the LECs of the nuclear EFT, including those that occur at next-to- and next-to-next-to-leading order in the EFT power-counting.

Pionless EFT is particularly powerful at the large quark masses used in existing LQCD calculations of nuclei. However, as the masses used in such calculations become closer to the physical light-quark masses, extensions of the finite-volume matching approach presented here to chiral EFTs that include explicit pionic degrees of freedom will likely be important. Alternative finite-volume many-body methods such as quantum Monte-Carlo [40, 41] and nuclear lattice EFT [42, 43] are promising approaches.

Appendix A: Integral definitions and automatic differentiation

This section presents explicit analytic formulae for the matrix elements required to compute variational energy bounds with Gaussian wavefunctions, and provides details of the computation of the derivatives of those matrix elements as required for the DP optimisation procedure.

The normalisation matrix $\mathbb{N}$ (defined in Eq. (15)), with elements labelled by the i-th and the j-th terms of wavefunctions, can be computed as

$$
\begin{aligned}
[\mathbb{N}]_{ij} &\equiv \int \Psi_L^{\mathrm{sym}}(A_i, B_i, \mathbf{d}_i; \mathbf{x})^* \Psi_L^{\mathrm{sym}}(A_j, B_j, \mathbf{d}_j; \mathbf{x}) d\mathbf{x} \\
&= \sum_{\mathcal{P}, \mathcal{P}'} \prod_{\alpha \in \{x,y,z\}} \sqrt{\frac{(2\pi)^n}{\mathrm{Det}[C_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)}]}} \sum_{\mathbf{b}^{(\alpha)}}^{|\mathbf{b}^{(\alpha)}| \leq \tilde{b}} \exp\left[-\frac{1}{2} \Omega_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)} \right],
\end{aligned}
\tag{A1}
$$

where $\mathcal{P}$ and $\mathcal{P}'$ denote permutations over the n bodies in each wavefunction and $\Omega_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)}$ and $C_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)}$ are quantities defined in Ref. [26] that depend on the parameters of the wavefunctions. The finite-volume sum is controlled by the integer cutoff $\tilde{b}$.

The matrix representation of the Hamiltonian $\mathbb{H}$ (Eq. (16)) can be broken up into three parts as

$$
\mathbb{H} = \mathbb{K} + \mathbb{V}_2 + \mathbb{V}_3,
\tag{A2}
$$

where $\mathbb{K}$, $\mathbb{V}_2$, and $\mathbb{V}_3$ denote the kinetic energy, the two-body potential, and the three-body potential, respectively. These are given by

$$
\begin{aligned}
[\mathbb{V}_2]_{ij} &\equiv C \sum_{a<b}^{n} \int \Psi_L^{\mathrm{sym}}(A_i, B_i, \mathbf{d}_i; \mathbf{x})^* g_\Lambda(\mathbf{x}_a - \mathbf{x}_b, L) \Psi_L^{\mathrm{sym}}(A_j, B_j, \mathbf{d}_j; \mathbf{x}) d\mathbf{x} \\
&= C \frac{\Lambda^3}{8\pi^{3/2}} \sum_{\mathcal{P}, \mathcal{P}'} \sum_{a<b}^{n} \prod_{\alpha \in \{x,y,z\}} \sqrt{\frac{(2\pi)^n}{\mathrm{Det}[C_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)}]}} \sqrt{\frac{\widetilde{C}_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)}}{\widetilde{C}_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)} + 2\rho}} \sum_{\mathbf{b}^{(\alpha)}}^{|\mathbf{b}^{(\alpha)}| \leq \tilde{b}} \exp\left[-\frac{1}{2} \Omega_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)} \right] \\
&\quad \times \sum_{q^{(\alpha)}=-\tilde{q}}^{\tilde{q}} \exp\left[-\frac{\rho \widetilde{C}_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)}}{\widetilde{C}_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)} + 2\rho} \left([(C_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)})^{-1} \cdot \mathbf{\Xi}^{(\alpha)}]_a - [(C_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)})^{-1} \cdot \mathbf{\Xi}^{(\alpha)}]_b - Lq^{(\alpha)} \right)^2 \right],
\end{aligned}
\tag{A3}
$$

where $C = \frac{1}{2}n(n-1)C_0 + 2(S_h(S_h+1) - \frac{3}{4}n)C_1$ for an n-body nucleus of spin S_h,

$$
\begin{aligned}
[\mathbb{V}_3]_{ij} &\equiv D_0 \sum_{a\neq b\neq c}^{\mathrm{cyc}} \int \Psi_L^{\mathrm{sym}}(A_i, B_i, \mathbf{d}_i; \mathbf{x})^* g_\Lambda(\mathbf{x}_a - \mathbf{x}_b, L) g_\Lambda(\mathbf{x}_b - \mathbf{x}_c, L) \Psi_L^{\mathrm{sym}}(A_j, B_j, \mathbf{d}_j; \mathbf{x}) d\mathbf{x} \\
&= D_0 \left(\frac{\Lambda^3}{8\pi^{3/2}} \right)^2 \sum_{\mathcal{P}, \mathcal{P}'} \sum_{a\neq b\neq c}^{\mathrm{cyc}} \prod_{\alpha \in \{x,y,z\}} \sqrt{\frac{(2\pi)^n}{\mathrm{Det}[\widehat{C}_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)}]}} \exp\left[-\frac{1}{2} \left(\mathbf{d}_{i\mathcal{P}}^{(\alpha)} \cdot B_{i\mathcal{P}}^{(\alpha)} \cdot \mathbf{d}_{i\mathcal{P}}^{(\alpha)} + \mathbf{d}_{j\mathcal{P}'}^{(\alpha)} \cdot B_{j\mathcal{P}'}^{(\alpha)} \cdot \mathbf{d}_{j\mathcal{P}'}^{(\alpha)} \right) \right] \\
&\quad \times \sum_{\mathbf{b}^{(\alpha)}}^{|\mathbf{b}^{(\alpha)}| \leq \tilde{b}} \exp\left[-\frac{1}{2} \left((L\mathbf{b}^{(\alpha)}) \cdot (A_{i\mathcal{P}}^{(\alpha)} + B_{i\mathcal{P}}^{(\alpha)}) \cdot (L\mathbf{b}^{(\alpha)}) + 2\mathbf{d}_{i\mathcal{P}}^{(\alpha)} \cdot B_{i\mathcal{P}}^{(\alpha)} \cdot (L\mathbf{b}^{(\alpha)}) - \mathbf{\Xi}^{(\alpha)} \cdot [\widehat{C}_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)}]^{-1} \cdot \mathbf{\Xi}^{(\alpha)} \right) \right] \\
&\quad \times \sum_{q^{(\alpha)}=-\tilde{q}}^{\tilde{q}} \exp\left[-\frac{L^2}{r_0^2}q^{(\alpha)2} + \frac{q^{(\alpha)2}L^2}{2r_0^4} \mathfrak{P}_v^{[a,b]} \cdot [\widehat{C}_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)}]^{-1} \cdot \mathfrak{P}_v^{[a,b]} + \frac{q^{(\alpha)}L}{r_0^2} \mathbf{\Xi}^{(\alpha)} \cdot [\widehat{C}_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)}]^{-1} \cdot \mathfrak{P}_v^{[a,b]} \right] \\
&\quad \times \sum_{t^{(\alpha)}=-\tilde{q}}^{\tilde{q}} \exp\left[-\frac{L^2}{r_0^2}t^{(\alpha)2} + \frac{t^{(\alpha)2}L^2}{2r_0^4} \mathfrak{P}_v^{[b,c]} \cdot [\widehat{C}_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)}]^{-1} \cdot \mathfrak{P}_v^{[b,c]} + \frac{t^{(\alpha)}L}{r_0^2} \mathbf{\Xi}^{(\alpha)} \cdot [\widehat{C}_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)}]^{-1} \cdot \mathfrak{P}_v^{[b,c]} \right] \\
&\quad \times \exp\left[\frac{t^{(\alpha)}q^{(\alpha)}L^2}{r_0^4} \mathfrak{P}_v^{[b,c]} \cdot [\widehat{C}_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)}]^{-1} \cdot \mathfrak{P}_v^{[a,b]} \right),
\end{aligned}
\tag{A4}
$$

$$[\mathbb{K}]_{ij} \equiv -\frac{1}{2M_N} \sum_{a=1}^{n} \int \Psi_L^{\text{sym}}(A_i, B_i, \mathbf{d}_i; \mathbf{x})^* \nabla_a^2 \Psi_L^{\text{sym}}(A_j, B_j, \mathbf{d}_j; \mathbf{x}) d\mathbf{x}$$

$$= \frac{1}{2M_N} \sum_{\mathcal{P},\mathcal{P}'} \sum_{\alpha \in \{x,y,z\}} \sqrt{\frac{(2\pi)^n}{\text{Det}[C_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)}]}} \sum_{\mathbf{b}^{(\alpha)}}^{|\mathbf{b}^{(\alpha)}| \leq \tilde{b}} \Theta_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)} \exp\left[-\frac{1}{2}\Omega_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)}\right]$$

$$\times \prod_{\substack{\beta \in \{x,y,z\} \\ \beta \neq \alpha}} \sqrt{\frac{(2\pi)^n}{\text{Det}[C_{i\mathcal{P};j\mathcal{P}'}^{(\beta)}]}} \sum_{\mathbf{b}^{(\beta)}}^{|\mathbf{b}^{(\beta)}| \leq \tilde{b}} \exp\left[-\frac{1}{2}\Omega_{i\mathcal{P};j\mathcal{P}'}^{(\beta)}\right], \qquad (A5)$$

where $\hbar = 1$ is used, $\sum_{a \neq b \neq c}^{\text{cyc}}$ indicates a sum over cyclic permutations of particles a, b and c, and the integer cutoff $\tilde{q}$ governs finite-volume effects in the interaction terms. The (wavefunction-parameter dependent) quantities $\Xi_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)}$, $\widetilde{C}_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)}$, $\widehat{C}_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)}$ and $\Theta_{i\mathcal{P};j\mathcal{P}'}^{(\alpha)}$, as well as the projection operators $\mathfrak{P}_v^{[a,b]}$ and $\mathfrak{P}_m^{[a,b]}$, are defined in Ref. [26].

The variational function $\mathcal{E}$ to be minimised (Eq. (8)) can be represented in terms of these matrices as

$$\mathcal{E}\left[\Psi_h^{(N)}(\theta)\right] = \frac{\mathbf{c} \cdot (\mathbb{K} + \mathbb{V}_2 + \mathbb{V}_3) \cdot \mathbf{c}}{\mathbf{c} \cdot \mathbb{N} \cdot \mathbf{c}}, \qquad (A6)$$

where $\mathbf{c} = (c_1, \ldots, c_N)^T$ collects the numerical coefficients (c_i of Eq. (A6)) parameterising $\Psi_h^{(N)}(\theta)$ as a linear combination of the Gaussian wavefunction terms $\Psi_L^{\text{sym}}(A_i, B_i, \mathbf{d}_i; \mathbf{x})$, and $\theta = \{\{A_i, B_i, \mathbf{d}_i, c_i\}, i \in \{1, \ldots N\}\}$. Storing the computational graph for Eq. (A6) and its gradients with respect to the parameters θ requires a large amount of memory. To reduce the memory usage, the chain rule is applied manually to compute the gradient of $\mathcal{E}$ as

$$\nabla_\theta \mathcal{E} = -\frac{\mathbf{c} \cdot (\mathbb{K} + \mathbb{V}_2 + \mathbb{V}_3) \cdot \mathbf{c}}{(\mathbf{c} \cdot \mathbb{N} \cdot \mathbf{c})^2} \nabla_\theta(\mathbf{c} \cdot \mathbb{N} \cdot \mathbf{c}) + \frac{1}{\mathbf{c} \cdot \mathbb{N} \cdot \mathbf{c}} (\nabla_\theta(\mathbf{c} \cdot \mathbb{K} \cdot \mathbf{c}) + \nabla_\theta(\mathbf{c} \cdot \mathbb{V}_2 \cdot \mathbf{c}) + \nabla_\theta(\mathbf{c} \cdot \mathbb{V}_3 \cdot \mathbf{c})). \qquad (A7)$$

The computation of $\nabla_\theta(\mathbf{c} \cdot \mathbb{X} \cdot \mathbf{c})$ for $\mathbb{X} \in \{\mathbb{N}, \mathbb{K}, \mathbb{V}_2, \mathbb{V}_3\}$ can be further broken up into a sum involving the gradient of each matrix element

$$\nabla_\theta(\mathbf{c} \cdot \mathbb{X} \cdot \mathbf{c}) = \sum_{i,j} \left(\nabla_\theta(c_i c_j)[\mathbb{X}]_{ij} + c_i c_j \nabla_\theta[\mathbb{X}]_{ij}\right). \qquad (A8)$$

Due to the permutation symmetry in this system, there are only $n! n(n-1)$ independent terms in the summation over permutations $\mathcal{P}$ and $\mathcal{P}'$ in $\mathbb{V}_2$ (Eq. (A3)) and $n! n(n-1)(n-2)$ terms in $\mathbb{V}_3$ (Eq. (A4)). Their gradients can be written as a sum of gradients on each independent term whose computational graph is discarded after its gradient is computed.

Appendix B: Numerical implementation details

A key component of the calculations presented here is the evaluation of the matrix elements $\mathbb{N}_{ij}$ and $\mathbb{H}_{ij}$ that enter both the DP and GEVP blocks. The numerical accuracy of these matrix elements is controlled by the integer cutoffs $\tilde{b}$ and $\tilde{q}$ used in the summations in Eqs. (A1)–(A5); for small values of these cutoffs, numerical instabilities appear with $\mathbb{N}$ potentially becoming non-positive-definite. Since the goal of the DP block is simply to produce trial wavefunctions, the accuracy criteria on matrix elements in the DP block is somewhat milder than in the GEVP block where a rigorous energy bound is sought. Consequently, $\tilde{b} = 15$ and $\tilde{q} = 6$ are used during automatic differentiation and $\tilde{b} = 30$ and $\tilde{q} = 12$ are chosen for the solution of the GEVP. These values avoid numerical stability issues but allow evaluation of the matrix elements for $N \in \{2, 3\}$-body systems at finite volume. For the optimisation and the solution of GEVP of four-body system binding energies $\tilde{b} = 8$ and $\tilde{q} = 3$ are chosen due to computational limitations.

The DP process depends on the initialisation of the wavefunction parameters θ. As in Ref. [28], the matrices A and B are generated from single-particle Gaussian widths d_a and two-body Gaussian widths d_{ab}. The particle displacement vectors $\mathbf{d}$ have components d_i and the weights of each wavefunction are written as $c_j = \tan \hat{c}_j$ to ensure both positive and negative values are accessed. The parameters d_a, d_{ab}, d_i and $\hat{c}_j$ are drawn from a normal distribution $N(1, 0.01)$, which in practice leads to stable results.

In the optimisation step in the DP block, a self-adaptive gradient descent method with a stepping clip is applied. For each step, the learning rate is increased by 20% if the energy decreases but is decreased by 60% if the energy

increases. Steps in which the energy increases are rejected. In addition, a maximum allowed step size is implemented for each parameter. The step in parameter $\theta_i \in \theta$ is $-\eta \partial_{\theta_i} \mathcal{E}$ if its absolute value is smaller than $f(\eta)$. Otherwise, the change in θ_i is $-\text{sgn}(\partial_{\theta_i} \mathcal{E}) f(\eta)$, where

$$f(\eta) = \begin{cases} 10^{-2}, & \eta > 0.2 \\ 10^{-3}, & 0.001 < \eta \le 0.2 \\ 10^{-4}, & \eta \le 0.001 \end{cases} \quad \text{(B1)}$$

Since the uncertainties of the LQCD results increase with A, the precision necessary in the variational optimisation for optimisation uncertainties to be sub-dominant relative to the LQCD uncertainties, or to the uncertainties propagated from the matching of the two and three-body LECs, decreases with A. For two and three-body systems, the training process is iterated until the relative change of the energy over the last 10 iterations is less than 10^{-5}. Under this condition, the relative differences between the upper bounds obtained using different seeds and the same set of LECs are less than 0.5% for all of the results presented in Sections III B–III D. For the four-body system, the relative changes in the last 10 steps of optimisation are less than 10^{-4} and the variations between initialisations are less than 1%. For the infinite volume calculations, the same convergence bounds hold for the two, three, and four-body systems. For five and six-body systems, the relative changes over the last ten steps are less than 10^{-3}, and the relative differences between results obtained with different initialisations are less than 2%.

[1] S. R. Beane *et al.* (Nplqcd), Phys. Rev. Lett. **106**, 162001 (2011), arXiv:1012.3812 [hep-lat].
[2] T. Inoue, N. Ishii, S. Aoki, T. Doi, T. Hatsuda, Y. Ikeda, K. Murano, H. Nemura, and K. Sasaki (Hal Qcd), Phys. Rev. Lett. **106**, 162002 (2011), arXiv:1012.5928 [hep-lat].
[3] S. R. Beane, E. Chang, W. Detmold, H. W. Lin, T. C. Luu, K. Orginos, A. Parreño, M. J. Savage, A. Torok, and A. Walker-Loud (Nplqcd), Phys. Rev. **D85**, 54511 (2012), arXiv:1109.2889 [hep-lat].
[4] S. R. Beane, E. Chang, S. D. Cohen, W. Detmold, H. W. Lin, T. C. Luu, K. Orginos, A. Parreño, M. J. Savage, and A. Walker-Loud (Nplqcd), Phys. Rev. **D87**, 34506 (2013), arXiv:1206.5219 [hep-lat].
[5] T. Yamazaki, K.-i. Ishikawa, Y. Kuramashi, and A. Ukawa, Phys. Rev. **D86**, 74514 (2012), arXiv:1207.4277 [hep-lat].
[6] T. Yamazaki, K.-i. Ishikawa, Y. Kuramashi, and A. Ukawa, Phys. Rev. **D92**, 14501 (2015), arXiv:1502.04182 [hep-lat].
[7] E. Berkowitz, T. Kurth, A. Nicholson, B. Joó, E. Rinaldi, M. Strother, P. M. Vranas, and A. Walker-Loud, Phys. Lett. **B765**, 285 (2017), arXiv:1508.00886 [hep-lat].
[8] K. Orginos, A. Parreño, M. J. Savage, S. R. Beane, E. Chang, and W. Detmold, Phys. Rev. **D92**, 114512 (2015), arXiv:1508.07583 [hep-lat].
[9] A. Francis, J. R. Green, P. M. Junnarkar, C. Miao, T. D. Rae, and H. Wittig, Phys. Rev. **D99**, 74505 (2019), arXiv:1805.03966 [hep-lat].
[10] T. Inoue (Hal Qcd), *Proceedings, 8th International Workshop on Chiral Dynamics (CD15): Pisa, Italy, June 29-July 3, 2015*, PoS **Cd15**, 20 (2016), arXiv:1511.04871 [hep-lat].
[11] N. Ishii, S. Aoki, T. Doi, T. Hatsuda, Y. Ikeda, T. Inoue, K. Murano, H. Nemura, and K. Sasaki (Hal Qcd), Phys. Lett. **B712**, 437 (2012), arXiv:1203.3642 [hep-lat].
[12] T. Inoue (Hal Qcd), *Proceedings, 13th International Conference on Hypernuclear and Strange Particle Physics (HYP 2018): Portsmouth Virginia, USA, June 24-29, 2018*, AIP Conf. Proc. **2130**, 20002 (2019), arXiv:1809.08932 [hep-lat].
[13] D. Kawai *et al.* (Hal Qcd), PTEP **2018**, 043b04 (2018), arXiv:1711.01883 [hep-lat].
[14] T. Iritani *et al.* (Hal Qcd), Phys. Lett. **B792**, 284 (2019), arXiv:1810.03416 [hep-lat].
[15] J. R. Green, A. D. Hanlon, P. M. Junnarkar, and H. Wittig, (2021), arXiv:2103.01054 [hep-lat].
[16] A. Parreño, P. E. Shanahan, M. L. Wagman, F. Winter, E. Chang, W. Detmold, and M. Illa (NPLQCD), Phys. Rev. D **103**, 074511 (2021), arXiv:2102.03805 [hep-lat].
[17] D. B. Kaplan, M. J. Savage, and M. B. Wise, Nucl. Phys. **B478**, 629 (1996), arXiv:nucl-th/9605002 [nucl-th].
[18] D. B. Kaplan, M. J. Savage, and M. B. Wise, Phys. Lett. **B424**, 390 (1998), arXiv:nucl-th/9801034 [nucl-th].
[19] D. B. Kaplan, M. J. Savage, and M. B. Wise, Nucl. Phys. **B534**, 329 (1998), arXiv:nucl-th/9802075 [nucl-th].
[20] U. van Kolck, Nucl. Phys. **A645**, 273 (1999), arXiv:nucl-th/9808007 [nucl-th].
[21] P. F. Bedaque, H. Hammer, and U. van Kolck, Phys. Rev. Lett. **82**, 463 (1999), arXiv:nucl-th/9809025.
[22] J.-W. Chen, G. Rupak, and M. J. Savage, Nucl. Phys. **A653**, 386 (1999), arXiv:nucl-th/9902056 [nucl-th].
[23] P. F. Bedaque, G. Rupak, H. W. Griesshammer, and H.-W. Hammer, Nucl. Phys. A **714**, 589 (2003), arXiv:nucl-th/0207034.
[24] P. F. Bedaque and U. van Kolck, Ann. Rev. Nucl. Part. Sci. **52**, 339 (2002), arXiv:nucl-th/0203055.
[25] M. Eliyahu, B. Bazak, and N. Barnea, Phys. Rev. C **102**, 044003 (2020), arXiv:1912.07017 [nucl-th].
[26] W. Detmold and P. E. Shanahan, Phys. Rev. D **103**, 074503 (2021), arXiv:2102.04329 [nucl-th].
[27] K. Varga and Y. Suzuki, Phys. Rev. C **52**, 2885 (1995), arXiv:nucl-th/9508023.
[28] X. Y. Yin and D. Blume, Phys. Rev. A **87**, 063609 (2013).
[29] M. Lüscher, Commun. Math. Phys. **105**, 153 (1986).

[30] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, Journal of Machine Learning Research **18**, 1 (2018).

[31] M. F. Kasim, S. Lehtola, and S. M. Vinko, arXiv e-prints , arXiv:2110.11678 (2021), arXiv:2110.11678 [physics.chem-ph].

[32] H.-J. Liao, J.-G. Liu, L. Wang, and T. Xiang, Physical Review X **9**, 031041 (2019), arXiv:1903.09650 [cond-mat.str-el].

[33] J. M. Arrazola, S. Jahangiri, A. Delgado, J. Ceroni, J. Izaac, A. Szva, U. Azad, R. A. Lang, Z. Niu, O. D. Matteo, R. Moyard, J. Soni, M. Schuld, R. A. Vargas-Hernandez, T. Tamayo-Mendoza, A. Aspuru-Guzik, and N. Killoran, "Differentiable quantum computational chemistry with pennylane," (2021), arXiv:2111.09967 [quant-ph].

[34] T. Tamayo-Mendoza, C. Kreisbeck, R. Lindh, and A. Aspuru-Guzik, ACS Central Science **4**, 559566 (2018).

[35] L. Coopmans, D. Luo, G. Kells, B. K. Clark, and J. Carrasquilla, PRX Quantum **2** (2021), 10.1103/prxquantum.2.020332.

[36] F. Schfer, M. Kloc, C. Bruder, and N. Lrch, Machine Learning: Science and Technology **1**, 035009 (2020).

[37] I. Khait, J. Carrasquilla, and D. Segal, "Optimal control of quantum thermal machines using machine learning," (2021), arXiv:2108.12441 [quant-ph].

[38] N. Barnea, L. Contessi, D. Gazit, F. Pederiva, and U. van Kolck, Phys. Rev. Lett. **114**, 052501 (2015), arXiv:1311.4966 [nucl-th].

[39] S. Amarasinghe, R. Baghdadi, Z. Davoudi, W. Detmold, M. Illa, A. Parreño, A. V. Pochinsky, P. E. Shanahan, and M. L. Wagman, (2021), arXiv:2108.10835 [hep-lat].

[40] J. Carlson, S. Gandolfi, F. Pederiva, S. C. Pieper, R. Schiavilla, K. E. Schmidt, and R. B. Wiringa, Rev. Mod. Phys. **87**, 1067 (2015), arXiv:1412.3081 [nucl-th].

[41] S. Gandolfi, J. Carlson, A. Roggero, J. Lynn, and S. Reddy, Phys. Lett. B **785**, 232 (2018), arXiv:1712.10236 [nucl-th].

[42] E. Epelbaum, H. Krebs, D. Lee, and U.-G. Meißner, Eur. Phys. J. A **45**, 335 (2010), arXiv:1003.5697 [nucl-th].

[43] T. A. Lähde and U.-G. Meißner, *Nuclear Lattice Effective Field Theory: An introduction*, Vol. 957 (Springer, 2019).

[44] A. Reuther, J. Kepner, C. Byun, S. Samsi, W. Arcand, D. Bestor, B. Bergeron, V. Gadepally, M. Houle, M. Hubbell, M. Jones, A. Klein, L. Milechin, J. Mullen, A. Prout, A. Rosa, C. Yee, and P. Michaleas, in *2018 IEEE High Performance extreme Computing Conference (HPEC)* (IEEE, 2018) pp. 1–6.