



Cite this: *Phys. Chem. Chem. Phys.*,
2022, 24, 25191

From vibrational spectroscopy and quantum tunnelling to periodic band structures – a self-supervised, all-purpose neural network approach to general quantum problems†

Jakob Gamper,^{‡a} Florian Kluibenschedl,^{‡a} Alexander K. H. Weiss^b and
Thomas S. Hofer^{‡a*}

In this work, a feed-forward artificial neural network (FF-ANN) design capable of locating eigensolutions to Schrödinger's equation via self-supervised learning is outlined. Based on the input potential determining the nature of the quantum problem, the presented FF-ANN strategy identifies valid solutions solely by minimizing Schrödinger's equation encoded in a suitably designed global loss function. In addition to benchmark calculations of prototype systems with known analytical solutions, the outlined methodology was also applied to experimentally accessible quantum systems, such as the vibrational states of molecular hydrogen H₂ and its isotopologues HD and D₂ as well as the torsional tunnel splitting in the phenol molecule. It is shown that in conjunction with the use of SIREN activation functions a high accuracy in the energy eigenvalues and wavefunctions is achieved without the requirement to adjust the implementation to the vastly different range of input potentials, thereby even considering problems under periodic boundary conditions.

Received 24th August 2022,
Accepted 3rd October 2022

DOI: 10.1039/d2cp03921d

rsc.li/pccp

1 Introduction

Ever since the emergence of quantum mechanics in the early 1900s, a large number of approaches providing analytical or at the very least numerical solutions to Schrödinger's equation¹ have been proposed. The Schrödinger picture enables *inter alia* a holistic representation of electrons in the non-relativistic limit. For this reason, solutions to the Schrödinger equation are of critical importance in many fields of physical and chemical sciences. Knowledge of the associated eigenfunctions provides a unique and unified approach to calculate the properties of virtually any molecular system as well as periodic solid-state structures. However, as a consequence of the inherent complexity associated to the subatomic interactions, no analytical solutions beyond those reported for one-electron systems are known to date. Recently, it has even been argued that no such solutions may be formulated for comparably simple systems such as the He atom.² While routine applications in quantum chemistry rely on numerical frameworks such as

Hartree–Fock (HF)^{3–6} and post-HF^{6,7} approaches as well as various implementations of density functional theory (DFT),^{8–10} the pursuit of alternative routes to provide more efficient as well as more accurate solutions to electronic structure problems is still a highly active field of research.^{11–16} Typically, simplified one-dimensional prototype systems such as harmonic and anharmonic oscillators, various instances of the particle-in-a-box system and multi-well potentials act as starting points to formulate alternative strategies to solve Schrödinger's equation. Following these initial model applications, the methodology can then be extended to more general frameworks involving an increased number of quantum particles, higher dimensionality and even antisymmetric constraints enabling the treatment of fermionic systems. However, despite the dimensional simplification, there is still a plethora of systems in quantum chemistry which are essentially 1D in nature. Typical examples are the stretch vibrations of diatomic molecules like H₂, HD and D₂,¹⁷ which due to their low effective mass represent highly sensitive quantum systems. Similarly, isolated stretch vibrations of carbon–halogen and oxygen–hydrogen bonds in organic compounds can be considered as 1D oscillators.^{18,19} Furthermore, a broad range of systems can be described by 1D double- and multi-well potentials. Prominent examples are the rotation of the hydroxy group in phenol^{20,21} and aliphatic alcohols like ethanol,^{22,23} described by periodic double- and triple-well potentials, as well as the transposition of atoms around the ideal bond axis in amorphous solids such as glasses.^{24,25}

^a Theoretical Chemistry, Division, Institute of General, Inorganic and Theoretical Chemistry, Center for Chemistry and Biomedicine, University of Innsbruck, Innrain 80-82, A-6020 Innsbruck, Austria

^b Research Institute for Biomedical Aging Research, University of Innsbruck, Rennweg 10, A-6020 Innsbruck, Austria

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d2cp03921d>

‡ J. Gamper and F. Kluibenschedl contributed equally to this work.



Recently, instances of heavy atom (*i.e.* carbon, nitrogen, oxygen) tunneling have been reported, greatly extending the importance of quantum mechanical tunnel effects for the understanding of chemical reactions.^{26,27} When turning to 1D periodic systems a frequent example encountered is the rectangular Kronig–Penney potential,^{28,29} which is a simplified model system corresponding to the periodic generalization of a prototype quantum particle bound to a finite square-well potential.³⁰ Despite its simplicity this system already gives rise to an analytic dispersion relation, similar to band structures known from solid-state physics. Additionally, many 1D systems are analytically solvable, which makes them ideal candidates to assess the performance and applicability of newly formulated numerical techniques to solve the Schrödinger equation. Established numerical methods to solve the 1D Schrödinger equation for a given potential energy surface (PES) of general form, are the grid-based Numerov method,^{17,20,31,32} discrete variable representation (DVR) techniques,^{33,34} the Chebychev collocation approach^{35,36} as well as other finite-difference methods.³⁷ However, recent developments in the theory and implementation of neural networks (NNs) paved the way for a new research area, focused on solving partial differential equations (PDEs) with NNs.^{38,39} Among others, neural networks have been successfully constructed to solve PDEs describing cosmological phase transitions⁴⁰ and processes in fluid dynamics,^{41,42} both of which are known to be highly challenging problems in physical sciences. The innovation in these approaches is to regard the NN as a universal function approximator^{43–45} acting as test function for the solution of a given PDE. Network parameters are obtained in a self-supervised learning process aimed at solving the PDE within the constraints of preset boundary conditions.

In this work, a feed-forward artificial neural network (FF-ANN)⁴⁶ with 1 input, 1 output and a single hidden layer with $2n$ units was successfully implemented to solve the 1D Schrödinger equation given as

$$\begin{aligned}\hat{H}\psi_m(x) &= \left(-\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial x^2} + \hat{V}\right)\psi_m(x) \\ &= E_m\psi_m(x),\end{aligned}\quad (1)$$

where E_m and $\psi_m(x)$ denote the eigenenergy and corresponding eigenstate, given in position space representation and labelled by the quantum number $m \in \{0, 1, \dots, k\}$. The respective Hamiltonian \hat{H} consists of the potential operator \hat{V} and the operator of the kinetic energy, with \hbar being the reduced Planck constant and μ the (effective) mass of the quantum system. The output of the proposed FF-ANN implementation $\psi_m(x)$ is a continuous function over the input domain, enabling inter- and extrapolation from any set of grid points employed in the learning process. This feature represents a distinct advantage of the introduced FF-ANN approach and distinguishes the method from the above mentioned grid-based frameworks, providing only output information at the employed grid points. Notable work in this direction already exists in ref. 47 and 48, however so far the discussion is restricted to relatively simple analytic systems, without any application to complex quantum mechanical systems

characterized *via* experimental measurements. Furthermore, the presented approach differs in a significantly improved version of the network structure, which is considered to be more flexible and even permits the modelling of complex wave functions, a feature not discussed in existing NN-based approaches to solve Schrödinger's equation. In addition, the approach is applied to non-analytic PES, the solutions of which can be directly related to experimental data.

2 Methods

Network architecture and loss function

The target of the present approach is to locate the lowest k eigenenergies E_m and their associated eigenstates $\psi_m(x)$ of the spatial one-dimensional stationary Schrödinger equation as given in eqn (1). A feed-forward artificial neural network (FF-ANN) with 1 input, 1 output and a single hidden layer with $2n$ units (*i.e.* neurons) and special internal structure, as schematically depicted in Fig. 1, is considered as test function for every eigenstate $\psi_m(x)$. In order to also enable the modeling of complex-numbered eigenfunctions, required for instance in the treatment of periodic systems based on Bloch's theorem,^{30,49} the hidden layer is grouped into sub-layers \mathcal{R} and \mathcal{I} , comprised of real and imaginary neurons, which enables the FF-ANN to access the complex plane.

The input of the FF-ANN is a set of N tuples (x_i, V_i) , where x_i is the discretized input coordinate ranging from x_{\min} to x_{\max} and $V_i = V(x_i)$ is the respective value for the potential. Notably, this structure enables the use of various sources to generate the input potential, either *via* analytical potential energy functions such as harmonic or Morse oscillators or alternatively, input data obtained for instance *via* a point-wise potential energy scan from quantum chemical calculations.^{17,20} In contrast to the majority of grid-based formalisms such as the Numerov approach and DVR this potential input does not have to be provided on an equispaced grid. The output of the FF-ANN is denoted by $\psi_m(x)$, given as

$$\begin{aligned}\psi_m(x) &= \sum_{l=1}^n \mathcal{R}_l + i \sum_{l=1}^n \mathcal{I}_l \\ &= \sum_{l=1}^n a_l^{\mathcal{R}} g(w_l^{\mathcal{R}} x + b_l^{\mathcal{R}}) \\ &\quad + i \sum_{l=1}^n a_l^{\mathcal{I}} \tilde{g}(w_l^{\mathcal{I}} x + b_l^{\mathcal{I}}),\end{aligned}\quad (2)$$

with $g, \tilde{g}: \mathbb{R} \rightarrow \mathbb{R}$ denoting the respective activation functions. For the sake of generality, the latter can also be chosen differently for individual neurons. Unless otherwise specified, the activation function of choice in this work is the periodic function $g(x) = \tilde{g}(x) = \sin(x)$, also referred to as SIREN (sinusoidal representation networks) activation function presented by Sitzmann *et al.* in 2020.⁵⁰ Another approach realized for the presented FF-ANN is the use of Gaussian activation functions.^{51,52} A detailed comparison between these functions is provided in a separate section.

To obtain approximate solutions of the Schrödinger equation for E_m and ψ_m , the output of the FF-ANN is set equal to the



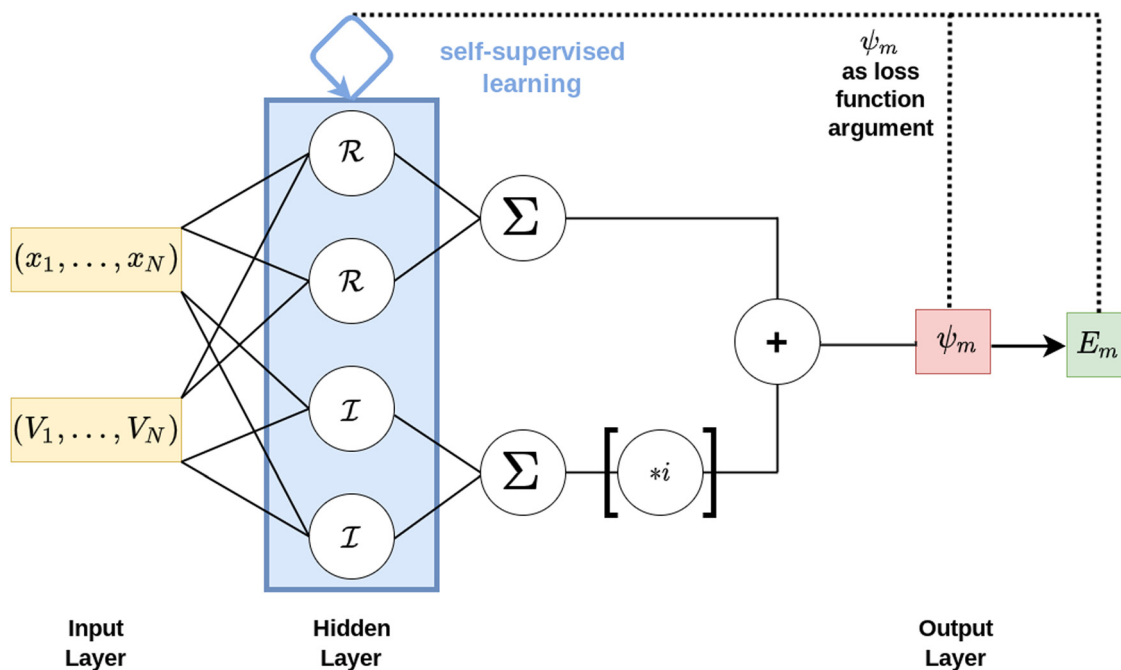


Fig. 1 Schematic illustration of the FF-ANN consisting of 1 input, 1 output and a single hidden layer containing $2n$ individual neurons. The hidden layer is further structured into a real and optional imaginary sub-layer, denoted by \mathcal{R} and \mathcal{I} , each containing a total of n neurons. Multiplication of \mathcal{I} with the imaginary unit i enables the FF-ANN to access the complex plane. The input is a set of N tuples of spatial coordinates x_i with corresponding potential point V_i . In contrast to the majority of grid-based approaches this potential does not have to be provided on an equispaced grid. The FF-ANN output corresponds to an eigenstate $\psi_m(x)$ represented as function over the entire input domain, while simultaneously providing an estimate for the eigenvalue E_m . To determine the optimal network parameters, an appropriately designed loss function is minimized via self-supervised learning.

wave function. From this the corresponding value for E_m is then calculated according to

$$E_m = \frac{\int dx \psi_m^*(x) \hat{H} \psi_m(x)}{\int dx \psi_m^*(x) \psi_m(x)}, \quad (3)$$

reducing the problem to the determination of optimal values for the parameter set $\{a_i, \omega_i, b_i\}^{\mathcal{R}, \mathcal{I}}$. This can be achieved by minimizing the respective loss function \mathcal{L}_m given as

$$\mathcal{L}_m = \gamma_1 E_m + \gamma_2 \phi_m^{\text{diff}} + \gamma_3 \phi_m^b + f(\phi_m^o), \quad (4)$$

with real weighting factors γ_i and a non-linear continuous function $f(\phi_m^o)$ depending on a suitable measure of orthogonality ϕ_m^o . This non-linearity is in contrast to the contributions arising from the boundary conditions, ϕ_m^b , and the differential equation, ϕ_m^{diff} , which are contributing linearly to \mathcal{L}_m . The explicit form for f is included below.

The general notion of the presented loss function was motivated from earlier works focusing on the application of NNs to find solutions to differential eigenvalue problems^{42,53} as for instance cosmological phase transitions presented by Piscopo *et al.* in 2019.⁴⁰ In this work, these approaches are extended to account for the outlined non-linear contribution and associated choice of weighting factors, which is discussed in more detail below. In the following, the individual contributions that appear in eqn (4) are highlighted. The first term, $\gamma_1 E_m$, corresponds to a linear eigenenergy penalty implying that the loss function is

minimized for the lowest sensible eigenvalue. The second term

$$\phi_m^{\text{diff}} = \frac{\int dx \psi_m^*(x) [\hat{H} - E_m] \psi_m(x)}{\int dx \psi_m^*(x) \psi_m(x)} \quad (5)$$

quantifies to what extent the FF-ANN output satisfies the underlying differential equation. In fact, this term should not result in any contribution once the FF-ANN output ψ_m and its respective eigenvalue E_m correspond to a valid eigensolution of the investigated Hamiltonian \hat{H} . As a distinct feature, ϕ_m^{diff} inherently represents a direct measure for the quality of the solution, that can be monitored in real-time as the optimization of the FF-ANN is progressing.

The behavior of the wave function at the boundaries of the spatial region is encoded in ϕ_m^b . All systems considered in this study can be categorized into (a) finite PES with localized solutions ($\varepsilon \rightarrow 0$) corresponding to Dirichlet boundary conditions on the input domain ranging from x_{\min} to x_{\max} , given as

$$\phi_m^b = |\psi_m(x_{\min}) - \varepsilon| + |\psi_m(x_{\max}) - \varepsilon|, \quad (6)$$

and (b) periodic PES with period L and periodic boundary conditions according to

$$\begin{aligned} 2\phi_m^b = & |\psi_m(x_{\min}) - \psi_m(x_{\min} + L)| \\ & + \left| \frac{\partial}{\partial x} \psi_m(x_{\min}) - \frac{\partial}{\partial x} \psi_m(x_{\min} + L) \right| \\ & + \left| \frac{\partial^2}{\partial x^2} \psi_m(x_{\min}) - \frac{\partial^2}{\partial x^2} \psi_m(x_{\min} + L) \right|. \end{aligned} \quad (7)$$



Another key property to be considered is the mutual orthogonality of different eigensolutions to the same Hamiltonian, which is encoded in $f(\phi_m^o)$. For the m -th state ϕ_m^o is written in the form

$$\phi_m^o = \sum_{l=0}^{m-1} \int dx \psi_m^*(x) \psi_l(x), \quad (8)$$

which is the sum over all projections of the currently optimized state ψ_m onto all previously found states ψ_l for $m > l \geq 0$.

Following these definitions, the contribution of the eigenenergy is the only term that effectively contributes to the loss function \mathcal{L}_m once a converged solution has been identified, *i.e.* the contributions ϕ_m^{diff} , ϕ_m^b and $f(\phi_m^o)$ are expected to be zero. In other words, any deviation from the correct eigenstate results in positive penalties to \mathcal{L}_m . Since the goal is to minimize \mathcal{L}_m , the first obtained solution is the one with the smallest eigenenergy corresponding to the quantum mechanical ground state. For all subsequent states the orthogonality condition applies, ensuring that the next minimum of the loss function corresponds to an excited eigenstate. Thus, the spectrum of the Hamiltonian can be found iteratively by minimizing \mathcal{L}_m , always taking into account all previously identified states for the given Hamiltonian. Schematically, this procedure for the training of the FF-ANN is illustrated in Algorithm 1.

Algorithm 1 Training of the FF-ANN

```

LOSS( $m, \psi_m, \psi_{l, l < m}$ ):
   $E_m \leftarrow \text{ENERGY}(\psi_m, \hat{H})$                                 \\\text{eqn (3)}
   $\phi_m^{\text{diff}} \leftarrow \text{DIFF\_EQ}(\psi_m, \hat{H}, E_m)$                     \\\text{eqn (5)}
  if periodic_boundaries = false:
     $\phi_m^b \leftarrow \text{DIRICHLET\_BOUND}(\psi_m)$                       \\\text{eqn (6)}
  else periodic_boundaries = true:
     $\phi_m^b \leftarrow \text{PERIODIC\_BOUND}(\psi_m)$                       \\\text{eqn (7)}
  end if
   $\phi_m^o \leftarrow \text{ORTHOGONAL}(m, \psi_m, \psi_{l, l < m})$             \\\text{eqn (8)}
   $\mathcal{L}_m \leftarrow \gamma_1 E_m + \gamma_2 \phi_m^{\text{diff}} + \gamma_3 \phi_m^b + f(\phi_m^o)$ 
  return  $\mathcal{L}_m$ 
end LOSS

CALC_SPECTRUM( $\hat{H}, \text{number\_of\_states}$ ):
  spectrum  $\leftarrow \{\}$ 
  for  $m \leftarrow 0$  to NUMBER_OF_STATES - 1 do:
     $\psi_m \leftarrow \text{INIT\_NETWORK\_PARAMS}()$ 
     $\psi_m \leftarrow \text{NORMALIZE\_NEURONS}()$ 
     $\psi_m \leftarrow \text{OPT\_LOSS}(\text{LOSS})$ 
     $E_m \leftarrow \text{ENERGY}(\psi_m, \hat{H})$                                 \\\text{eqn (3)}
    spectrum  $\leftarrow \text{spectrum} \cup \{(\psi_m, E_m)\}$ 
  end for
  return spectrum
end CALC_SPECTRUM

```

Activation functions

In the following the different aspects in the design of the FF-ANN approach are outlined in detail. The SIREN activation function⁵⁰ and its second derivative are given as

$$g(x) = \sin(wx + b) \quad (9)$$

$$\frac{d^2 g(x)}{dx^2} = -w^2 \sin(wx + b),$$

while the Gaussian activation function and the corresponding second derivative evolve to

$$g(x) = e^{(-wx^2 - b)} \quad (10)$$

$$\frac{d^2 g(x)}{dx^2} = 2we^{(-wx^2 - b)}(2wx^2 - 1).$$

The latter approach is motivated by the most widely used basis functions in quantum mechanical electronic structure calculation schemes, which are based on minimizing a basis set consisting of pre-parametrized Gaussian type orbitals (GTOs).⁵⁴ When comparing the Gaussian ansatz with the SIREN activation method, which was routinely applied within this work, both methods seem to be a valid option, with the SIREN approach featuring some distinct advantages. It is shown in Table 1, that the SIREN activation function performs slightly better regarding the determination of the eigenvalues for H₂ and its isotopologues HD and D₂. The main reason to choose the SIREN ansatz as state-of-the-art was to obtain an FF-ANN applicable for a diverse set of potentials. For the description of the torsional tunnel splitting in phenol as well as the band structure calculation in case of the Kronig–Penney potential, the solutions of the Schrödinger equation follow the same periodicity as the underlying periodic potential. Hence, the inherently periodic nature of the SIREN activation function provides an adequate approach compared to the Gaussian ansatz.

For the construction of the loss function only analytic derivatives for both activation functions are considered, while all integrations were performed using the Simpsons method.⁵⁵ Here, the SIREN activation function shows one further advantage over its Gaussian counterpart: The integral to calculate the scalar product between two eigenstates can be performed analytically for the SIREN type *via* eqn (S1) and (S2) given in the ESI.† However, this is in general not possible for the

Table 1 Wave numbers of the fundamental and first overtone excitation, $\bar{\nu}_{01}$ and $\bar{\nu}_{02}$, in cm^{−1} for three different isotopologues of the hydrogen molecule: H₂, HD and D₂ at Full-CI/cc-pVQZ level obtained *via* the FF-ANN approach using 40 neurons compared to the associated Numerov results and experimental reference data

	FF-ANN (SIREN)		FF-ANN (Gauss)		Numerov ¹⁷		Exp.	
	$\bar{\nu}_{01}$	$\bar{\nu}_{02}$	$\bar{\nu}_{01}$	$\bar{\nu}_{02}$	$\bar{\nu}_{01}$	$\bar{\nu}_{02}$	$\bar{\nu}_{01}$	$\bar{\nu}_{02}$
H ₂	4162.2	8089.5	4162.4	8090.0	4162.2	8089.5	4161.1 ^a	8087 ^b
HD	3632.7	7088.3	3632.8	7088.4	3632.7	7088.3	3632.3 ^a	7087 ^b
D ₂	2993.7	5868.6	2993.7	5868.6	2993.7	5868.6	2993.6 ^a	5868.5 ^b

^a Ref. 77. ^b Ref. 78.



Gaussian ansatz, which is explicitly shown in eqn (S3) of the ESI.†

Orthogonality contribution

In the following the choice for the orthogonality contribution $f(\phi_m^o)$, with ϕ_m^o as given in eqn (8) is discussed. Empirically, it was found that an expression according to

$$f(\phi_m^o) = 100 \cdot \phi_m^o (|\phi_m^o - 0.5| + 0.5) \quad (11)$$

yields efficient results. In the limit $\phi_m^o \rightarrow 0$ this contribution vanishes, $f \rightarrow 0$, as expected. The motivation for choosing this ϕ_o dependence is the behaviour of f in the limit $\phi_o \rightarrow 1$, which is steeper than the more convenient linear choice, as illustrated in Fig. S1 (ESI†). Hence, the formulation in eqn (11) helps to ensure that the excited states are pushed away from all known lower lying states during training. This choice for f was motivated by the observation that for problems in which the energy difference between neighboring states is larger than 1 caused the network to optimize to the same state as before, at the cost of a penalty in the loss function from the orthogonality contribution. Notably, this result can be still obtained, however it is more unlikely since the steeper slope for $\phi_o \rightarrow 1$ can be interpreted as a repulsion in this minimum of the multi-dimensional parameter space which was found to be highly effective in the initial steps of the training process.

Choice of weighting factors γ_i

The loss function as defined in eqn (4) contains the real weighting factors γ_1 , γ_2 and γ_3 . Ideally, these are set in such a way that no term in eqn (4) outweighs the other contribution by orders of magnitude. Otherwise the optimizer is steered to adjust the parameters in such a way that the FF-ANN output mainly minimizes the contribution associated to the highest weighted contribution, while poorly satisfying all other terms. While it can be argued that given a sufficient number of training periods the optimization algorithm might escape local minima, this is certainly not considered to be good practice. Choosing the weighting factors according to

$$\gamma_1 = |E_0^{-1}| \quad \text{and} \quad \gamma_2 = \gamma_3 = \begin{cases} |E_0^{-1}| & E_0 < 1 \\ |E_0| & E_0 > 1 \end{cases} \quad (12)$$

proved to be effective during training for all systems considered in this study. With this choice the order of magnitude for the first contribution in eqn (4), $\gamma_1 E_m$, is of order 1 for systems where the spacing between different energy levels is of order E_0 . This rescaling is especially important for eigenenergies which are smaller than any sensibly achievable value for ϕ_m^{diff} and ϕ_m^b . The choice for γ_2 and γ_3 ensures that deviations in ϕ_m^{diff} and ϕ_m^b from their ideal values are always strongly penalized, which is especially important in the early stages of the training process. Of course, the choice for the weighting factors as given in eqn (12) is not sensible when training the ground state, which is why $\gamma_{1,2,3}$ are set to 1 in this case.

Initialization and training of the network

The parameters of the FF-ANN employing SIREN activation functions are initialized based on (a) a random uniform distribution or (b) a Fourier-transformation of an approximate harmonic oscillator state. For (a) the initial parameter values $a_i^{\mathcal{R},\mathcal{J}}$, $b_i^{\mathcal{R},\mathcal{J}}$ and $w_i^{\mathcal{R},\mathcal{J}}$ are randomly chosen from the intervals $[0,2]$, $[0,2\pi]$ and $w_0[-\sqrt{c}, \sqrt{c}]$ ($w_0 = 30$, $c = 6$), respectively. Sitzmann *et al.* have shown that this random initialization provides an effective strategy for a large class of problems.⁵⁰ For (b) a harmonic approximation of the potential in each minimum is generated. The eigenstates of the corresponding quantum harmonic oscillator Hamiltonian are well-known and serve as an initial guess for the desired network output. To obtain the parameter values for $a_i^{\mathcal{R},\mathcal{J}}$, $b_i^{\mathcal{R},\mathcal{J}}$ and $w_i^{\mathcal{R},\mathcal{J}}$ a fast Fourier-transformation of the harmonic oscillator states is performed. This approach was motivated from the resemblance of the network output with a truncated Fourier-series in the case of SIREN activation functions. The basin-hopping algorithm^{56–59} was employed for the training of the FF-ANN due to its capability to locate global minima in high-dimensional parameter spaces. The algorithm works by randomly perturbing the network parameters followed by local minimization. Parameters are accepted based on a Metropolis criterion similar to Monte-Carlo algorithms,^{60,61} providing also the possibility of simulated annealing as the optimization is progressing. This procedure is repeated several times and outputs parameter values that minimize the loss function \mathcal{L}_m . In this work local minimization is performed with the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm.^{62–65} For the implementation of the basin-hopping algorithm in conjunction with local BFGS minimization the scipy library⁶⁶ was utilized. Convergence criteria of the local minimizer were set such that the change of the loss function is less than 10^{-12} for 5000 consecutive minimization steps.

3 Results

Analytically solvable systems

To demonstrate the efficiency of the method and benchmark the achievable accuracy, a number of analytically solvable systems has been analyzed. All FF-ANN calculations in this section have been performed excluding the \mathcal{J} layer containing the imaginary functions in eqn (2), since the wave functions are exclusively real in these cases. Network parameters were initialized based on a Fourier-transform of a test eigenstate solving the harmonic approximation of the potential in each minimum. For convenience all calculations in this section were carried out employing atomic units, setting $\hbar = 1$ and $\mu = 1$. This implies that all energies and distances are given in units of hartree and bohr, respectively. First, the analytically solvable single-well potentials of the harmonic and Morse oscillators are considered, which are both widely employed to represent bonded vibrational states in simple molecules.⁶⁷ The Hamiltonian for the harmonic oscillator is defined according to

$$\hat{H}_{\text{harm}} = -\frac{\hbar^2}{2\mu} \frac{\partial^2}{\partial x^2} + \frac{1}{2} \mu \omega^2 x^2, \quad (13)$$



with ω being the angular frequency. Similarly, the Hamiltonian of the Morse oscillator is given as

$$H_{\text{Morse}}(x) = -\frac{\hbar^2}{2\mu} \frac{\partial^2}{\partial x^2} + D_e \left(1 - e^{-\alpha(x-x_0)}\right)^2, \quad (14)$$

with D_e corresponding to the dissociation energy, α being a measure for the width of the potential-well and x_0 denoting the respective equilibrium distance.⁶⁸ Analytic solutions for both systems can be found in the literature.^{30,68} The wave functions for the harmonic and Morse oscillator, obtained using 40 and 65 neurons are depicted in Fig. 2. A quantitative comparison of the respective ground and 3rd excited states with the corresponding analytic wave functions shows that the largest spatial difference on the considered solution domain is of order 10^{-5} to 10^{-2} , for the two systems. Additional examples of associated wave functions together with their analytic counterparts are depicted in the ESI,[†] Fig. S2 and S3 to provide further support for the efficiency of the presented neural network strategy. The corresponding FF-ANN and analytic eigenenergies

are summarized in Table S1 of the ESI.[†] To establish a lower bound for the number of neurons required to obtain sensible results for the harmonic oscillator, the FF-ANN has been successfully trained employing neuron numbers ranging from 1 to 50. The respective differences to the analytic eigenenergies are depicted in Fig. S4 of the ESI.[†] A high accuracy, up to eigenenergy differences below 10^{-4} hartree, can already be obtained with as few as 20 neurons. On the contrary, an increase in the number of neurons does not significantly improve the eigenenergies but may increase the computation time significantly.

In general, the vibrational states of diatomic chemical species show relatively large spacings between the different excitations compared to those associated to molecular rotation. The rotational Hamiltonian H_{rot} in absence of any vibrational coupling about a single axis has a much simpler form compared to the vibrational Hamiltonians by not including any external potential, *i.e.*:

$$\hat{H}_{\text{rot}} = -\frac{\hbar^2}{2I} \frac{\partial^2}{\partial \phi^2} \quad (15)$$

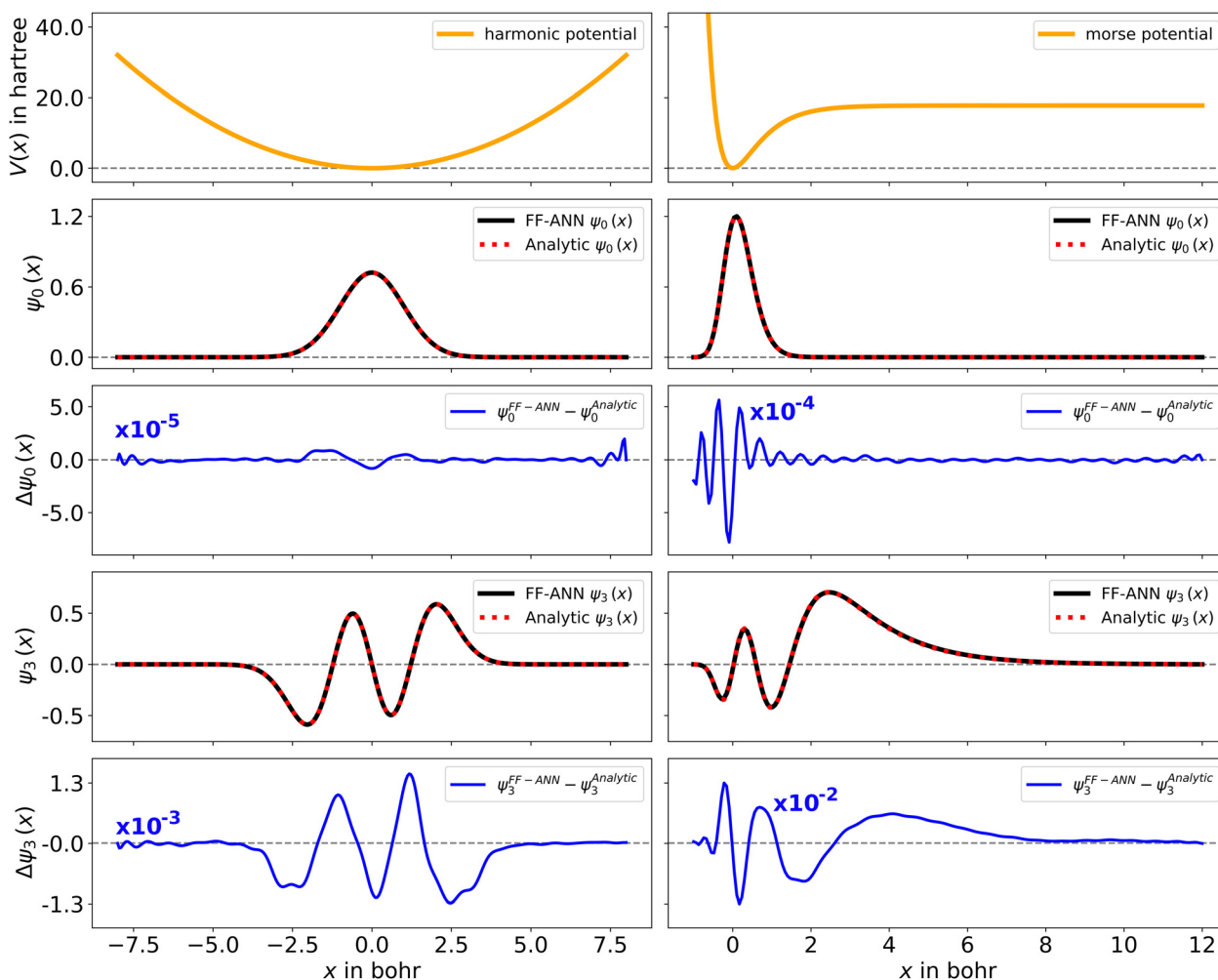


Fig. 2 Quantum mechanical wave functions obtained from the FF-ANN with 40 and 65 neurons for the harmonic oscillator with μ and ω set to 1 (left) and the Morse oscillator with $D_e = 20$, $a = 1.5$ and $x_0 = 0$ (right) defined through the associated potentials (top, orange), employing 40 neurons. The depiction of the output function ψ_m (black) reveals excellent agreement with the analytical wave functions (red dashed), with very small absolute deviations (blue).



with I being the molecular moment of inertia. The Laplacian is given in terms of polar coordinates including only the rotation angle ϕ . To represent the rotation of a chemical species a HCl³⁵ molecule was treated as a linear rigid rotator and its rotational eigensolutions were calculated using the FF-ANN approach employing 40 neurons. The deviation of the lowest three transition wavenumbers to the analytical solutions is below 10^{-12} cm^{-1} and does not rise above 10^{-8} cm^{-1} for the first 10 transitions. All calculated transition wavenumbers are listed in Table S2 and shown in Fig. S5 of the ESI.†

Even more challenging, yet highly relevant benchmark systems for the neural network are applications to double-well potentials. This way the adaptability of the FF-ANN to model wave functions of increasing complexity in terms of curvature in conjunction with small differences in values for the eigenenergy between ground and first excited state, an effect known as tunnel splitting, can be assessed.^{69,70} The first double-well potential under consideration is the Razavy potential,^{71,72} which is parameterized according to

$$V(x) = [\xi \cosh(2x) - M]^2, \quad (16)$$

with parameters ξ and M . Double-well structures were reported to appear for $M > \xi$ along with exact solutions for positive integer values of M .⁷³ Another analytically solvable potential with double-well character is the hyperbolic potential reported by Dong *et al.*,⁷⁴ which is given as

$$V(x) = a^2 \sinh^2(x) - k \tanh^2(x), \quad (17)$$

with parameters a and k . The results for the two considered double-well potentials, obtained using the FF-ANN with 55 and 65 neurons, are depicted in Fig. S6 and S7 of the ESI.† The expected tunnel splitting was successfully reproduced with high accuracy as can be seen in Table S3 (ESI†).

Analytically solvable periodic system

In the following, the FF-ANN was used to solve a challenging periodic quantum system, namely the rectangular Kronig–Penney potential introduced already in 1931.²⁸ It represents the periodic extension of the particle in a finite square-well potential, and despite its simplicity already leads to an analytically solvable dispersion relation associated to band structures known from solid-state physics.²⁹ According to Bloch's theorem the wave function of a system subject to a periodic potential can be described by $\psi_k^m(x) = e^{ikx}u_k^m(x)$, with crystal momentum k , band index m and the lattice periodic function u_k^m .⁴⁹ In order to obtain u_k^m the following equation has to be solved

$$\left[\frac{\hbar^2}{2\mu} (k - i\nabla)^2 + \hat{V} \right] u_k^m = E_k^m u_k^m, \quad (18)$$

which is obtained after insertion of ψ_k^m into the Schrödinger equation. Here, E_k^m denotes the dispersion relation for the m -th band. The associated band structure can thus be obtained by computing the lowest m eigenstates u_k^m and eigenenergies E_k^m of eqn (18) for different values in k , as is illustrated in Algorithm 2.

Algorithm 2 Band structure calculations

```

import CALC_SPECTRUM from Algorithm 1
 $a \leftarrow \text{length\_1D\_unit\_cell}$ 
 $\text{first\_brillouin\_zone} \leftarrow \left[0, \frac{\pi}{a}\right]$ 
ENABLE_IMAGINARY_NEURONS ()
 $\text{disp} \leftarrow \{\}$  \ \ set to store dispersion relation
for all  $k \in \text{first\_brillouin\_zone}$  do:
     $\{(u_k^m, E_k^m)\} \leftarrow \text{CALC\_SPECTRUM}(\hat{H}, n\_bands)$ 
     $\text{disp} \leftarrow \text{disp} \cup \{(k, E_k^m) | 0 \leq m < n\_bands - 1\}$ 
end for
  
```

Since ψ_k^m is unique up to a reciprocal lattice vector it is sufficient to only consider crystal momenta in the first Brillouin zone, *i.e.* $k \in [0, \pi/a]$ with a being the size of the 1D unit cell.^{30,49} To enable a correct treatment of the Hamiltonian shown in eqn (18), the neurons permitting the inclusion of imaginary activation functions of sub-layer \mathcal{J} have to be activated. In the following, a Kronig–Penney potential as depicted in Fig. 3a is considered. Energy values of the first and second band for selected values of the crystal momentum obtained using the FF-ANN with 70 neurons are compared with the corresponding analytical dispersion relation $E(k)^{28,29}$ in Fig. 3b.

Probability densities for selected values of the crystal momentum are shown in Fig. 3c. The successful reproduction of the band structure demonstrates the capability to properly describe periodic PDEs with complex eigensolutions, a feature rarely encountered in the literature.⁵⁰ A special characteristic of the presented FF-ANN is the possibility to determine the periodic function $u_k^m(x)$ at any given k -point, an analysis of which can be related to the finding that the tunnelling probability decreases with an increase in crystal momentum. To the best of knowledge the real-space wave functions $u_k^m(x)$ for general values of k have not yet been reported in the literature, since typically the solutions only provide the dispersion relation E_k^m .

Bond vibrations of H₂, HD and D₂

In addition to the analytical benchmark calculations, experimentally measured quantum mechanical systems were investigated to provide examples for the applicability of the presented approach beyond analytically solvable potentials. In the first example, the binding potential of molecular H₂ was determined *via* accurate quantum chemical calculations and the associated vibrational eigenstates were evaluated using the FF-ANN approach. The bond length between the two hydrogen atoms was varied equidistantly with a step size of 0.025 Å starting from 0.4 Å up to 7.0 Å. For each generated configuration a single point configuration interaction (CI)⁶ calculation at the cc-pVQZ⁷⁵ level using single and double excitations was performed (CISD) *via* gaussian16.⁷⁶ Since the hydrogen molecule possesses only two electrons, the CISD approach can be considered as equivalent to Full-CI.⁶ According to the Born–Oppenheimer approximation,⁶ the Schrödinger equation can



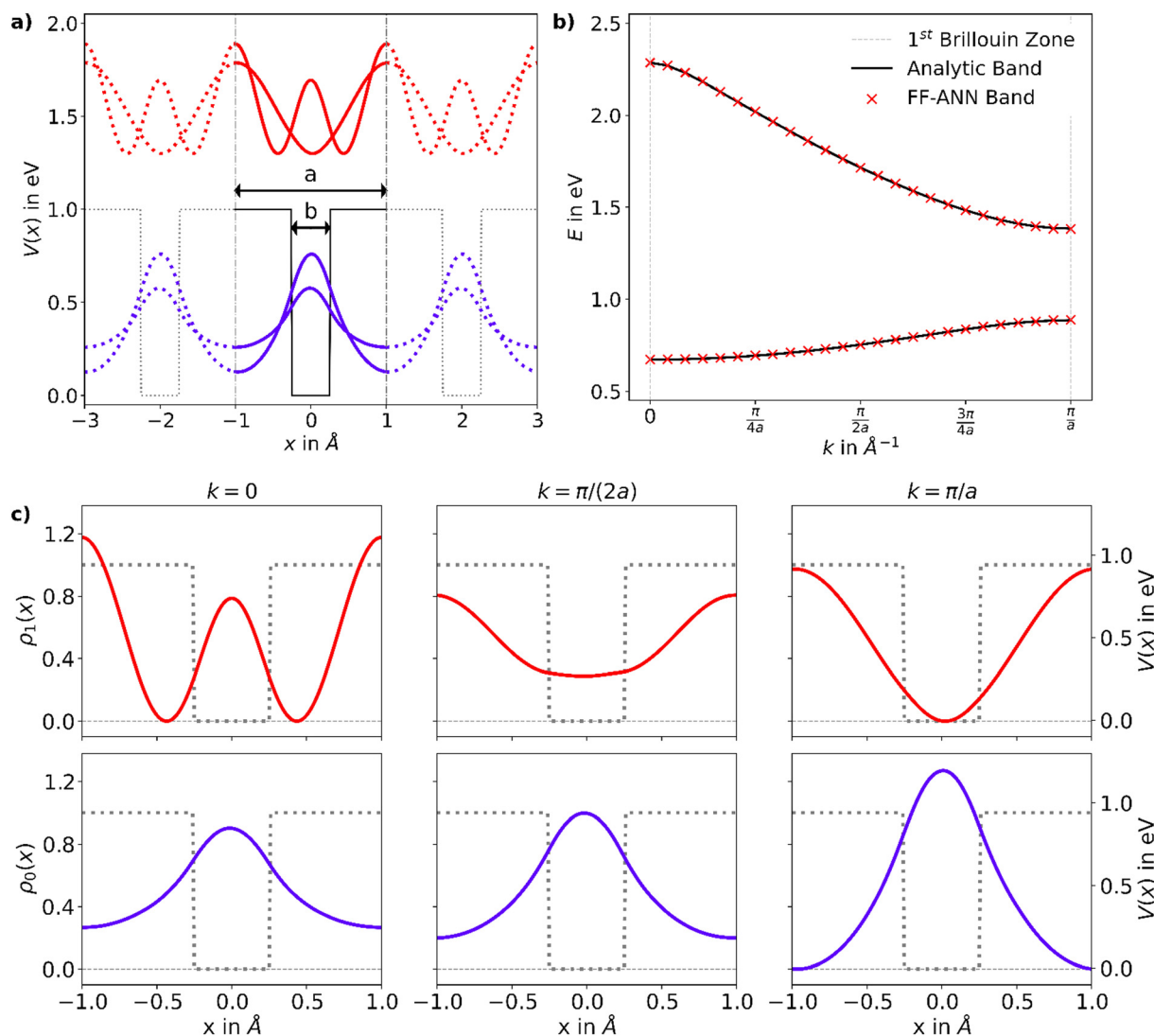


Fig. 3 (a) Rectangular Kronig–Penney potential with $a = 2$ Å, $b = 0.5$ Å and $V = 1$ eV, along with the probability densities $\rho_i = u_k^{*i} u_k^i$ of the first ($i = 0$, blue) and second ($i = 1$, red) band, depicted for two different values of the crystal momentum $k = 0$ and π/a , corresponding to k -points in the center and at the border of the 1st Brillouin zone. (b) Comparison of the associated dispersion relation $E(k)$ obtained from the FF-ANN against the analytical solution.²⁹ (c) Probability densities for the first (blue) and second (red) band for k -values of 0 , $\frac{\pi}{2a}$ and $\frac{\pi}{a}$. The probability densities and band energies were obtained using 70 neurons.

be applied on the same PES for H₂, HD and D₂ respectively, thereby only varying the effective mass μ within the vibrational Hamilton operator to obtain the associated eigenvalues and eigenstates. Therefore, three calculations were performed using the presented FF-ANN network on the described grid-based potential using 40 neurons. The effective masses were set to $\mu_{\text{H}_2} = 0.5039125$ g mol⁻¹, $\mu_{\text{HD}} = 0.6717112$ g mol⁻¹ and $\mu_{\text{D}_2} = 1.007051$ g mol⁻¹.

For the FF-ANN routine two different activation functions were employed. In addition to the SIREN activation functions,⁵⁰ their Gaussian-type analogues as described in Section 2 were used to perform the calculation for the respective isotopologues.

The resulting fundamental and first overtones are summarized in Table 1, while Fig. S8 of the ESI,[†] depicts the respective

potential along with the calculated ground state and the seven lowest excited state wavefunctions for the three isotopologues H₂, HD and D₂. In addition, a comparison between the three lowest eigenstates of all three considered systems is shown in Fig. S8 (ESI[†]).

The data obtained *via* the FF-ANN approach for H₂ are in excellent agreement with the experimental reference values, with the errors for the fundamental and first overtone being 1.1 cm⁻¹ and 2.5 cm⁻¹ respectively, corresponding to relative deviations of 0.026% and 0.031%. These deviations are greatly reduced in case of HD and D₂, showing a difference of only 0.1 cm⁻¹ and 0.4 cm⁻¹ (*i.e.* 0.01% and 0.003%) for the fundamental excitation. This systematic improvement can be explained by the decreasing impact of rotational–vibrational coupling upon increasing mass of the oscillator. Since the 1D

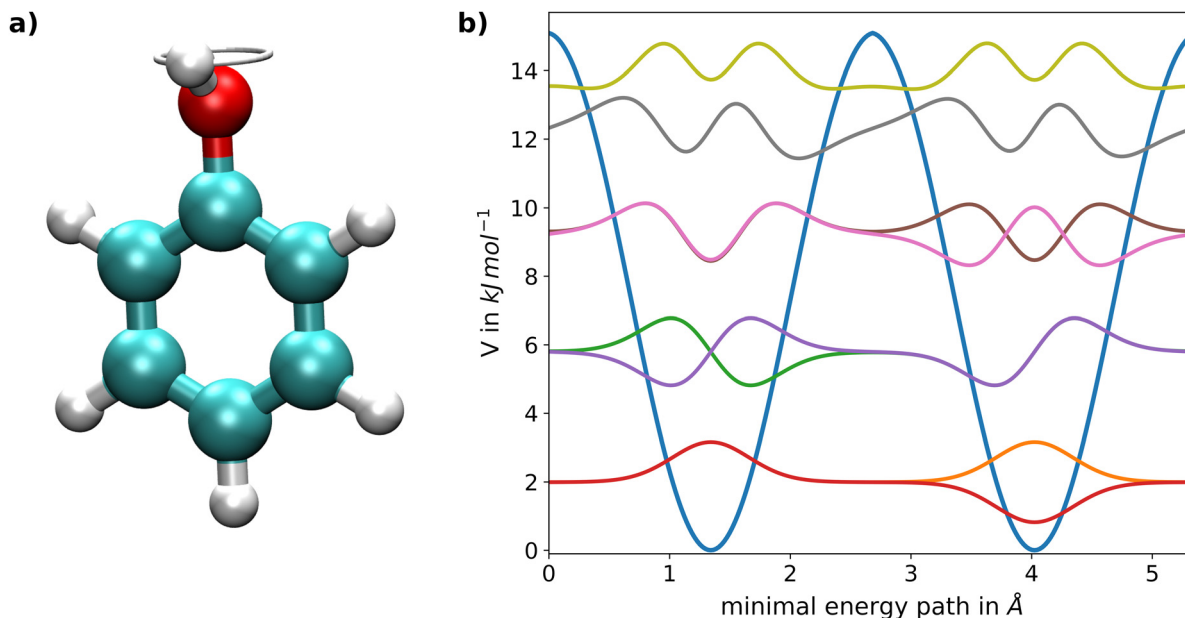


Fig. 4 (a) Graphical representation of the intramolecular torsion of the OH-group in phenol. The snapshot depicts the configuration of the transition state approximately at 2.7 Å; of the minimal energy path. (b) Potential energy surface along the respective periodic minimum energy path including the ground state as well as the seven lowest excited states obtained *via* the FF-ANN approach using a total of 20 neurons.

vibrational Schrödinger equation does not take these contributions into account, the largest deviations are observed in the H_2 case, which is the lightest isotopologue considered. Furthermore, the neural network calculations including the SIREN activation functions match exactly the results obtained *via* the Numerov approach,¹⁷ which was applied to the same input potential.

The comparison of the data obtained by treating the test sets with the same FF-ANN settings, differing only in the choice of activation function, shows some minor advantages in favor of the SIREN approach, when referring to the Numerov and experimental data. Hence, the latter were chosen as default, which is further supported by the inherent periodicity of the sine function. This leads to further major advantages applying the FF-ANN on periodic potentials as already seen in case of the Kronig–Penney potential discussed above.

Torsional tunnel splitting in phenol

As a second application a complex target system featuring torsional tunnel splittings was investigated, namely the intramolecular rotation of the OH-group in the phenol molecule depicted in Fig. 4a. This degree of freedom represents a prime example of a quantum mechanical tunnelling phenomenon and has been investigated in detail from both the experimental²¹ and theoretical perspective.^{20,21} The associated potential energy along the reaction coordinate is shown in Fig. 4b and was taken from a previous study.²⁰ The minimum energy path was determined *via* constrained energy minimizations at B3LYP level,⁷⁹ thereby fixing the dihedral angle of the OH-group followed by re-calculation of the respective configurations using 2nd order Møller–Plesset perturbation theory (MP2).⁸⁰ For the determination of the eigenvalues and eigenstates, the presented neural

network approach was applied on the latter PES with a total of 20 neurons. The effective mass of $0.98868 \text{ g mol}^{-120}$ is determined as the reduced mass of the four C–H groups in *ortho*- and *meta*-position, since two carbon atoms and one hydrogen atom of the phenyl molecule are located along the torsional axis and should therefore not be considered.

The results for the first four torsional tunnel splittings obtained *via* the FF-ANN approach are listed in Table 2, along with a comparison to data determined *via* a periodic Numerov approach,²⁰ quasiharmonic reaction path Hamiltonian (RPH) calculations²¹ as well as experimental measurements.²¹ Furthermore, the deviations of the three theoretical descriptions from the experimental reference are included.

The calculation results of the FF-ANN show a remarkably good agreement with the experimental data, especially when considering the low absolute wave numbers associated to the first and second tunnel splitting as well as the increase from the lowest (0.0019 cm^{-1}) to the highest experimentally measured wave number (1.77 cm^{-1}), spanning three orders of magnitude. Considering the relative deviation it can be observed that both the Numerov and RPH method tend to underestimate the torsional tunnel splitting. Thus, for the $3^0 \rightarrow 3^1$ splitting energy, for which no experimental data is provided, the result of 17.84 cm^{-1} can be regarded as a very reliable estimation.

4 Conclusion

In this work the capabilities of a feed-forward artificial neural network designed to identify solutions to Schrödinger's equation based on self-supervised learning were demonstrated.



Table 2 Comparison of the torsional tunnel splittings in cm^{-1} for the intramolecular torsion of the OH-group in phenol obtained via the FF-ANN approach at MP2/cc-pVTZ level against the results obtained using an adapted Numerov approach at the same level, a quasi-harmonic reaction path Hamiltonian (RPH) at B3LYP/6-311+G(d,p)//CCSD(T)/cc-pVTZ level and experimental data. The respective relative deviations from the experimental reference values are given in parenthesis

$\bar{\nu}$	FF-ANN MP2	Numerov ²⁰ MP2	RPH ²¹ CCSD(T)	Exp. ²¹
$0^0 \rightarrow 0^1$	0.0019 (+0.0%)	0.0017 (−10.1%)	0.0013 (−31.6%)	0.0019
$1^0 \rightarrow 1^1$	0.0921 (+2.8%)	0.0840 (−6.2%)	0.070 (−21.9%)	0.0896
$2^0 \rightarrow 2^1$	1.79 (+1.1%)	1.69 (−4.5%)	1.49 (−15.8%)	1.77
$3^0 \rightarrow 3^1$	17.84	17.11	15.85	—

It has been shown that the presented strategy shows a large flexibility despite the variety in the considered one-dimensional quantum problems. Different classes of one-dimensional potentials were investigated, ranging from localized to periodic solutions with the extension to determine band structures for the associated potential energy surface. For the latter, half of the hidden layer neurons are set to act as imaginary units, which highlights a unique feature of the presented FF-ANN approach.

Furthermore, it was demonstrated that the applicability of the FF-ANN is not limited to analytic benchmark scenarios. The presented implementation is capable of determining the vibrational excitation energies of the hydrogen molecule H_2 and its isotopologues HD and D_2 in very good agreement with experimental reference values. Additionally, the prominent torsional tunnel splitting of the intramolecular torsion of the OH-group in phenol could be resolved with exceptional accuracy without relying on any experimental input.

A remarkable feature regarding the flexibility of the FF-ANN is strongly related to the presented architecture. The neural network is able to perform calculations on any number of given input tuples (x_i, V_i) , without constraining the resulting eigenstates of the solution to a pre-specified output grid. Furthermore, the approach presented within this work is not limited to an equidistant potential input, which is a key advantage over the majority of grid-based approaches such as the Numerov framework and related methods. Thus, for the input of the PES only selected supporting points have to be provided, preferably in regions where the curvature in the eigenstate can be expected to be low. For example, a possible strategy is to provide more points near the respective minimum, while fewer are given at larger displacements.

In general, grid-point approaches only provide output data at the location of the associated input points. Hence, the eigenstate information in between these grid points can only be accessed via approximate interpolation schemes. In contrast, the output of the presented FF-ANN is represented as a function of the neurons in the hidden layer. Therefore, no interpolations have to be performed to access the output over the entire domain providing also the possibility of extrapolation outside the provided input without any further computational cost. The presented one-dimensional implementation displays a larger computational demand compared to grid-based methods

such as DVR or Numerov approach. However, when aiming at a generalization of the framework to problems in higher dimensions, the computational demand as well as memory requirements of the grid-based problems increase by orders of magnitude even when employing a sparse matrix representation,^{17,81} since in this case the problem needs to be represented in terms of block matrices. This gives the FF-ANN the potential to provide a more efficient scalability compared to these established methodologies, in particular when combined with the advantages of providing a non-equispaced input potential along with a wavefunction output over the entire domain as discussed above.

Due to the FF-ANN's high flexibility and wide applicability, the extension to n -dimensional problems and the subsequent application to fermionic states by including an antisymmetric contribution in the loss function is considered as a promising outlook for further development of the presented FF-ANN. Comparing the methodology of the neural network to standard quantum mechanical calculation schemes such as Hartree Fock,^{3–6} which are based on pre-defined basis sets, a major distinction can be identified. While the commonly employed Gaussian-type basis functions are typically pre-parametrized for each type of element and contribute only linearly within these schemes by applying a linear combination approach, both the SIREN and Gaussian activation functions of the presented FF-ANN show no inherent parameterization and are optimized regarding all contributing parameters. This property holds great promise to consider the outlined FF-ANN implementation as a primer for the formulation of novel *ab initio* calculation schemes aimed at the treatment of many-electron systems.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Financial supports for this work via a PhD scholarship for J. Gamper issued by the Leopold-Franzens-University of Innsbruck (Vicerector Prof. Dr Ulrike Tanzer) are gratefully acknowledged. The computational results presented have been achieved (in part) using the HPC infrastructure of the University of Innsbruck.

Notes and references

- 1 E. Schrödinger, *Ann. Phys.*, 1926, **386**, 109–139.
- 2 I. Toli and S. Zou, *Chem. Phys. Lett.*, 2019, **737**, 100021.
- 3 P. Lykos and G. W. Pratt, *Rev. Mod. Phys.*, 1963, **35**, 496–501.
- 4 J. R. Stone, *J. Phys. G*, 2005, **31**, R211–R230.
- 5 P. Echenique and J. L. Alonso, *Mol. Phys.*, 2007, **105**, 3057–3098.
- 6 T. Helgaker, P. Jorgensen and J. Olsen, *Molecular electronic-structure theory*, John Wiley & Sons, Nashville, TN, 2013.



- 7 J. Townsend, J. K. Kirkland and K. D. Vogiatzis, *Mathematical Physics in Theoretical Chemistry*, Elsevier, 2019, pp. 63–117.
- 8 P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864–B871.
- 9 W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133–A1138.
- 10 M. Orio, D. A. Pantazis and F. Neese, *Photosynth. Res.*, 2009, **102**, 443–453.
- 11 P. Popelier, *Solving the Schrödinger equation*, Imperial College Press, London, England, 2011.
- 12 J. Hermann, Z. Schätzle and F. Noé, *Nat. Chem.*, 2020, **12**, 891–897.
- 13 D. Pfau, J. S. Spencer, A. G. D. G. Matthews and W. M. C. Foulkes, *Phys. Rev. Res.*, 2020, **2**, 033429.
- 14 K. T. Schütt, M. Gastegger, A. Tkatchenko, K.-R. Müller and R. J. Maurer, *Nat. Commun.*, 2019, **10**, 5024.
- 15 K. Mills, M. Spanner and I. Tamblyn, *Phys. Rev. A*, 2017, **96**, 042113.
- 16 X. Wang, A. Kumar, C. R. Shelton and B. M. Wong, *Phys. Chem. Chem. Phys.*, 2020, **22**, 22889–22899.
- 17 U. Kuenzer, J.-A. Sorarù and T. S. Hofer, *Phys. Chem. Chem. Phys.*, 2016, **18**, 31521–31533.
- 18 J. G. Dojahn, E. C. M. Chen and W. E. Wentworth, *J. Phys. Chem.*, 1996, **100**, 9649–9657.
- 19 G. Cavallo, P. Metrangolo, R. Milani, T. Pilati, A. Priimagi, G. Resnati and G. Terraneo, *Chem. Rev.*, 2016, **116**, 2478–2601.
- 20 U. Kuenzer and T. S. Hofer, *Chem. Phys. Lett.*, 2019, **728**, 195–200.
- 21 S. Albert, P. Lerch, R. Prentner and M. Quack, *Angew. Chem., Int. Ed.*, 2012, **52**, 346–349.
- 22 J. Castillo-Chará and E. L. Sibert, *J. Chem. Phys.*, 2003, **119**, 11671–11681.
- 23 J. R. Durig, W. E. Bucy, C. J. Wurrey and L. A. Carreira, *J. Phys. Chem.*, 1975, **79**, 988–993.
- 24 P. W. Anderson, B. I. Halperin and C. M. Varma, *Philos. Mag.*, 1972, **25**, 1–9.
- 25 A. Alexander and S. Ben, *Condensed Matter Field Theory*, Cambridge University Press, 2nd edn, 2010, vol. 1.
- 26 C. Castro and W. L. Karney, *Angew. Chem., Int. Ed.*, 2020, **59**, 8355–8366.
- 27 S. Karmakar and A. Datta, *J. Chem. Sci.*, 2020, **132**, 127.
- 28 R. D. L. Kronig and W. G. Penney, *Proc. R. Soc. A*, 1931, **130**, 499–513.
- 29 R. L. Pavelich and F. Marsiglio, *Am. J. Phys.*, 2015, **83**, 773–781.
- 30 C. Cohen-Tannoudji, B. Diu and F. Laloe, *Quantum Mechanics*, Wiley-VCH, 1997, vol. 1.
- 31 B. V. Numerov, Trudy Glavnoï rossiiskoi astrofizicheskoi observatorii; t. 2, 1923, vol. 2, pp. 188–288.
- 32 B. V. Numerov, *Mitteilungen der Nikolai-Hauptsternwarte zu Pulkowo*, 1924, vol. 10, pp. 58–155.
- 33 D. T. Colbert and W. H. Miller, *J. Chem. Phys.*, 1992, **96**, 1982–1991.
- 34 A. Bulgac and M. M. Forbes, *Phys. Rev. C*, 2013, **87**, 051301.
- 35 B. Fornberg, *A practical guide to pseudospectral methods*, Cambridge University Press, 1998.
- 36 J. C. Mason and D. C. Handscomb, *Chebyshev polynomials*, Chapman and Hall/CRC, 2002.
- 37 J. C. Strikwerda, *Finite difference schemes and partial differential equations*, SIAM, 2004.
- 38 I. Lagaris, A. Likas and D. Fotiadis, *IEEE trans. neural netw.*, 1998, **9**, 987–1000.
- 39 J. Blechschmidt and O. G. Ernst, *GAMM Mitteilungen*, 2021, **44**, e202100006.
- 40 M. L. Piscopo, M. Spannowsky and P. Waite, *Phys. Rev. D*, 2019, **100**, 016002.
- 41 M. Raissi, P. Perdikaris and G. Karniadakis, *J. Comput. Phys.*, 2019, **378**, 686–707.
- 42 T. Dockhorn, CoRR, 2019, abs/1904.07200.
- 43 K. Hornik, M. Stinchcombe and H. White, *Neural Netw.*, 1989, **2**, 359–366.
- 44 K. Hornik, *Neural Netw.*, 1991, **4**, 251–257.
- 45 B. C. Csáji, et al., *Faculty of Sciences*, MSc thesis, Eötvös Loránd University, Hungary, 2001, vol. 24, p. 7.
- 46 M. Christopher, *Pattern Recognition and Machine Learning*, Springer, New York, NY, 1st edn, 2006.
- 47 M. Sugawara, *Comput. Phys. Commun.*, 2001, **140**, 366–380.
- 48 Y. Shirvany, M. Hayati and R. Moradian, *Commun. Nonlinear Sci. Numer. Simul.*, 2008, **13**, 2132–2145.
- 49 F. Bloch, *Z. Phys.*, 1929, **52**, 555–600.
- 50 V. Sitzmann, J. Martel, A. Bergman, D. Lindell and G. Wetzstein, *Advances in Neural Information Processing Systems*, 2020, pp. 7462–7473.
- 51 E. J. Hartman, J. D. Keeler and J. M. Kowalski, *Neural Comput.*, 1990, **2**, 210–215.
- 52 B. Yuen, M. T. Hoang, X. Dong and T. Lu, *Sci. Rep.*, 2021, **11**, 18757.
- 53 I. Ben-Shaul, L. Bar and N. Sochen, Deep Learning Solution of the Eigenvalue Problem for Differential Operators, 2021, <https://openreview.net/forum?id=m4baHw5LZ7M>.
- 54 P. M. Gill, *Advances in Quantum Chemistry*, Elsevier, 1994, pp. 141–205.
- 55 K. Atkinson, *An introduction to numerical analysis*, John Wiley and Sons (WIE), Brisbane, QLD, Australia, 1989.
- 56 D. J. Wales and J. P. K. Doye, *J. Phys. Chem. A*, 1997, **101**, 5111–5116.
- 57 D. J. Wales and H. A. Scheraga, *Science*, 1999, **5432**, 1368–1372.
- 58 D. J. Wales, *Energy landscapes*, Cambridge University Press, Cambridge, UK, 2003.
- 59 B. Olson, I. Hashmi, K. Molloy and A. Shehu, *Adv. Artif. Intell.*, 2012, **2012**, 1–19.
- 60 *Monte Carlo methods in statistical physics*, Institut Fur Physik Kurt Binder, ed. K. Binder and D. M. Ceperley, Springer, New York, NY, 2nd edn, 1986.
- 61 D. P. Landau and K. Binder, *A guide to Monte Carlo simulations in statistical physics*, Cambridge University Press, Cambridge, England, 4th edn, 2014.
- 62 C. G. Broyden, *IMA J. Appl. Math.*, 1970, **6**, 76–90.
- 63 R. Fletcher, *Comput. J.*, 1970, **13**, 317–322.
- 64 D. F. Shanno, *Math. Comput.*, 1970, 647–656.
- 65 D. Goldfarb, *Math. Comput.*, 1970, 23–26.



- 66 P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt and SciPy 1.0 Contributors, *Nat. Methods*, 2020, **17**, 261–272.
- 67 Z. Rong, H. G. Kjaergaard and M. L. Sage, *Mol. Phys.*, 2003, **101**, 2285–2294.
- 68 J. P. Dahl and M. Springborg, *J. Chem. Phys.*, 1988, **88**, 4535–4547.
- 69 A. Garg, *Am. J. Phys.*, 2000, **68**, 430–437.
- 70 A. Sitnitsky, *Comput. Theor. Chem.*, 2018, **1138**, 15–22.
- 71 M. Baradaran and H. Panahi, *Adv. High Energy Phys.*, 2017, **2017**, 1–13.
- 72 F. Finkel, A. González-López and M. A. Rodríguez, *J. Phys. A: Math. Theor.*, 1999, **32**, 6821–6835.
- 73 M. Razavy, *Am. J. Phys.*, 1980, **48**, 285–288.
- 74 Q. Dong, G.-H. Sun, J. Jing and S.-H. Dong, *Phys. Lett. A*, 2019, **383**, 270–275.
- 75 T. H. Dunning, *J. Chem. Phys.*, 1989, **90**, 1007–1023.
- 76 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian 16 Revision C.01*, Gaussian Inc., Wallingford, CT, 2016.
- 77 G. D. Dickenson, M. L. Niu, E. J. Salumbides, J. Komasa, K. S. E. Eikema, K. Pachucki and W. Ubachs, *Phys. Rev. Lett.*, 2013, **110**, 193601.
- 78 H.-O. Hamaguchi, I. Suzuki and A. Buckingham, *Mol. Phys.*, 1981, **43**, 963–973.
- 79 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 1372–1377.
- 80 C. Møller and M. S. Plesset, *Phys. Rev.*, 1934, **46**, 618–622.
- 81 U. Kuenzer and T. S. Hofer, *Chem. Phys.*, 2019, **520**, 88–99.

