

<https://doi.org/10.1038/s41524-024-01501-5>

# Machine learning Hubbard parameters with equivariant neural networks

**Martin Uhrin** <sup>1,2</sup>✉, **Austin Zadoks** <sup>1</sup>, **Luca Binci** <sup>1,4</sup>, **Nicola Marzari** <sup>1,3</sup> & **Iurii Timrov** <sup>1,3</sup>✉

Density-functional theory with extended Hubbard functionals ( $\text{DFT} + U + V$ ) provides a robust framework to accurately describe complex materials containing transition-metal or rare-earth elements. It does so by mitigating self-interaction errors inherent to semi-local functionals which are particularly pronounced in systems with partially-filled d and f electronic states. However, achieving accuracy in this approach hinges upon the accurate determination of the on-site  $U$  and inter-site  $V$  Hubbard parameters. In practice, these are obtained either by semi-empirical tuning, requiring prior knowledge, or, more correctly, by using predictive but expensive first-principles calculations. Here, we present a machine learning model based on equivariant neural networks which uses atomic occupation matrices as descriptors, directly capturing the electronic structure, local chemical environment, and oxidation states of the system at hand. We target here the prediction of Hubbard parameters computed self-consistently with iterative linear-response calculations, as implemented in density-functional perturbation theory (DFPT), and structural relaxations. Remarkably, when trained on data from 12 materials spanning various crystal structures and compositions, our model achieves mean absolute relative errors of 3% and 5% for Hubbard  $U$  and  $V$  parameters, respectively. By circumventing computationally expensive DFT or DFPT self-consistent protocols, our model significantly expedites the prediction of Hubbard parameters with negligible computational overhead, while approaching the accuracy of DFPT. Moreover, owing to its robust transferability, the model facilitates accelerated materials discovery and design via high-throughput calculations, with relevance for various technological applications.

A fundamental tool in investigating compounds involving transition-metal (TM) and rare-earth (RE) compounds is density-functional theory (DFT)<sup>1,2</sup>, a cornerstone for first-principles simulations in physics, chemistry, and materials science. In practical applications, DFT necessitates approximations to the exchange-correlation functional, with the local spin-density approximation and spin-polarized generalized-gradient approximation ( $\sigma$ -GGA) being the most prevalent choices. However, these approximations yield unsatisfactory outcomes for various properties of TM and RE compounds, primarily due to significant self-interaction errors (SIEs)<sup>3–5</sup> that are particularly pronounced for localized d and f electrons. To address these challenges, more accurate approaches surpassing the limitations of “standard DFT” have been devised. Noteworthy among these are Hubbard-corrected DFT (so-called  $\text{DFT} + U$ <sup>6–8</sup> and its extension

$\text{DFT} + U + V$ <sup>9–11</sup>, whose role in addressing SIEs, rather than correlation errors, was first pointed out in ref. 4), meta-GGA functionals<sup>12–14</sup>, and hybrid functionals<sup>15–17</sup>. While these methods offer valuable insights, each comes with inherent limitations and challenges, as discussed in, e.g., ref. 18. Hubbard-corrected DFT in particular stands out for its greater accuracy with only a marginal increase in computational cost over standard DFT functionals<sup>9</sup>. In the  $\text{DFT} + U + V$  scheme, a corrective Hubbard energy  $E_{U+V}$  is introduced alongside the approximate DFT energy  $E_{\text{DFT}}$ :

$$E_{\text{DFT}+U+V} = E_{\text{DFT}} + E_{U+V}. \quad (1)$$

In the simplified rotationally-invariant formulation<sup>8</sup>, the extended Hubbard correction energy for a manifold with angular momentum  $\ell$

<sup>1</sup>Theory and Simulation of Materials (THEOS), and National Centre for Computational Design and Discovery of Novel Materials (MARVEL), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. <sup>2</sup>Université Grenoble Alpes, St Martin D’Hères, France. <sup>3</sup>Laboratory for Materials Simulations (LMS), Paul Scherrer Institut (PSI), Villigen, Switzerland. <sup>4</sup>Present address: Department of Materials Science & Engineering, University of California Berkeley, Berkeley, CA, USA. e-mail: [martin@digitalbaker.net](mailto:martin@digitalbaker.net); [iurii.timrov@psi.ch](mailto:iurii.timrov@psi.ch)

takes the form<sup>20</sup>:

$$E_{U+V} = \frac{1}{2} \sum_{I\sigma} U^I \text{Tr} [\mathbf{n}_\ell^{II\sigma} (\mathbf{1} - \mathbf{n}_\ell^{II\sigma})] - \frac{1}{2} \sum_I^* \sum_{J(J \neq I)} \sum_\sigma V^{IJ} \text{Tr} [\mathbf{n}_\ell^{IJ\sigma} \mathbf{n}_\ell^{JI\sigma}], \quad (2)$$

where  $I$  and  $J$  are atomic site indices, and  $\sigma$  is the spin index.  $U^I$  and  $V^{IJ}$  are effective on-site and inter-site Hubbard parameters, respectively. The asterisk in the sum signifies that, for each atom  $I$ , the index  $J$  encompasses all its neighbors up to a given distance. The generalized occupation matrices  $\mathbf{n}_\ell^{II\sigma}$  are derived from the projection of the Kohn-Sham (KS) states onto localized atom-centered orbitals  $\phi_m^I(\mathbf{r})$  (Hubbard projector functions) of neighboring atoms:

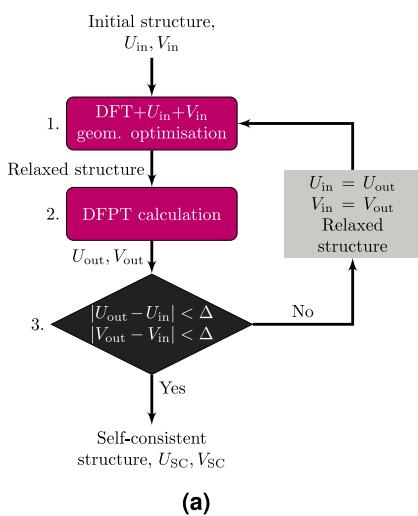
$$\mathbf{n}_\ell^{II\sigma} \equiv n_{mm'}^{II\sigma} = \sum_{v,\mathbf{k}} f_{v,\mathbf{k}}^\sigma \langle \psi_{v,\mathbf{k}}^\sigma | \phi_{m'}^I \rangle \langle \phi_m^I | \psi_{v,\mathbf{k}}^\sigma \rangle, \quad (3)$$

where  $v$  represents the band index of the KS wavefunctions  $\psi_{v,\mathbf{k}}^\sigma(\mathbf{r})$ ,  $m$  and  $m'$  denote magnetic quantum numbers associated with a specific angular momentum  $\ell$ ,  $\mathbf{k}$  denotes points in the first Brillouin zone (BZ), and  $f_{v,\mathbf{k}}^\sigma$  are the occupations of the KS states. The two terms in eq. (2)—proportional to on-site  $U^I$  and inter-site  $V^{IJ}$ —counteract one another. The on-site term promotes localization on atomic sites, suppressing hybridization with neighbors, while the inter-site term favors hybridized states with components on neighboring atoms. Consequently, the values of  $U^I$  and  $V^{IJ}$  are critical to optimizing the extent of the localization and hybridization within Hubbard-corrected DFT. However, these parameters are not known a priori and must be determined in some way. In passing we note that for the sake of simplicity, hereafter we drop the superscripts  $I$  and  $J$  in the notations of Hubbard parameters and occupation matrices, unless required for clarity.

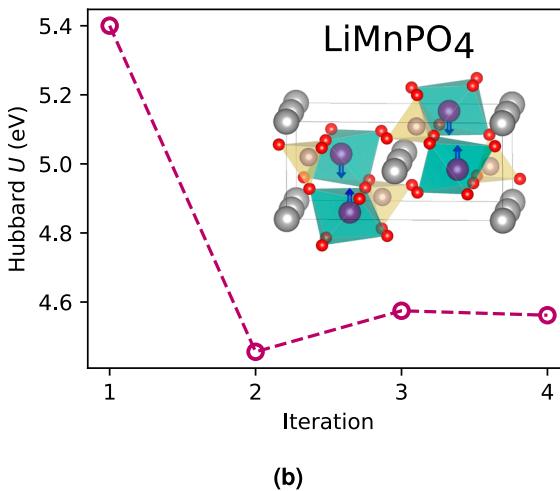
Hubbard  $U$  can be fit semi-empirically to reproduce a target property from experimental data<sup>21–25</sup> or from other advanced first-principles methods (e.g., GW<sup>26</sup> or hybrids<sup>15–17</sup>); however, this approach has many limitations. Fitted parameters are not guaranteed to, and often do not, accurately predict properties other than those used for fitting, and calibrating a single Hubbard  $U$  to multiple properties is non-trivial, requiring advanced algorithms like Bayesian optimization (BO)<sup>27,28</sup>. This method also precludes materials discovery, where properties are by definition unknown, and targeting results of advanced first-principles methods, like DFT with hybrid functionals,

inherits the limitations of those methods<sup>29,30</sup>. Moreover, the inter-site  $V$  parameters, which are necessary to properly describe materials with pronounced covalent interactions, are difficult to fit semi-empirically due to the high-dimensional regression procedure they would require. An attractive alternative is computing Hubbard parameters using first-principles methods such as constrained DFT (cDFT)<sup>31–39</sup>, Hartree-Fock-based approaches<sup>10,11,40–43</sup>, and the constrained random phase approximation<sup>44–47</sup>. These methods do not rely on data from experiments or advanced simulation methods and, moreover, are able to provide values not only for  $U$  but also for  $V$ . However, first-principles approaches are considerably more computationally expensive than DFT +  $U(+V)$  ground-state calculations themselves, making their application feasible but demanding for large systems or high-throughput studies.

The linear-response (LR) formulation of cDFT (LR-cDFT)<sup>20</sup> has witnessed popularity due to its simplicity and accuracy; however, it demands computationally expensive supercell calculations. A recent reformulation of LR-cDFT in terms of density-functional perturbation theory (DFPT)<sup>48,49</sup> significantly reduces the computational burden for determining Hubbard parameters by replacing cumbersome supercell calculations with faster unit-cell calculations and by leveraging symmetries that further diminishes the computational cost by reducing the number of perturbations in reciprocal space (see Sec. S1 in the Supplemental Information (SI)). Its physical rationale relies on heuristically imposing piecewise linearity of the total energy of the system as a function of the population of the Hubbard manifold<sup>20</sup>. Despite the numerous successful applications of DFPT in computing Hubbard parameters<sup>18,50–57</sup>, this approach introduces a significant overhead compared to DFT +  $U+V$  ground-state calculations. Moreover, it has been shown that jointly optimizing Hubbard parameters and the crystal structure, rather than relying on the equilibrium geometry obtained using (semi-)local functionals, can significantly improve the accuracy of the final properties of interest<sup>58</sup>. This is due in part to taking into account the geometry dependence of  $U\{\mathbf{R}\}$  and thus the contribution  $dU/d\mathbf{R}$  to the Hellmann-Feynman forces<sup>59</sup>. To do so, a self-consistent (SC) procedure combining DFPT and structural optimizations (Fig. 1) can be used (see Sec. S2 in the SI)<sup>49,60</sup>. Figure 1 shows a typical SC determination of Hubbard parameters for LiMnPO<sub>4</sub> following this procedure, where each iteration costs approximately one order of magnitude more computational time than a Hubbard-corrected DFT electronic ground-state calculation. Consequently, the adoption of acceleration techniques is highly sought after to expedite the determination of Hubbard parameters while providing a



**Fig. 1 | Calculating Hubbard corrections self-consistently using density-functional perturbation theory.** **a** Protocol for the self-consistent calculation of Hubbard parameters using DFPT<sup>49</sup>.  $U_{\text{in}}$  and  $V_{\text{in}}$  represent the input Hubbard parameters, while  $U_{\text{out}}$  and  $V_{\text{out}}$  denote the output parameters, with  $\Delta$  representing the convergence threshold.  $U_{\text{SC}}$  and  $V_{\text{SC}}$  are the final SC Hubbard parameters.



**b** Convergence of the Hubbard  $U$  parameter for Mn-3d states in LiMnPO<sub>4</sub> using the self-consistent protocol<sup>18</sup>. The inset displays the crystal structure of the material, where arrows indicate the spin direction, and Li atoms are depicted in gray, O in red, Mn in violet, and P in yellow.

level of accuracy very close to that of DFPT and, prospectively, a simple estimate of  $dU/d\mathbf{R}$  and  $dV/d\mathbf{R}$ .

Using machine learning (ML) to accelerate parts of a DFT workflow is becoming increasingly common<sup>61,62</sup>, including applications to predicting Hubbard parameters. A recent example is the work of Cai et al.<sup>63</sup>, in which they use crystal structure parameters, such as bond lengths and angles (among others), as descriptors for a random forest regression model trained on a database of Mn oxides. Their training data consist of  $U$  values fitted using BO to replicate band gaps and band structures obtained from HSE hybrid functional calculations—a concept initially proposed in the earlier work of ref. 27 within the context of DFT +  $U$  and subsequently extended to the DFT +  $U$  +  $V$  framework<sup>64</sup>. The resulting ML model predicts  $U$  values such that subsequent DFT +  $U$  calculations using these values yield band gaps that are 2–3 times smaller than the reference HSE values, while the band structures appear qualitatively similar. Despite the significance of these efforts in applying ML to the prediction of Hubbard parameters, numerous critical issues persist, such as the accuracy of the reference training data (HSE often provides unreliable band gaps for solids<sup>29,30</sup>), the choice of the target properties for the training data (fitting  $U$  to reproduce band gaps is questionable), the importance of self-consistency between Hubbard parameters and the crystal structure, as well as the lack of inter-site  $V$ , to name a few. Addressing these points is highly relevant and necessitates further investigation, which serves as the motivation for this work.

Here, we introduce a novel ML approach based on equivariant neural networks (ENNs) which aims to replace computationally demanding first-principles DFPT calculations of Hubbard  $U$  and  $V$  parameters while providing a negligible loss of accuracy for the vast majority of practical applications. Crucially, the model employs: (1) atomic occupation matrices within the DFT +  $U$  +  $V$  framework as descriptors of the geometry, electronic structure (e.g., oxidation states (OSs)) and local chemical environments in materials, as well as (2) DFPT-based Hubbard parameters, and (3) interatomic distances. This model is general and can be applied to materials with ionic, covalent, and mixed ionic-covalent interactions. It is trained on all intermediate Hubbard parameters and occupation matrices obtained during the SC cycle and directly provides the final SC values of  $U$  and  $V$ , thus also bypassing intermediate structural optimizations. The utilization of ENNs facilitates the exploitation of the inherent O(3) group structure of the occupation matrices, ensuring excellent model performance even with scarce training data. Equivariant models have demonstrated state-of-the-art accuracy and transferability in ML interaction potentials<sup>65,66</sup>, while our work is the first to incorporate electronic-structure degrees of freedom as explicit features in solids.

## Results

### Occupation matrices as the model inputs

The goal of our ML approach is to replace the SC procedure represented in Fig. 1 and provide the final Hubbard parameters using as input results from an initial DFT(+ $U$  +  $V$ ) calculation. Conventional ML methods for atomistic systems primarily use the ionic structures as inputs (see e.g., refs. 67–70), side-stepping the need to explicitly calculate or even consider the electronic structure. However, as a consequence such models may not be particularly sensitive to changes in electronic structure that only lead to subtle changes in local atomic geometry, as can happen during a change of OS. To overcome this limitation, our models take as input the on-site occupation matrices  $\mathbf{n}_\ell^\sigma \equiv \mathbf{n}_\ell^{II\sigma}$  [see Eq. (3),  $I = J$ ], which describe the local electronic structure around an atom, including the OS which is reflected in the occupation matrix eigenvalues<sup>71</sup>. This is a particularly compelling choice given the significant variations in Hubbard parameters seen for different OSs of TM elements<sup>18,52</sup>. We emphasize that relying solely on the trace of the occupation matrix as a descriptor may not adequately distinguish between different OSs and chemical environments.

As with any learning task, the ML model should respect the way in which the target physical properties transform under global rotations, translations, reflections and permutations of labels. In the present case, the output  $U$  and  $V$  values are left unchanged by these transformations, however

the entries of the occupation matrix do change under rotation, making  $n_{mm}^\sigma$  unsuitable for use as inputs to non-symmetry aware models. For this reason, we make use of an *equivariant* learning model<sup>72–75</sup>, which has the property that the learned function  $f: X \rightarrow Y$  obeys the following relation:

$$D_Y[g]f(x) = f(D_X[g]x), \quad \forall g \in G, \quad \forall x \in X, \quad (4)$$

where the equivariance is with respect to a group  $G$  (in our case SE(3)), and  $D_Y[g]$  and  $D_X[g]$  are representations that act on the vector spaces  $Y$  and  $X$ , respectively. For example, applying a rotation to the inputs and then applying  $f$  must produce the same outputs as applying the rotation to the outputs of  $f$  given the original (unrotated) input. While our model outputs scalars (which are symmetry invariant), by using an equivariant model we ensure that the inputs and all intermediate (hidden) features transform together under group actions. This property has been shown to give state-of-the-art accuracy and transferability, particularly within the domain of ML interaction potentials<sup>65,70,76</sup>.

### Equivariant descriptors

By virtue of being atom-centered, the on-site occupation matrix  $\mathbf{n}_\ell^\sigma$  is naturally invariant to global translations. However, to build a rotationally invariant model, it is convenient to re-express the occupation matrix in terms of the irreducible representations (irreps) of the O(3) group. If we let  $I^{(\ell, P)}$  be the irrep with degree  $\ell$  and parity  $P$ , we can express the occupation matrices as the tensor products  $I^{(\ell, P)} \otimes I^{(\ell, P)}$ , where e.g.,  $\ell = 1$  for p orbitals, and  $\ell = 2$  for d orbitals. These can then be decomposed into a direct sum of irreps, e.g., for  $\ell = 1$  and  $P = -1$  the occupation matrix is a  $3 \times 3$ , rank-2, tensor  $I^{(1, -1)} \otimes I^{(1, -1)}$ , that is symmetric, as  $\mathbf{n}_\ell^\sigma = (\mathbf{n}_\ell^\sigma)^T$ . In practice, this decomposition is achieved by applying a change of basis for  $\mathbf{n}_\ell^\sigma$  that transforms it into the irrep basis (see Supplementary Section 3 for details). At this stage, the remaining symmetry to be addressed is that of label permutation. In the case of the atom labels for the inter-site term, the model should be invariant to permutation of  $I$  and  $J$ , however in our training data the Hubbard  $V$  correction is always applied to a d-block/p-block atom pairs which have occupation matrices of different dimension ( $5 \times 5$  for d and  $3 \times 3$  for p), thereby making them distinguishable.

There remains a permutational invariance to be imposed on the spin labels ( $\sigma = \uparrow$  or  $\downarrow$ ). We achieve this using permutationally invariant polynomials:

$$\mathbf{x}_\ell^1 = \mathbf{n}_\ell^\uparrow + \mathbf{n}_\ell^\downarrow, \quad (5)$$

$$\mathbf{x}_\ell^2 = \mathbf{n}_\ell^\uparrow \otimes \mathbf{n}_\ell^\downarrow. \quad (6)$$

In the following, for the sake of convenience, we label these tensors in their irrep form as  $\mathbf{x}_p^i$  for p orbitals and  $\mathbf{x}_d^i$  for d orbitals, where  $i$  can be 1 or 2. We note in passing that in the non-spin-polarized case, one can set  $\mathbf{n}_\ell^\uparrow = \mathbf{n}_\ell^\downarrow = \mathbf{n}_\ell/2$ , where  $\mathbf{n}_\ell$  is divided by two to account for spin degeneracy, and then proceed as above. In addition, atomic species (i.e., atomic types) are encoded as one-hot vectors in our ML model. A summary of the irreps of all the model inputs is presented in Table 1.

In the case of the ML model for Hubbard  $V$ , the interatomic distances  $r_{IJ}$  are also included as an additional input, giving some basic information on the ionic structure. It is important to note that we exclusively focus on the on-site occupation matrices  $\mathbf{n}_\ell^{II\sigma}$  for practical reasons. While inter-site matrices (see, e.g.,<sup>77,78</sup>) could be naturally included in our models, Quantum ESPRESSO does not currently output these. Additionally, this approach ensures consistency with the LR approach used to calculate the Hubbard  $V$  parameters from first principles, which relies on the response of on-site occupation matrices only<sup>9,49</sup>. However, using inter-site occupation matrices  $\mathbf{n}_\ell^{II\sigma}$  as descriptors instead of  $r_{IJ}$  presents an interesting alternative that could be explored in future studies. Finally, while additional descriptors, such as the derivatives of the occupation matrix with respect to the Hubbard parameters, could potentially enhance the performance of the ML model, this would significantly increase the cost of generating training data.

## Equivariant neural-network model

To construct our ML model, we use the e3nn library<sup>75,79</sup> and PyTorch<sup>80</sup>; our codebase is open-source and freely available (see “code availability” below). As shown in Fig. 2, we define two separate models, one for predicting the on-site  $U$  values and the other for the inter-site  $V$  ones. By using separate models, we provide flexibility for the model to be used in calculations where only Hubbard  $U$  is applied (i.e., DFT +  $U$ ). In each case, the model starts with one or more nodes that carry attributes, expressed as a direct sum of irreps, which represent the inputs to the learned function. The inputs pass through a series of repeated layers made up of a tensor product (analogous to all-to-all connected layers in a traditional neural network) and a gated non-linearity (an equivariant version of a traditional activation function<sup>79</sup>) followed by a final tensor product before readout. For each experiment, we rescale the model outputs to have zero mean and a standard deviation of one based on the Hubbard parameters found in the training set which simultaneously accelerates training and improves the final loss. All of the tensor products have learnable weights, meaning that every pair of input irreps that contribute to one output irrep has a learnable scalar parameter that is optimized during training. We use the AdamW<sup>81</sup> optimizer to minimize the loss, which is the mean squared error between predicted and training Hubbard parameters.

## Training and validation datasets

To train and validate our ENN model, we curated a dataset comprising materials with diverse crystal structures. This dataset is constructed based on an investigation of various Li-ion cathode materials, covering olivine-type<sup>18</sup>, spinel-type<sup>52</sup>, and layered-type<sup>82</sup> structures at different Li concentrations, employing the SC DFT +  $U$  and DFT +  $U$  +  $V$  protocols outlined above. Additionally, we include materials such as tunnel- and rutile-type  $\text{MnO}_2$ <sup>51,83</sup> and perovskite-type RE nickelates<sup>53</sup> in our dataset. The crystal and electronic structure details, magnetic ordering, and various other properties of these materials are extensively documented in their respective publications. These

are the compounds containing TM elements Ni, Mn, and Fe, and a list of all these materials is provided in Table 2. While our current dataset is relatively small in terms of material count, ongoing efforts are directed towards a high-throughput exploration involving hundreds of materials, which will serve as an expanded training and validation set for a future, general-purpose, model. This initial dataset demonstrates the efficacy of our ML model, showcasing its ability to predict Hubbard parameters even with limited data (see Table 3). Despite the size of the dataset, the model exhibits accurate

**Table 2 | A list of the materials that are used to train and validate the ML model**

Crystal structure type	Chemical composition	#
Olivine	$\text{Li}_x\text{FePO}_4$	5
	$\text{Li}_x\text{MnPO}_4$	5
	$\text{Li}_x\text{Fe}_{0.5}\text{Mn}_{0.5}\text{PO}_4$	5
Spinel	$\text{Li}_x\text{Mn}_2\text{O}_4$	2
	$\text{Li}_x\text{Mn}_{1.5}\text{Ni}_{0.5}\text{O}_4$	2
Layered	$\text{Li}_x\text{NiO}_2$	2
	$\text{Li}_x\text{MnO}_2$	2
Tunnel	Fe-doped $\alpha\text{-MnO}_2$	1
Tunnel	$\alpha\text{-MnO}_2$	1
Rutile	$\beta\text{-MnO}_2$	1
Perovskite	$\text{YNiO}_3$	1
	$\text{PrNiO}_3$	1
Total	12	28

The last column shows the total number of materials. For the olivine-type materials  $x = 0, 0.25, 0.50, 0.75, 0$ , while for the spinel- and layered-type materials  $x = 0, 1$ .

**Table 1 | Irreducible representations of various attributes (inputs and outputs) of the ML model**

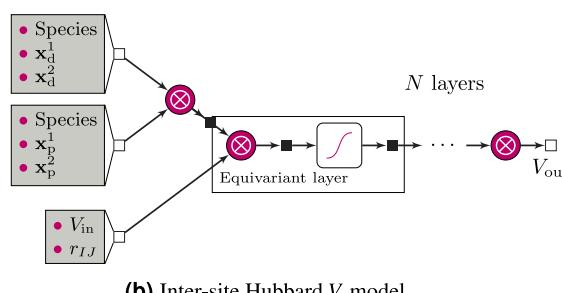
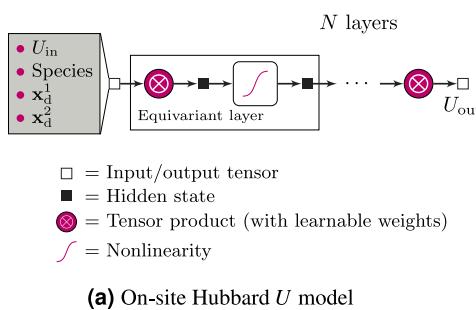
Attribute	Irreducible representation
$U^l, V^J, r_{IJ}$	$\Gamma^{(0, 1)}$
Atomic species	$\bigoplus N_s \Gamma^{(0, 1)}$
$\mathbf{n}_p^\sigma$	$\Gamma^{(0, 1)} \oplus \Gamma^{(2, 1)}$
$\mathbf{n}_d^\sigma$	$\Gamma^{(0, 1)} \oplus \Gamma^{(2, 1)} \oplus \Gamma^{(4, 1)}$

$U^l$  and  $V^J$  are the on-site and inter-site Hubbard parameters, respectively,  $r_{IJ}$  is the interatomic distance between species  $I$  and  $J$ ,  $N_s$  is the number of atomic species (one-hot vectors),  $\mathbf{n}_p^\sigma$  and  $\mathbf{n}_d^\sigma$  are the occupation matrices for the p and d orbitals, respectively, in the special representation for the ML model [see Eqs. (5) and (6)]. The irreducible representations of the occupation matrices can be readily obtained using the e3nn library<sup>75,79</sup>.

**Table 3 | Chemical elements and the total number of data points for both on-site  $U$  and inter-site  $V$  Hubbard parameters determined using the SC protocol illustrated in Fig. 1**

Chemical element	Number of data points	
	$U$	$V$
Ni	396 (124)	67,802 (63,332)
Mn	856 (284)	162,272 (153,232)
Fe	138 (120)	22,511 (21,483)
Total	1390 (528)	252,585 (238,047)

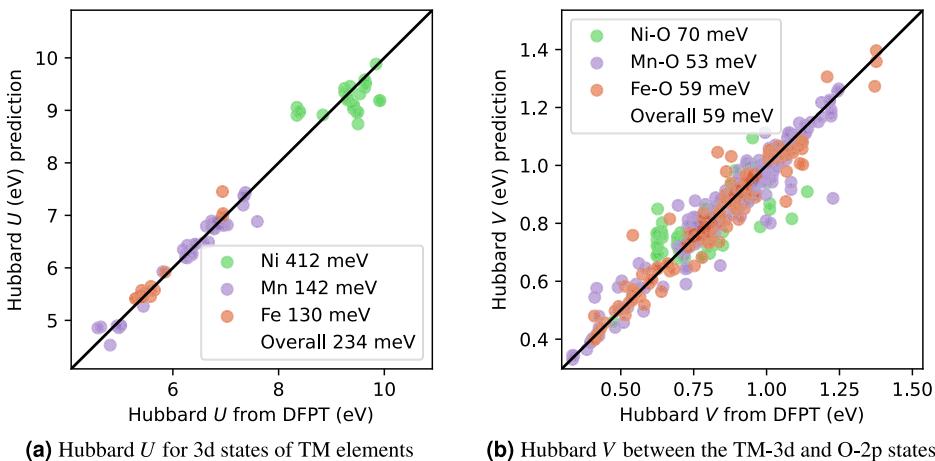
These data points are aggregated across all the materials listed in Table 2 that contain the respective chemical element. Numbers in brackets indicate the unique data points after de-duplication (see text).



**Fig. 2 | Schematic illustration of the equivariant neural-network ML model for predicting Hubbard parameters.** **a** The model for predicting on-site  $U$  values, **b** the model for inter-site  $V$  values.  $U_{\text{in}}$  and  $V_{\text{in}}$  are the input Hubbard parameters, while  $U_{\text{out}}$  and  $V_{\text{out}}$  are the outputs. The atomic species enter as one-hot tensors, and  $r_{IJ}$  is

the interatomic distance between sites  $I$  and  $J$ .  $\mathbf{x}_d^1$  and  $\mathbf{x}_d^2$  are the occupation matrices for the d orbitals, while  $\mathbf{x}_p^1$  and  $\mathbf{x}_p^2$  for the p orbitals, all in the special representation for the ML model [see Eqs. (5) and (6)]. Tensors are represented as squares: open for inputs and outputs, and filled for intermediate features.

**Fig. 3 | Parity plots showing the prediction accuracy on an unseen validation dataset.** **a** Results from Hubbard  $U$ , and, **b**  $V$  predictions. The energies in the legend are the RMSE categorized by element(s) and the overall RMSE across all elements. All attributes listed in Table 1 are used as inputs for the ML model.



predictions, setting the stage for further refinement and expansion with the upcoming comprehensive investigation.

For each material in our dataset, SC Hubbard  $U$  and  $V$  parameters are computed using the SC protocol illustrated in Fig. 1. We include all  $V$  parameters for Hubbard-active atom pairs whose DFPT values are greater than 0.3 eV, in contrast to ref. 64 where only  $V$  for the nearest neighbor couples were included in the training of their ML model. Throughout the iterative process, all intermediate converged occupation matrices and Hubbard parameters are systematically saved. The dataset for training  $V$  comprises approximately seven times more data points than that for  $U$  since  $V$  is an atom-pair quantity, in contrast to the local and consequently sparser nature of  $U$ . For each material and chemical element, this data is randomly split, holding back 20% from each material for validation.

In the first-principles calculations, certain input Hubbard parameters ( $U_{in}$  and  $V_{in}$ ) yield a specific ground state and occupation matrices, and the response of the system due to a perturbation through DFPT provides corresponding output Hubbard parameters ( $U_{out}$  and  $V_{out}$ ). We use these quantities to predict either the final SC Hubbard parameter values, or the result of a single DFPT calculation to test various aspects of the model (see details for each experiment below). It is worth noting that many of the structures contain symmetry-equivalent sites, and correlations between results from successive SC steps are prevalent. To address this issue, we implement a de-duplication procedure aimed at mitigating potential leakage of identical data from the training to the validation set. This involves calculating a symmetry-invariant distance between all pairs of ENN inputs for each atomic species in each material, which is then used to cluster duplicates based on a specified distance threshold (see Supplementary Section 4). In random train/validate splits, we sample from these clusters rather than individual data points, ensuring that only one example from each cluster is included in the validation set.

#### Accuracy of the ML model in predicting the SC Hubbard parameters

First, we demonstrate the model's ability to directly learn the final SC Hubbard parameters for the TM elements across all the materials in our dataset. This is achieved using all attributes listed in Table 1 as inputs for our ML model. Figure 3 shows the parity plots obtained for the on-site  $U$  and inter-site  $V$  Hubbard parameters. The mean absolute relative error (MARE) over all species consistently remains below 3% and 5% for the  $U$  and  $V$  parameters, respectively (per-species distributions are reported in Sec. S5 in the SI). The result for Hubbard  $U$  is particularly promising, considering the relatively small amount of available training data. The higher MARE obtained for  $V$  is likely due to the significantly increased number of degrees of freedom in its model, owing to the involvement of pairs of atoms. It is worth noting that in order to attain this accuracy in predicting the  $U$  and  $V$  values, utilizing a relatively small batch size of 8 – 16 was beneficial for preventing overfitting.

**Table 4 | Comparison of the open-circuit voltages  $\Phi$  (in V) and magnetic moments for TM elements (in  $\mu_B$ ) for  $\text{Li}_x\text{MnPO}_4$  and  $\text{Li}_x\text{FePO}_4$  computed within DFT +  $U + V$ , with  $U$  and  $V$  obtained from first principles using DFPT and predicted using the ML model**

$x$	Property	$\text{Li}_x\text{MnPO}_4$	$\text{Li}_x\text{FePO}_4$
0 – 1	$\Phi$ (DFPT)	4.205	3.544
	$\Phi$ (ML)	4.194	3.544
	$\Delta\Phi$	– 0.26%	0.00%
0	$m$ (DFPT)	3.9738	4.1828
	$m$ (ML)	3.9720	4.1832
	$\Delta m$	– 0.05%	0.01%
1	$m$ (DFPT)	4.7482	3.7388
	$m$ (ML)	4.7486	3.7391
	$\Delta m$	0.01%	0.01%

The voltages are computed using the total energy differences for the Li concentrations  $x = 0$  and  $x = 1$ , while the magnetic moments are computed as the trace of the difference between the spin-up and spin-down occupation matrices<sup>18</sup>.  $\Delta\Phi$  and  $\Delta m$  are the relative differences between the voltages and magnetic moments based on DFPT and ML, respectively.

Second, we examine the extent to which our ML model depends upon the input Hubbard parameters for its prediction accuracy. To investigate this, we repeat the same numerical experiment as above, eliminating the input Hubbard parameters ( $U_{in}$  and  $V_{in}$ ) from our model. As detailed in Sec. S6 in the SI, the parity plots closely resemble those in Fig. 3. The reduction in accuracy of our ML model resulting from this simplification is relatively small, namely the overall root mean square error (RMSE) increases by 29 and 18 meV for the on-site  $U$  and inter-site  $V$ , respectively. This finding suggests that the input Hubbard parameters may not be the most critical input attributes for our ML model, and that the input occupation matrices contain sufficient information to make low-error predictions.

In addition to evaluating the model accuracy, it is also important to evaluate its impact on downstream properties, such as voltages in Li-ion battery cathode materials. Table 4 shows a comparison of calculated open-circuit voltage (OCV)<sup>18</sup>  $\Phi$  and magnetic moments within the DFT +  $U + V$  framework, using first-principles SC Hubbard parameters from DFPT and those predicted by our ML model. Such OCV have been calculated using standard definitions<sup>84,85</sup>, which, when employed in the present context, imply that the energies of different Li concentrations are evaluated using different Hubbard parameters. This approach was used and discussed in more detail in previous works<sup>18,83,86</sup>, while our goal here is to showcase the consistency between the results obtained with first-principles Hubbard parameters and the ones calculated through the ML model. We find that the

differences between the computed and predicted  $U$  and  $V$  values are less than 0.12 eV for  $\text{Li}_x\text{MnPO}_4$  and  $\text{Li}_x\text{FePO}_4$  ( $x = 0$  and  $x = 1$ ). Generally, such a small variation in the values of Hubbard parameters has a negligible impact on the vast majority of various physical and chemical properties of materials. The differences observed in the OCV and magnetic moments in Table 4 are indeed negligible, indicating the accuracy and reliability of the ML-predicted Hubbard parameters.

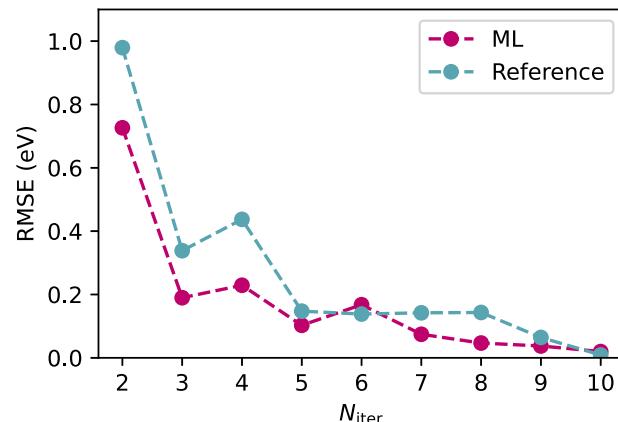
### ML model's performance using a reduced number of iterations in the SC protocol

In the preceding section, we assessed the ML model's performance in predicting final SC Hubbard parameters using training data that consists of a subset of DFPT calculations from all iterations of the SC protocol (see Fig. 1). However, this approach relies on conducting numerous computationally intensive DFPT calculations to generate these training data (often 2–5 but occasionally up to 10 per material to reach self-consistency).

In this experiment, we instead investigate how well the model can predict the results of a *single* DFPT calculation as a function of how many SC iterations are performed in generating training data. This task is more targeted towards high-throughput applications where data generation for model training could be more efficiently performed by sampling the first few DFPT calculations in the SC workflows of many materials rather than the highly-correlated DFPT calculations of full workflows for a few materials. Effectively, this approach aims to stop the SC cycle early and use a model trained on the calculations already performed to refine the not-yet-SC result to be closer to self-consistency, essentially for free. We investigate this use case by training the model on only the first  $N_{\text{iter}} - 1$  DFPT results, and asking it to predict the outcome of the next DFPT calculation and study how the model performs for different values of  $N_{\text{iter}}$ .

Figure 4 shows the average RMSE over all elements as the number of SC iterations used for training  $N_{\text{iter}}$  increases. The reference data correspond to RMSE evaluated using DFPT calculations, i.e., excluding the ML component. This RMSE is computed by determining the difference between the Hubbard parameters at the current  $N_{\text{iter}}$ th and the previous ( $N_{\text{iter}} - 1$ )th iteration. It is evident that the reference RMSE decreases non-monotonically, which was observed in previous studies<sup>49</sup>, and eventually diminishes at large  $N_{\text{iter}}$ . It is noteworthy that the optimal  $N_{\text{iter}}$  for achieving self-consistency in the SC protocol varies for different materials listed in Table 2. Therefore, for larger  $N_{\text{iter}}$  values in Fig. 4, fewer data points are available compared to smaller  $N_{\text{iter}}$  values.

The ML-based RMSE in Fig. 4 at the  $N_{\text{iter}}$ th iteration is computed by training the model on data from all previous  $N_{\text{iter}} - 1$  iterations and predicting the Hubbard parameters for the  $N_{\text{iter}}$ th iteration. Section S7 in the SI contains parity plots for various values of  $N_{\text{iter}}$ , ranging from 2 to 7, alongside the reference results broken down by element as well as the results for the training set. It is evident from Fig. 4 that the ML-based RMSE steadily decreases, and for all  $N_{\text{iter}}$  values except  $N_{\text{iter}} = 6$ , it is smaller than the RMSE of the reference dataset. This result shows that our ML model can improve upon a DFPT-based SC result that was terminated early, particularly during the first few iterations, thereby facilitating faster convergence of Hubbard parameters. For instance, the reference value for  $N_{\text{iter}} = 2$  indicates that on average, a single-shot DFPT calculation yields Hubbard parameters with an RMSE of ~1 eV, while the ML model trained on single-shot data can reduce this error to under 0.8 eV. Incorporating additional  $N_{\text{iter}}$  iterations in the ML training data leads to steady improvements, although the relative performance compared to the reference gradually diminishes until reaching the floor of the model's accuracy. This floor is partly determined by the convergence of the underlying DFPT-based Hubbard parameters in the SC protocol, conducted with a tolerance of approximately  $\Delta = 0.01\text{--}0.1$  eV (see Fig. 1). Therefore, the most significant computational savings in predicting the SC Hubbard parameters can be achieved using our ML model, which acts as a surrogate, after only a few iterations in the DFPT-based SC protocol. This substantially reduces the overall computational cost and renders it a feasible approach, particularly for high-throughput screening scenarios where a priori training data is unavailable for all materials.

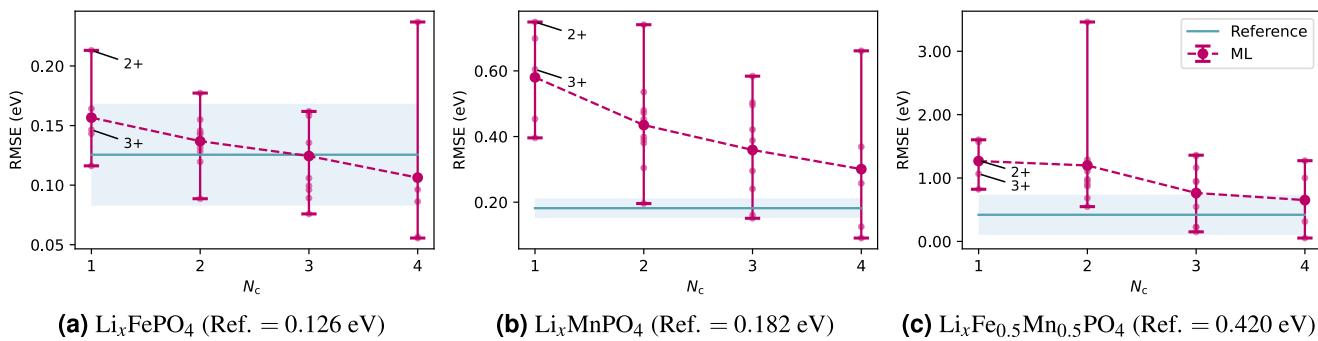


**Fig. 4 | Evaluation of model performance using a reduced number of self-consistent steps.** RMSE for all materials as a function of the number of iterations  $N_{\text{iter}}$  in the SC protocol (see Fig. 1). At each  $N_{\text{iter}}$ , the plotted reference value is the RMSE between the DFPT result obtained from the previous SC iteration and the current one (i.e., a measure of how much the Hubbard  $U$  changes by doing one more iteration). The ML value shows the prediction made from having trained on results from all the preceding  $N_{\text{iter}} - 1$  iterations and predicting the  $N_{\text{iter}}$ th value. This helps to answer the question of how much the ML model can improve the Hubbard parameter, at essentially no cost, when limiting the number of DFPT steps as might be done during the high-throughput screening.

### Transferability of the ML model

To test the transferability of the model, we isolate training data from the olivines ( $\text{Li}_x\text{MPO}_4$  with  $M = \text{Fe}$ ,  $\text{Mn}$ , or  $\text{Fe}_{0.5}\text{Mn}_{0.5}$ ) for which we have calculations at Li concentrations  $x = 0, 0.25, 0.5, 0.75, 1$ . Within this class of materials, the OS of TM elements Mn and Fe changes from +2 to +3 upon delithiation<sup>86</sup>. As mentioned earlier, this change of OS is directly reflected in changes to the occupation matrices<sup>18</sup>. Consequently, the Hubbard parameters for TM elements at different OS also exhibit variation. For instance, the SC Hubbard  $U$  parameter for Mn changes from 4.56 to 6.26 eV when transitioning from +2 to +3, while for Fe this change is from 5.29 to 5.43 eV<sup>18</sup>. In addition, in Sec. S8 of the SI we present the distribution of Hubbard  $U$  parameters for Fe and Mn ions in the olivines extracted from the DFPT-based SC protocol for various Li concentrations. Our ML model effectively captures these changes in the Hubbard parameters and accurately predicts their values based on the occupation matrices for each TM ion at different Li concentrations.

The outcomes of our numerical experiments are shown in Fig. 5. As a reference, we compute the RMSE individually for each material by training the ML model on 80% of the data from all five concentrations and validating it on the remaining 20% of the data after de-duplication. This yields RMSE values of 126, 182, and 420 meV for the Fe, Mn, and mixed Fe-Mn olivines, respectively (see the horizontal lines in Fig. 5). Subsequently, we investigate the RMSE values for scenarios in which the ML model is trained on fewer concentrations, denoted as  $N_c$ , to assess the sensitivity of the ML model to the amount of the training data and its transferability for predicting Hubbard parameters at other concentrations. To accomplish this, we train the ML model on  $N_c$  concentrations and validate it on the remaining  $5 - N_c$  concentrations, and then we average out over all possible permutations of the concentrations between the training and validation datasets. The average RMSE values are represented by larger dots, smaller dots show each individual result, while the error bars in Fig. 5 indicate the maximum and minimum RMSE values resulting from the various permutations. When training on data from a single concentration ( $N_c = 1$ ) we have labeled the results from  $x = 0$  and 1 which correspond to 3+ and 2+ formal OSs respectively, in these cases the model is being tested on OSs it has not seen during training. Four out of six of these give a lower RMSE than experiments where the training data included a mixture of 2/3+ states, suggesting that the model is capable of making meaningful predictions outside of the OS

**Fig. 5 | Evaluation of model transferability to unseen electronic structures.**

a Results for  $\text{Li}_x\text{FePO}_4$ , (b)  $\text{Li}_x\text{MnPO}_4$ , (c)  $\text{Li}_x\text{Fe}_{0.5}\text{Mn}_{0.5}\text{PO}_4$ . RMSE as a function of  $N_c$  (the number of Li concentrations) for three olivines. The model is trained using  $N_c$  concentrations and validated on the remaining  $5 - N_c$ . Error bars indicate the range of RMSE values obtained by considering various permutations of concentrations of the training and validation datasets, large dots represent the average,

small dots show individual results, and experiments that only contain training data from a single oxidation state are labeled (2+ and 3+). The reference is the RMSE computed by training the ML model on 80% of the Hubbard parameters from all five concentrations, with validation performed on the remaining 20% after de-duplication. The blue line represents the mean results over three runs with different random initializations, the confidence intervals show the standard deviation.

**Table 5 | Hubbard  $U$  results when training on  $\alpha$ - $\text{MnO}_2$  data and predicting on  $\beta$ - $\text{MnO}_2$  and vice versa as well as a reference trained on 80% of all of the deduplicated data and predicting on the remainder**

Trained on	RMSE (meV)
80% of $\alpha$ & $\beta$ - $\text{MnO}_2$	144
$\alpha$ - $\text{MnO}_2$	132
$\beta$ - $\text{MnO}_2$	157

In this case, the model shows good transferability to the unseen crystal structures.

seen during training. More generally, as  $N_c$  increases, the average RMSE for each material decreases, reaching values of approximately 106, 301, and 652 meV at  $N_c = 4$  for the Fe, Mn, and mixed Fe-Mn olivines, respectively.

To test whether our model can transfer between different crystal structures we perform an additional experiment where we isolate data from the tunnel ( $\alpha$ ) and rutile ( $\beta$ ) phases of  $\text{MnO}_2$ , training on one and testing predictions on the other. Table 5 shows that the model can transfer well between the two phases, achieving RMSEs comparable to those found when training on 80% of the data from both phases and predicting on the rest.

Overall, these experiments demonstrate that, depending on the desired accuracy, it is not necessary to include training data from the exact target system that the model is being evaluated on, and even using examples of one or two similar structures may be sufficient to extrapolate to other compositions or crystal structures. Given the computational cost of SC calculations, this can lead to a significant speedup, particularly in cases requiring larger supercells and various configurations. These findings demonstrate the good transferability of our ML model, indicating its potential utility and reliability for predicting Hubbard parameters in materials not included in the model's training dataset.

## Discussion

We have introduced a novel equivariant ML model designed for predicting the SC on-site  $U$  and inter-site  $V$  Hubbard parameters, thereby circumventing the computationally intensive DFPT-based protocols. The model incorporates three input descriptors: Hubbard parameters, interatomic distances, and, notably, atomic occupation matrices. The latter plays a pivotal role in encoding essential information about the electronic structure and local chemical environment within materials. Such an ML model holds significant promise, particularly for high-throughput investigations and large-scale systems, scenarios where DFPT-based approaches become too computationally expensive. Furthermore, the model demonstrates good transferability, rendering it reliable for predicting Hubbard parameters in materials not included in its training dataset, or as a very accurate first guess for further DFPT refinements.

The usage of our ML model is straightforward and entails two DFT-based calculations for a given material. Initially, a ground-state calculation employing DFT +  $U + V$  with initial guesses for  $U$  and  $V$  (which can be set to zero) is conducted to determine the atomic occupation matrices required as input for the model. Subsequently, a final structural optimization using the model-predicted SC Hubbard parameters yields a SC structural-electronic ground state. The computational cost of these two calculations and a model evaluation is negligible compared to the DFPT-based SC protocol they replace (see Sec. S10 in the SI). Furthermore, we have demonstrated that the full SC protocol can be reduced to just a few iterations, which prove adequate for achieving a RMSE within a few percent of the real value.

Unlike other ML models designed for predicting Hubbard parameters<sup>27,28,63,64</sup>, our model does not rely on experimental data or information from other state-of-the-art computational methods such as GW and hybrid functionals, which possess inherent limitations and specific ranges of applicability. Instead, our model exclusively relies on LR theory through DFPT, which provides material-specific Hubbard parameters directly reflecting the local chemistry and OS of TM elements. Additionally, the current ML model can be trained on different spin configurations of the same material, with these variations reflected in the occupation matrices, though the resulting changes in Hubbard parameters are minor<sup>51,83</sup>. Notably, our model not only predicts on-site Hubbard  $U$  parameters but also inter-site Hubbard  $V$  parameters, crucial for materials characterized by significant covalent interactions, an aspect that was disregarded in previous studies (Except ref. 64 which is based on the fitting of  $U$  and  $V$  to reproduce the results obtained using hybrid functionals, but it does not provide a ML model that can be used to predict the Hubbard parameters for new materials.). Furthermore, the architecture of our ML model is highly versatile, permitting easy integration of additional inputs and outputs, facilitating exploration of diverse learning tasks beyond learning Hubbard parameters. We expect that similar hybrid ML-accelerated electronic structure methodologies, maintaining accuracy and transferability, will become prevalent, potentially yielding a comparable impact on the field as observed with ML interaction potentials.

Lastly, it is essential to highlight the limitations of the trained models we have presented. These were trained on data generated using a specific computational setup (see below), which must remain exactly the same when the model is applied to other systems. In other words, the  $U$  and  $V$  values predicted by the model are not transferable across electronic-structure codes and even across different pseudopotentials within the same code<sup>87</sup>. Furthermore, the training data only encompass Fe, Mn, and Ni with Hubbard  $U$  corrections, while  $V$  parameters are available for pairs involving these elements and O. Consequently, these models can effectively predict Hubbard parameters for other materials with Hubbard corrections on these atoms or atom pairs, provided that the model input is generated using identical

pseudopotentials, Hubbard projectors, and functional (further elaborated in the “methods” section). Nevertheless, with a more diverse dataset, the models could be easily extended to accommodate a broader range of compositions. Indeed, ongoing efforts aim to establish a comprehensive database of Hubbard parameters for various TM-containing materials, akin to ref. 88, thereby significantly broadening the scope of our model to encompass numerous TM elements across diverse OS and chemical environments. While we have shown that the models can predict the results of fully SC parameters, models which predict the output of single DFPT calculations may be more applicable in certain situations (see Sec. S11 in the SI). We have shown that models can be trained for this task but have not explicitly tested their performance under recursive evaluation, i.e., performing the SC convergence procedure replacing DFPT completely with an ML model. The performance of the models in this use case and the potential for active learning during this process could be investigated in future work. Moreover, our model can integrate into automated AiiDA workflows<sup>89,90</sup>, enabling non-experts to harness it effortlessly and access SC Hubbard parameters with minimal intervention. Consequently, we believe that our ML model represents a significant advancement in expediting materials discovery, design, and understanding based on the DFT + U + V approach, thereby unlocking new avenues for technological progress and breakthroughs.

## Methods

### Ab initio calculations

All calculations are performed using the plane-wave pseudopotential method as implemented in the QUANTUM ESPRESSO distribution<sup>91–93</sup>. We use the exchange-correlation functional constructed using  $\sigma$ -GGA with the PBEsol prescription<sup>94</sup>, and the pseudopotentials are taken from the SSSP library v1.1 (efficiency)<sup>95,96</sup>. For metallic ground states, we use Gaussian smearing. To construct the Hubbard projectors, we use atomic orbitals which are orthonormalized using Löwdin’s method<sup>97,98</sup>. Structural optimizations are performed using DFT + U + V<sup>58</sup> with the Broyden-Fletcher-Goldfarb-Shanno algorithm<sup>99</sup> and convergence thresholds for the total energy of  $10^{-6}$  Ry, for forces of  $10^{-5}$  Ry/Bohr, and for pressure of 0.5 Kbar. The DFPT calculations of Hubbard parameters are performed using the HP code<sup>100</sup>, with an accuracy of 0.01–0.1 eV for the computed values of U and V. The information about the kinetic-energy cutoff,  $\mathbf{k}$  and  $\mathbf{q}$  points sampling of the Brillouin zone for each system are detailed in Sec. S9 in the SI. More technical details can be found directly in the source files publicly available through the Materials Cloud Archive (see below).

### Data availability

The data used to generate the results presented in this paper are accessible in the Materials Cloud Archive<sup>101</sup>.

### Code availability

The latest version of the code for the equivariant neural networks is available at <https://github.com/camml-lab/hubbardml/>. Results presented in this paper were generated using version 0.2.0<sup>102</sup>, which is also available on the Materials Cloud Archive<sup>101</sup>.

Received: 4 June 2024; Accepted: 19 December 2024;

Published online: 25 January 2025

## References

- Hohenberg, P. & Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **136**, B864 (1964).
- Kohn, W. & Sham, L. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133 (1965).
- Perdew, J. & Zunger, A. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B* **23**, 5048 (1981).
- Kulik, H., Cococcioni, M., Scherlis, D. & Marzari, N. Density functional theory in transition-metal chemistry: a self-consistent Hubbard U approach. *Phys. Rev. Lett.* **97**, 103001 (2006).
- Mori-Sánchez, P., Cohen, A. & Yang, W. Many-electron self-interaction error in approximate density functionals. *J. Chem. Phys.* **125**, 201102 (2006).
- Anisimov, V., Zaanen, J. & Andersen, O. Band theory and Mott insulators: Hubbard U instead of Stoner. *I. Phys. Rev. B* **44**, 943 (1991).
- Liechtenstein, A., Anisimov, V. & Zaanen, J. Density-functional theory and strong interactions: orbital ordering in Mott-Hubbard insulators. *Phys. Rev. B* **52**, R5467 (1995).
- Dudarev, S., Botton, G., Savrasov, S., Humphreys, C. & Sutton, A. Electron-energy-loss spectra and the structural stability of nickel oxide: an LSDA+U study. *Phys. Rev. B* **57**, 1505 (1998).
- Campo Jr, V. L. & Cococcioni, M. Extended DFT+U+V method with on-site and inter-site electronic interactions. *J. Phys. Condens. Matter* **22**, 055602 (2010).
- Tancogne-Dejean, N. & Rubio, A. Parameter-free hybridlike functional based on an extended Hubbard model: DFT+U+V. *Phys. Rev. B* **102**, 155117 (2020).
- Lee, S.-H. & Son, Y.-W. First-principles approach with a pseudoohybrid density functional for extended Hubbard interactions. *Phys. Rev. Res.* **2**, 043410 (2020).
- Sun, J., Ruzsinszky, A. & Perdew, J. Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.* **115**, 036402 (2015).
- Bartók, A. & Yates, J. Regularized SCAN functional. *J. Chem. Phys.* **150**, 161101 (2019).
- Furness, J., Kaplan, A., Ning, J., Perdew, J. & Sun, J. Accurate and numerically efficient r<sup>2</sup>SCAN meta-generalized gradient approximation. *J. Phys. Chem. Lett.* **11**, 8208 (2020).
- Adamo, C. & Barone, V. Toward reliable density functional methods without adjustable parameters: the PBE0 model. *J. Chem. Phys.* **110**, 6158 (1999).
- Heyd, J., Scuseria, G. & Ernzerhof, M. Hybrid functionals based on a screened Coulomb potential. *J. Chem. Phys.* **118**, 8207 (2003).
- Heyd, J., Scuseria, G. & Ernzerhof, M. Erratum: “Hybrid functionals based on a screened Coulomb potential” [J. Chem. Phys. 118, 8207 (2003)]. *J. Chem. Phys.* **124**, 219906 (2006).
- Timrov, I., Aquilante, F., Cococcioni, M. & Marzari, N. Accurate electronic properties and intercalation voltages of olivine-type Li-ion cathode materials from extended hubbard functionals. *PRX Energy* **1**, 033003 (2022).
- Himmetoglu, B., Floris, A., de Gironcoli, S. & Cococcioni, M. Hubbard-corrected DFT energy functionals: the LDA+U description of correlated systems. *Int. J. Quant. Chem.* **114**, 14 (2014).
- Cococcioni, M. & de Gironcoli, S. Linear response approach to the calculation of the effective interaction parameters in the LDA+U method. *Phys. Rev. B* **71**, 035105 (2005).
- Wang, L., Maxisch, T. & Ceder, G. Oxidation energies of transition metal oxides within the GGA+U framework. *Phys. Rev. B* **73**, 195107 (2006).
- Kubaschewski, O., Alcock, C. & Spencer, P. *Materials Thermochemistry* 6th edn (Pergamon Press, 1993).
- Le Bacq, O., Pasturel, A. & Bengone, O. Impact on electronic correlations on the structural stability, magnetism, and voltage of LiCoPO<sub>4</sub> battery. *Phys. Rev. B* **69**, 245107 (2004).
- Aykol, M. & Wolverton, C. Local environment dependent GGA+U method for accurate thermochemistry of transition metal compounds. *Phys. Rev. B* **90**, 115105 (2014).
- Isaacs, E. & Marianetti, C. Compositional phase stability of strongly correlated electron materials within DFT+U. *Phys. Rev. B* **95**, 045141 (2017).
- Hedin, L. New method for calculating the one-particle green’s function with application to the electron-gas problem. *Phys. Rev.* **139**, A796 (1965).
- Yu, M., Yang, S., Wu, C. & Marom, N. Machine learning the Hubbard U parameter in DFT+U using Bayesian optimization. *npj Comput. Mater.* **6**, 180 (2020).

28. Tavadze, P. et al. Exploring DFT+*U* parameter space with a Bayesian calibration assisted by Markov chain Monte Carlo sampling. *npj Comput. Mater.* **7**, 182 (2021).
29. Skone, J., Govoni, M. & Galli, G. Self-consistent hybrid functional for condensed systems. *Phys. Rev. B* **89**, 195112 (2014).
30. Skone, J., Govoni, M. & Galli, G. Nonempirical range-separated hybrid functionals for solids and molecules. *Phys. Rev. B* **93**, 235106 (2016).
31. Dederichs, P., Blügel, S., Zeller, R. & Akai, H. Ground states of constrained systems: application to cerium impurities. *Phys. Rev. Lett.* **53**, 2512 (1984).
32. McMahan, A., Martin, R. & Satpathy, S. Calculated effective Hamiltonian for La<sub>2</sub>CuO<sub>4</sub> and solution in the impurity Anderson approximation. *Phys. Rev. B* **38**, 6650 (1988).
33. Gunnarsson, O., Andersen, O., Jepsen, O. & Zaanen, J. Density-functional calculation of the parameters in the Anderson model: Application to Mn in CdTe. *Phys. Rev. B* **39**, 1708 (1989).
34. Hybertsen, M., Schlüter, M. & Christensen, N. Calculation of Coulomb-interaction parameters for La<sub>2</sub>CuO<sub>4</sub> using a constrained-density-functional approach. *Phys. Rev. B* **39**, 9028 (1989).
35. Gunnarsson, O. Calculation of parameters in model Hamiltonians. *Phys. Rev. B* **41**, 514 (1990).
36. Pickett, W., Erwin, S. & Ethridge, E. Reformulation of the LDA+*U* method for a local-orbital basis. *Phys. Rev. B* **58**, 1201 (1998).
37. Solovyev, I. & Imada, M. Screening of Coulomb interactions in transition metals. *Phys. Rev. B* **71**, 045103 (2005).
38. Nakamura, K., Arita, R., Yoshimoto, Y. & Tsuneyuki, S. First-principles calculation of effective onsite Coulomb interactions of 3d transition metals: Constrained local density functional approach with maximally localized Wannier functions. *Phys. Rev. B* **74**, 235113 (2006).
39. Shishkin, M. & Sato, H. Self-consistent parametrization of DFT+*U* framework using linear response approach: Application to evaluation of redox potentials of battery cathodes. *Phys. Rev. B* **93**, 085135 (2016).
40. Mosey, N. & Carter, E. Ab initio evaluation of Coulomb and exchange parameters for DFT+*U* calculations. *Phys. Rev. B* **76**, 155123 (2007).
41. Mosey, N., Liao, P. & Carter, E. Rotationally invariant ab initio evaluation of Coulomb and exchange parameters for DFT+*U* calculations. *J. Chem. Phys.* **129**, 014103 (2008).
42. Andriotis, A., Sheetz, R. & Menon, M. LSDA+*U* method: a calculation of the *U* values at the Hartree-Fock level of approximation. *Phys. Rev. B* **81**, 245103 (2010).
43. Agapito, L., Curtarolo, S. & Buongiorno Nardelli, M. Reformulation of DFT+*U* as a pseudohybrid hubbard density functional for accelerated materials discovery. *Phys. Rev. X* **5**, 011006 (2015).
44. Springer, M. & Aryasetiawan, F. Frequency-dependent screened interaction in Ni within the random-phase approximation. *Phys. Rev. B* **57**, 4364 (1998).
45. Kotani, T. Ab initio random-phase-approximation calculation of the frequency-dependent effective interaction between 3d electrons: Ni, Fe, and MnO. *J. Phys. Condens. Matter* **12**, 2413 (2000).
46. Aryasetiawan, F. et al. Frequency-dependent local interactions and low-energy effective models from electronic structure calculations. *Phys. Rev. B* **70**, 195104 (2004).
47. Aryasetiawan, F., Karlsson, K., Jepsen, O. & Scönberger, U. Calculations of Hubbard *U* from first-principles. *Phys. Rev. B* **74**, 125106 (2006).
48. Timrov, I., Marzari, N. & Cococcioni, M. Hubbard parameters from density-functional perturbation theory. *Phys. Rev. B* **98**, 085127 (2018).
49. Timrov, I., Marzari, N. & Cococcioni, M. Self-consistent Hubbard parameters from density-functional perturbation theory in the ultrasoft and projector-augmented wave formulations. *Phys. Rev. B* **103**, 045141 (2021).
50. Timrov, I. et al. Electronic structure of Ni-substituted LaFeO<sub>3</sub> from near edge x-ray absorption fine structure experiments and first-principles simulations. *Phys. Rev. Res.* **2**, 033265 (2020).
51. Mahajan, R., Kashyap, A. & Timrov, I. Pivotal role of intersite Hubbard interactions in Fe-doped  $\alpha$ -MnO<sub>2</sub>. *J. Phys. Chem. C* **126**, 14353 (2022).
52. Timrov, I., Kotiuga, M. & Marzari, N. Unraveling the effects of intersite Hubbard interactions in spinel Li-ion cathode materials. *Phys. Chem. Chem. Phys.* **25**, 9061 (2023).
53. Binci, L., Kotiuga, M., Timrov, I. & Marzari, N. Hybridization driving distortions and multiferroicity in rare-earth nickelates. *Phys. Rev. Res.* **5**, 033146 (2023).
54. Haddadi, F., Linscott, E., Timrov, I., Marzari, N. & Gibertini, M. On-site and inter-site Hubbard corrections in magnetic monolayers: The case of FePS<sub>3</sub> and CrI<sub>3</sub>. *Phys. Rev. Mater.* **8**, 014007 (2024).
55. Bonfà, P. et al. Magnetostriction-driven muon localisation in an antiferromagnetic oxide. *Phys. Rev. Lett.* **132**, 046701 (2024).
56. Gelin, S. et al. Ternary oxides of s- and p-block metals for photocatalytic solar-to-hydrogen conversion. *PRX Energy* **3**, 013007 (2024).
57. Macke, E., Timrov, I., Marzari, N. & Ciacchi, L. C. Orbital-resolved DFT+*U* for molecules and solids. *J. Chem. Theory Comput.* **20**, 4824 (2024).
58. Timrov, I., Aquilante, F., Binci, L., Cococcioni, M. & Marzari, N. Pulay forces in density-functional theory with extended Hubbard functionals: from nonorthogonalized to orthogonalized manifolds. *Phys. Rev. B* **102**, 235159 (2020).
59. Kulik, H. & Marzari, N. Accurate potential energy surfaces with a DFT+*U*(R) approach. *J. Chem. Phys.* **135**, 194105 (2011).
60. Hsu, H., Umehoto, K., Cococcioni, M. & Wentzcovitch, R. First-principles study for low-spin LaCoO<sub>3</sub> with a structurally consistent Hubbard *U*. *Phys. Rev. B* **79**, 125124 (2009).
61. de Mendonça, J. P. A., Mariano, L. A., Devijver, E., Jakse, N. & Poloni, R. Artificial neural network-based density functional approach for adiabatic energy differences in transition metal complexes. *J. Chem. Theory. Comput.* **19**, 7555–7566 (2023).
62. Schubert, Y., Luber, S., Marzari, N. & Linscott, E. Predicting electronic screening for fast Koopmans spectral functional calculations. *npj Comput. Mater.* **10**, 299 (2024).
63. Cai, G. et al. Predicting structure-dependent Hubbard *U* parameters via machine learning. *Mater. Futures* **3**, 025601 (2024).
64. Yu, W. et al. Active learning the high-dimensional transferable Hubbard *U* and *V* parameters in the DFT+*U*+*V* scheme. *J. Chem. Theory Comput.* **19**, 6425 (2023).
65. Batzner, S. et al. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.* **13**, 2453 (2022).
66. Musaelian, A. et al. Learning local equivariant representations for large-scale atomistic dynamics. *Nat. Commun.* **14**, 579 (2023).
67. Bartók, A. P., Kondor, R. & Csányi, G. On representing chemical environments. *Phys. Rev. B* **87**, 184115 (2013).
68. Musil, F. et al. Physics-inspired structural representations for molecules and materials. *Chem. Rev.* **121**, 9759–9815 (2021).
69. Uhrin, M. Through the eyes of a descriptor: constructing complete, invertible descriptions of atomic environments. *Phys. Rev. B* **104**, 144110 (2021).
70. Batatia, I. et al. A foundation model for atomistic materials chemistry. *arXiv:2401.00096* <http://arxiv.org/abs/2401.00096> (2023).
71. Sit, P. H.-L., Car, R., Cohen, M. H. & Selloni, A. Simple, unambiguous theoretical approach to oxidation state determination via first-principles calculations. *Inorg. Chem.* **50**, 10259 (2011).
72. Kondor, R., Lin, Z. & Trivedi, S. Clebsch-Gordan nets: a fully Fourier space spherical convolutional neural network. *arXiv:1806.09231* <https://arxiv.org/abs/1806.09231> (2018).
73. Thomas, N. et al. Tensor field networks: rotation- and translation-equivariant neural networks for 3D point clouds. *arXiv:1802.08219* <https://arxiv.org/abs/1802.08219> (2018).

74. Grisafi, AandWilkins,D. MandWillatt,M. JandCeriotti,M. Atomic-scale representation and statistical learning of tensorial properties. *ACS Symposium Series* **1326**, 1–21 (2019).
75. Geiger, M. et al. e3nn/e3nn: 2022-04-13, <https://doi.org/10.5281/zenodo.3724963>. <https://zenodo.org/record/6459381#.YvDuJ-xBxH4>.
76. Merchant, A. et al. Scaling deep learning for materials discovery. *Nature* **624**, 80–85 (2023).
77. Nigam, J., Willatt, M. J. & Ceriotti, M. Equivariant representations for molecular Hamiltonians and N -center atomic-scale properties. *J. Chem. Phys.* **156**, 1–15 (2022).
78. Zhang, L. et al. Equivariant analytical mapping of first principles Hamiltonians to accurate and transferable materials models. *npj Comput. Mater.* **8**, 158 (2022).
79. Geiger, M. & Smidt, T. E. e3nn: Euclidean neural networks. *arXiv:2207.09453* <http://arxiv.org/abs/2207.09453> (2022).
80. Paszke, A. et al. *Pytorch: An Imperative Style, High-performance Deep Learning Library* 8024–8035 (Curran Associates, Inc., 2019).
81. Loshchilov, I. & Hutter, F. Decoupled weight decay regularization. *arXiv:1711.05101* <http://arxiv.org/abs/1711.05101> (2019).
82. Chakraborty, A., Dixit, M., Aurbach, D. & Major, D. Predicting accurate cathode properties of layered oxide materials using the SCAN meta-GGA density functional. *npj Comput. Mater.* **4**, 60 (2018).
83. Mahajan, R., Timrov, I., Marzari, N. & Kashyap, A. Importance of intersite Hubbard interactions in  $\beta$ -MnO<sub>2</sub>: a first-principles DFT+U +V study. *Phys. Rev. Materials* **5**, 104402 (2021).
84. Aydinol, M., Kohan, A., Ceder, G., Cho, K. & Joannopoulos, J. Ab initio study of lithium intercalation in metal oxides and metal dichalcogenides. *Phys. Rev. B* **56**, 1354 (1997).
85. Zhou, F., Marianetti, C., Cococcioni, M., Morgan, D. & Ceder, G. Phase separation in LiFePO<sub>4</sub> induced by correlation effects. *Phys. Rev. B* **69**, 201101(R) (2004).
86. Cococcioni, M. & Marzari, N. Energetics and cathode voltages of LiMPO<sub>4</sub> olivines (M=Fe, Mn) from extended Hubbard functionals. *Phys. Rev. Materials* **3**, 033801 (2019).
87. Kulik, H. & Marzari, N. A self-consistent Hubbard U density-functional theory approach to the addition-elimination reactions of hydrocarbons on bare FeO<sup>+</sup>. *J. Chem. Phys.* **129**, 134314 (2008).
88. Moore, G. et al. High-throughput determination of Hubbard U and Hund J values for transition metal oxides via linear response formalism. *Phys. Rev. Mater.* **8**, 014409 (2024).
89. Huber, S. P. et al. AiiDA 1.0, a scalable computational infrastructure for automated reproducible workflows and data provenance. *Sci. Data* **7**, 300 (2020).
90. Uhrin, M., Huber, S. P., Yu, J., Marzari, N. & Pizzi, G. Workflows in AiiDA: engineering a high-throughput, event-based engine for robust and modular computational workflows. *Comput. Mater. Sci.* **187**, 110086 (2021).
91. Giannozzi, P. et al. Quantum ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys.: Condens. Matter.* **21**, 395502 (2009).
92. Giannozzi, P. et al. Advanced capabilities for materials modelling with Quantum ESPRESSO. *J. Phys. Condens. Matter.* **29**, 465901 (2017).
93. Giannozzi, P. et al. Quantum ESPRESSO toward the exascale. *J. Chem. Phys.* **152**, 154105 (2020).
94. Perdew, J. et al. Restoring the density-gradient expansion for exchange in solids and surfaces. *Phys. Rev. Lett.* **100**, 136406 (2008).
95. Prandini, G., Marrazzo, A., Castelli, I. E., Mounet, N. & Marzari, N. Precision and efficiency in solid-state pseudopotential calculations. *npj Comput. Mater.* **4**, 72 (2018).
96. The SSSP library of the Materials Cloud: <https://www.materialscloud.org/discover/sssp/table/efficiency>.
97. Löwdin, P.-O. On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. *J. Chem. Phys.* **18**, 365 (1950).
98. Mayer, I. On Löwdin's method of symmetric orthogonalization. *Int. J. Quant. Chem.* **90**, 63 (2002).
99. Fletcher, R. *Practical Methods of Optimization* 2nd edn (Wiley, 1987).
100. Timrov, I., Marzari, N. & Cococcioni, M. HP—A code for the calculation of Hubbard parameters using density-functional perturbation theory. *Comput. Phys. Commun.* **279**, 108455 (2022).
101. Uhrin, M., Zadoks, A., Binci, L., Marzari, N. & Timrov, I. Dataset of self-consistent Hubbard parameters for Ni, Mn and Fe from linear-response. Materials Cloud Archive **2024.160** <https://archive.materialscloud.org/record/2024.160> (2024).
102. Uhrin, M. HubbardML: Version 0.2.0 <https://doi.org/10.5281/zenodo.1394209> (2024).

## Acknowledgements

We thank Francesco Aquilante, Mario Geiger, and Jigyasa Nigam for fruitful discussions. We acknowledge support by the NCCR MARVEL, a National Centre of Competence in Research, funded by the Swiss National Science Foundation (Grant number 205602). This work was supported by a grant from the Swiss National Supercomputing Centre (CSCS) under project ID s1073 (Piz Daint) and ID 465000416 (LUMI-G). This work has been partially supported by MIAI@Grenoble Alpes, (ANR-19-P3IA-0003).

## Author contributions

M.U. developed and trained ML model architecture; A.Z. wrote data processing code; L.B. and I.T. performed self-consistent calculations of Hubbard parameters; M.U., A.Z., L.B., N.M., and I.T. analyzed the results and wrote the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41524-024-01501-5>.

**Correspondence** and requests for materials should be addressed to Martin Uhrin or Iurii Timrov.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025