Acta Materialia 220 (2021) 117341

Contents lists available at ScienceDirect

Acta Materialia

journal homepage: www.elsevier.com/locate/actamat

Full length article

# A neural network potential for searching the atomic structures of pure and mixed nanoparticles. Application to ZnMg nanoalloys with an eye on their anticorrosive properties

P. Álvarez-Zapatero*, A. Vega, A. Aguado

*Departamento de Física Teórica, Atómica y Óptica, Universidad de Valladolid, E–47011 Valladolid, Spain*

ARTICLE INFO

*Article history:*
Received 2 March 2021
Revised 13 September 2021
Accepted 17 September 2021
Available online 24 September 2021

*Keywords:*
Atomistic simulation
Artificial neural networks
Density functional theory
Magnesium alloy
Corrosion

ABSTRACT

The accurate description of the potential energy landscape of moderate-sized nanoparticles is a formidable task, but of paramount importance if one aims to characterize, in a realistic way, their physical and chemical properties. We present here a Neural Network potential able to predict structures of pure and mixed nanoparticles with an error in energy and forces of the order of chemical accuracy as compared with the values provided by the theoretical method used in the training process, in our case the density functional theory. The neural network is integrated into a basin-hopping algorithm which dynamically feeds the training process. The main ingredients of the neural network algorithm as well as the protocol used for its implementation and training are detailed, with particular emphasis on those aspects that make it so efficient and transferable. As a first test, we have applied it to the determination of the global minimum structures of ZnMg nanoalloys with up to 52 atoms and stoichiometries corresponding to $MgZn_2$ and $Mg_2Zn_{11}$, of special interest in the context of anticorrosive coatings. We present and discuss the structural properties, chemical order, stability and pertinent electronic indicators, and we extract some conclusions on fundamental aspects that may be at the roots of the good performance of ZnMg nanoalloys as protective coatings. Finally, we comment on the step forward that the presented machine learning approach constitutes, both in the fact that it allows to accurately explore the potential energy surface of systems that other methodologies can not, and that it opens new prospects for a variety of problems in Materials Science.



## 1. Introduction

Structure determination in small nanoparticles remains as one of the fundamental problems in cluster research. On one hand, an accurate knowledge of a cluster structure is a prerequisite for a detailed understanding of its several physico-chemical properties, of broad interest in many nanotechnology applications; on the other hand, there are no direct experimental probes of structure for free-standing clusters, only indirect probes such as photoemission spectra, vibrational spectra, electron diffraction patterns, etc. [1–5], that have to be compared to theoretical predictions before obtaining a definitive structural assessment [6,7]. To make things worse, a theoretical determination of the global minimum (GM) structure of a cluster requires an exhaustive computational sampling of the potential energy surface. Although several global optimization methods, such as genetic [8,9] or Basin Hopping [10–13] algorithms, have been proposed to efficiently tackle this problem, they tend to be computationally impractical to sample the *ab initio* potential energy surface of all but the smallest nanoparticles. A well established practical solution to this problem consists of performing an exhaustive sampling of the approximate potential energy surface provided by an empirical potential (EP) to quickly determine trial structures that are then reoptimized at a first-principles density functional theory (DFT) level, and has come to be known as EP-DFT approach [14]. It introduces new problems, however: in order for this approach to be successful, the EP model must provide a reasonably accurate and transferable account of atomic interactions. In recent years, neural network (NN) potentials have emerged as a particular type of empirical potentials which may reach an unprecedented numerical accuracy, so combining global optimization

* Corresponding author.
*E-mail address:* pablo.alvarez.zapatero@uva.es (P. Álvarez-Zapatero).

https://doi.org/10.1016/j.actamat.2021.117341

with NN methods may be an optimal solution for the structure determination problem in clusters, taking into account the capabilities of presently available computers [15–21].

Machine Learning tools fall under several categories depending on the type of problem they are intended to tackle and how they are implemented. In this work, we focus on the so called supervised learning techniques, in which the computer is fed with a set of inputs and the desired outputs, so that it can find a logical rule that connects both. In practical terms, the computer will find the functional relationship between the numerical inputs and targets through a regression analysis, thus obtaining a continuous function. We have chosen the neural network approach as the particular method to perform the required regression. Neural Networks are computational algorithms inspired in the biological neural networks [22]. They are formed by an ensemble of interconnected units (or nodes) called artificial neurons, each of which receives a signal that propagates to the subsequent neurons through a non linear activation function. All Neural Networks have a common general structure: an input layer of nodes where the signal is created, followed by one or more layers of nodes (called hidden layers) through which the signal evolves in a non-linear way, and lastly one output layer where the signal dies providing the user with a result. We will be concerned with the simplest NN type: the feed-forward Neural Network [23–28], where the information evolves only in one direction, from the input layer to the output layer passing through each hidden layer once. Each node in a given layer is connected to all nodes of the subsequent layer via a non-linear matrix relationship, which must be optimized to give accurate output data during a training stage. In our case, the training involves fitting *ab initio* data including binding energies and atomic forces calculated on a large and diverse data pool of cluster structures. Despite its simplicity, the NN encloses an outstanding potentiality. According to the Universal Approximation Theorem [29,30], any continuous multidimensional function can be approximated to any desired degree of accuracy by increasing the number of neurons in a single hidden layer of a feed-forward neural network. In practice, the accuracy of the NN in interpolating the data is nevertheless limited by the finite size of the input data set and by the physical appropriateness of the input descriptors chosen by the user.

The structural and electronic properties of metallic nanoparticles in the size range up to several hundred atoms generally exhibit a marked non-monotonous dependence with size, to a large extent due to quantum-confinement effects. This fact complicates the training process, since unveiling such complex size-dependent behaviors is difficult without an explicit representation of the electronic degrees of freedom. Thus, developing a Neural Network potential to accurately describe the potential energy surface of small nanostructures is more challenging than for their bulk crystalline counterparts, where the absence of such non-monotonous electronic and geometric behaviors plus the dramatic reduction of structural relaxation possibilities, allows for a simpler and successful training. So far, considerable efforts have been devoted to develop NN techniques for bulk systems [31–35] but not as much for small nanostructures [36,37]. Besides, in most cases where nanostructures are the goal, the employed approaches in the training stage extrapolate from the bulk limit [38,39]. On the contrary, our Neural Network potential implementation is specifically tailored to train and test finite nanostructures comprising a wide range of sizes and chemical bondings, for both homoatomic clusters and nano-alloys with up to 3 different chemical elements.

In this work, we present both our computer implementation of the neural network approach for clusters, and its particular application to Zn-Mg nanoalloys with up to 52 atoms. Specifically, we focus on two compositions, $MgZn_2$ and $Mg_2Zn_{11}$, because they have been found to be optimal for use as sacrificial corrosion protective coatings [40–42]. We have previously reported the structures of small $Zn_xMg_{20-x}$ nanoalloys in the whole range of compositions [43], as well as the structures of equiatomic nanoalloys as a function of size [44], and analyzed their electronic structure in order to provide some preliminary clues about the enhanced corrosion protection capabilities of this alloy. We need to get accurate structural models of $MgZn_2$ and $Mg_2Zn_{11}$ before embarking into a dedicated study of their reactivity with external re-agents such as oxygen, chlorine or water, this being the main motivation for the present study.

## 2. The neural network approach and protocol for the structural search

The use of Neural Networks can potentially overcome the limitations and shortcomings of usual empirical potentials. When properly trained, Neural Network potentials can be used to accurately describe atomic interactions in arbitrary materials, so they become an ideal alternative in those systems where available empirical potentials are not able to capture the essential features of the true potential energy surface. Their main advantage relies on the flexibility offered by the machine learning technique: while classical interatomic potentials have a limited number of parameters to be tuned on a definite simple functional expression that itself does not account for all of the important geometrical and electronic effects, in the neural network approach one can, in principle, fit every functional form describing the real potential energy surface [29,30]. In order to take full advantage of this flexibility, we have written a Neural Network Fortran90 code specifically tailored to deal with cluster systems. As detailed in the following, the code is fed with separate training and testing data sets generated by an external code. Preferably, and for the sake of accuracy, those data sets are here generated at a first-principles level of theory, although other options would of course be possible. The code produces as output a NN potential providing both energies and analytic atomic forces, that can be later used by other external codes such as global optimization or molecular dynamics routines. At present, it can train both homoatomic and heteroatomic nanostructures of arbitrary size and with up to three different chemical species, though in this section we will describe its particular application to Zn-Mg binary nanoalloys.

### 2.1. Input and output layers. Structural descriptors

Our particular NN implementation for clusters is based on the Behler-Parrinello method [45–47]. In this method, the total cluster energy is written as a sum of atomic contributions:

$$E = \sum_{i}^{N} E_i, \tag{1}$$

where $N$ is the total number of atoms in the cluster. Each atomic energy forms the output layer of a separate feed-forward Neural Network dealing with that particular atom. The energy of atom $i$ is expressed in terms of its local atomic environment, described by adequate symmetry functions that form the input layer of the corresponding neural network. Although each atom is processed separately, atoms of the same kind are described by the same NN. This way, one has to train a single feed-forward Neural Network for every chemical species present in the system, whose separate outputs, $E_i$, are auxiliary quantities used to recover the total energy of the system.

The additivity assumption implied by Eq. (1) is to be considered an approximation (in the sense that it can not be demonstrated starting from fundamental quantum-mechanical laws such as Schrödinger equation), whose accuracy may be judged *a posteriori* by the quality of the training fit. We will show below that it is

accurate enough for the metallic nanoalloys dealt with in this paper. Its most important computational advantage is that it allows to generate a single neural network potential to describe nanoalloys of arbitrary size and composition [45], as the NN is dealing only with one particular atom at a given time. Machine learning models avoiding the additivity assumption and based on global descriptors [48–50] need to be fitted separately for each current system size or composition, which we consider an undesirable feature as our goal is to study a broad range of sizes or compositions.

Choosing an appropriate set of structural descriptors for the input layer is one of the key factors to successfully train the NN and obtain an accurate potential. The wiser the user is in providing the right physical descriptors (those that are statistically dominant in determining the cluster energy), the faster the training process will be, the lower the irreducible error of the fit, and fewer input descriptors will be needed. In our work, the symmetry functions are separately tailored for each chemical element, and are constructed taking into account the local atomic and chemical environments of the atom we are describing (an atom-centred approach) up to some cutoff radius. So in first place, we have to define a cutoff function $f_c$ around atom $i$, which defines the size of its atomic environment and the radial extension of the symmetry functions:

$$f_c(r_{ij}) = \begin{cases} \frac{1}{2}\left[cos\left(\frac{\pi r_{ij}}{r_c}\right)+1\right], & r_{ij} \le r_c \\ 0, & r_{ij} > r_c, \end{cases} \quad (2)$$

which smoothly decays to zero in value and slope at the cutoff radius $r_c$. The user has complete freedom in choosing the $r_c$ values, and can judge the appropriateness of the chosen cutoffs by the quality of the fit. In particular, for finite nanoscale systems which are our target in this paper, the cutoffs can always be chosen long enough so as to provide structural information about the complete set of atoms in the environment of a reference atom $i$. In bulk systems with significant long-range interactions (electrostatic, van der Waals, etc.), however, the energy of an atom $i$ is not expected to depend only on its local environment up to a predefined cutoff distance, so our NN energy expresion should be complemented with separate long-range energy terms [51]. We emphasize that our code, in the version presented in this paper, is specifically tailored to deal with finite-size atomic systems.

To describe the radial arrangement of atoms in the surroundings of atom $i$, radial symmetry functions are defined as a sum of products of Gaussians times the cutoff function:

$$G_i^r = \sum_j e^{-\eta(r_s - r_{ij})^2} f_c(r_{ij}), \quad (3)$$

where $\eta$ is a width parameter that determines the radial extension, and $r_s$ a shifting parameter that displaces the Gaussians to improve sensitivity at specific radii. For homoatomic clusters, the sum runs over all the neighbors within the cutoff sphere around atom $i$ and the $G_i^r$ functions incorporate purely geometrical information. In the case of binary nanoalloys (formed by $A$ and $B$ chemical species), separate radial functions are defined for $A-A$, $A-B$ and $B-B$ distance distributions in order to incorporate chemical ordering information as well.

Describing just the radial distribution of the atoms is not sufficient to obtain a suitable fingerprint of the atomic environment. A description of the angular distribution of the neighbors of atom $i$ is accomplished by employing the following angular symmetry functions:

$$G_i^\theta = 2^{(1-\zeta)} \sum_{j,k\neq j} (1+\lambda cos\theta_{ijk})^\zeta e^{-\eta(r_{ij}^2+r_{ik}^2+r_{jk}^2)} f_c(r_{ij}) f_c(r_{ik}) f_c(r_{jk}), \quad (4)$$

where $\theta_{ijk}$ is the angle conformed by the atomic triplet (i; j; k), $\eta$ determines again the width of the gaussian functions, and $\lambda = \pm 1$ is used to invert the shape of the cosine function for an improved sensitivity at different values of $\theta_{ijk}$. When dealing with binary nanoalloys, separate angular functions are employed to distinguish the chemical nature of the two neighbors of atom $i$ in each triplet (i; j; k). Therefore, there are three different versions of each $G_i^\theta$ function, as the (j; k) neighbors of a given triplet can be of *AA*, *AB* or *BB* types.

All the parameters appearing in these functions, as well as the total number of symmetry functions employed, must be customized beforehand to accurately characterize the cluster structures. Using too few of these functions would result in an incomplete description of the structural and chemical environment and, thus, to a poor relationship between descriptors and targets. On the other hand, the set of symmetry functions should also be kept as small as possible to increase the computational efficiency of the training and testing stages. One should in particular avoid that different symmetry functions are linearly related.

It is interesting to compare the structural information content of our local descriptors to that of global descriptors such as the Coulomb matrix [48–50], which essentially compiles the pair distances $r_{ij}$ between all atoms in the molecule. Therefore, in the limit of long enough cutoffs, the structural information contained in our radial descriptors for atom $i$ is essentially the same as that in the $i$th row of the Coulomb matrix. However, in global descriptor methods [48–50] it is the whole Coulomb matrix that is fed into the machine learning code in a single step, while in our local approach we are feeding, so to say, each row of that matrix at a time. The consequence is that, while the Coulomb matrix implicitly contains complete angular information, we have to separately add that information through the angular local descriptors, which parallels the inclusion of 3-body terms in a cluster expansion. We demonstrate below that the accuracy of our fitting is of the same quality as that obtained with recent methods based on the Coulomb matrix descriptor [50], even if our training set contains a wide variety of system sizes and compositions and thus represents a much more stringent transferability test. This finding demonstrates that 2-body and 3-body interactions, together with the non-linearity of the NN, suffice to faithfully represent the metallic nanoalloys considered in this paper. This is a very interesting conclusion because with the local method we can fit a single potential for arbitrary sizes and compositions, as explained above. Another advantage of the local descriptors is that they are easily generalized to deal with chemical order in nanoalloys, by simply replicating the descriptors for each possible pair or triplet of atoms in the nanoalloys.

### 2.2. Architecture and training of the neural network

Both the number of hidden layers and the number of nodes in each layer are set by the user, and together define the size of the neural network. All of the signal values in a given layer $k$ are collected in a vector array $\mathbf{X}_k$. The signals evolve from a given layer to the next one in a feed-forward way, according to the equation:

$$\mathbf{X}_{k+1} = \sigma(\mathbf{X}_k \mathbf{W}_{k,k+1} + \mathbf{B}_{k+1}), \quad (5)$$

where $\mathbf{W}_{k,k+1}$ is the matrix of weights that connects the nodes between layers $k$ and $k+1$, $\mathbf{B}_{k+1}$ is a bias term and $\sigma$ is the non-linear activation function which builds up the signals of the $k+1$ layer. All matrices and vector arrays are real-valued, and only the signal in the input layer must be provided by the user.

The activation function employed in our work is the Swish function [52], $\sigma(x) = \frac{x}{1+e^{-x}}$, except for the output layer where a linear function was used. Swish belongs to the family of rectifier functions, whose distinguishing feature is to be unbounded, as opposed to many previous implementations which relied on bounded sigmoid functions. We found rectifiers to produce a more accurate

fitting to the *ab initio* data in the training stage (see below), as compared to sigmoids such as the hyperbolic tangent function (see Table S1 in the ESI which compares the performance of different activation functions). Our observation is in line with the known superiority in performance of Swish over sigmoids in regression problems [52–54].

The optimal values for the several weight matrices and bias arrays are obtained by training the Neural Network to match the *ab initio* values of energies and forces calculated on a large data set of cluster structures. This is done by minimizing, in a least-squares sense, the value of the following cost function (or objective function):

$$\Gamma = \frac{1}{2M}\sum_{k}^{M}\left[\left(\frac{E_k^{DFT}-\widehat{E}_k}{N_k}\right)^2 + \epsilon\sum_{l}^{N_k}\sum_{\alpha}^{3}\left(\frac{F_{k,l\alpha}^{DFT}-\widehat{F}_{k,l\alpha}}{3N_k}\right)^2\right], \quad (6)$$

where $M$ is the training set size (in our case the number of cluster structures included in the training set), $N_k$ the size of cluster $k$ in the data pool, and $\epsilon$ a scaling factor (set to 0.05) used to obtain balanced relative errors in energies and forces. The NN energy of structure $k$, $\widehat{E}_k$, is given by the sum of the individual atomic contributions $E_{k,i}$ as stated by the Behler-Parrinello approach (Eq. (1)). The NN force on atom $l$ of structure $k$ is then obtained as:

$$\widehat{F}_{k,l\alpha} = -\frac{\partial\widehat{E}_k}{\partial r_{k,l\alpha}} = -\sum_{i}^{N_k}\frac{\partial E_{k,i}}{\partial r_{k,l\alpha}} = -\sum_{i}^{N_k}\sum_{s}^{M_s}\frac{\partial E_{k,i}}{\partial G_{i,s}}\frac{\partial G_{i,s}}{\partial r_{k,l\alpha}}, \quad (7)$$

where $M_s$ is the number of symmetry functions associated with atom $l$ of cluster $k$, and $\alpha$ denotes the cartesian direction. Thus, the force on atom $l$ depends on $\frac{\partial E_{k,i}}{\partial G_{i,s}}$, that is the derivative of the output layer value with respect to the symmetry functions in the input layer, and on $\frac{\partial G_{i,s}}{\partial r_{k,l\alpha}}$, that is the derivatives of the symmetry functions with respect to the cartesian coordinates. Both terms are obtained analytically, which makes our code computationally very efficient. Full details about the analytic forces are provided in the ESI. Notice that the atomic forces are not explicitly provided by the output layer of the NN (as they can be analytically obtained from the atomic energies in the output layer), yet they are included in the cost function to explicitly train the NN towards accurate forces as well.

The minimization of the cost function $\Gamma$ is afforded by employing the Nadam algorithm [55,56], a gradient descent method that exploits the first- and second-order moments of the gradients to improve computational performance and stability, based on the reliable Adam [57] algorithm. The gradients of $\Gamma$ with respect to the weights and biases appearing in Eq. (5), needed by Nadam, are calculated with the backpropagation algorithm [58]. An important technical point here is the rate at which the cost function is updated, as it significantly impacts the efficiency and accuracy of the minimization problem. In NN jargon, the evaluation of the whole dataset once is called an epoch. In the so-called batch gradient descent, the weights and biases are updated after one epoch, producing a relatively smooth error surface, but being very slow and requiring a lot of memory for large data sets. In stochastic gradient descent, one updates the variables after evaluating each individual in the training set. Being the fastest update rate, it results in a noisy gradient, but notice that some noise is not necessarily harmful since it helps the algorithm in escaping from shallow local minima. A good compromise is the mini-batch gradient descent, which updates the weights based on a small subset of individuals (32, 64 or 128 are typical choices in parallel computations). After one epoch, these subsets (or mini-batches) are randomly reselected from the whole data set. This way the gradients are more robust compared to stochastic gradient descent while having some white noise. This is the particular implementation used in our work, with a mini-batch size of 64 individuals. Other technical details are explained in the ESI.

A final important point concerning the training stage is the stopping criterion, which should take into consideration the bias-variance tradeoff problem [59,60]. Our code randomly divides the data set provided by the user into two parts: 90% of the structures go to the training set on which the cost function $\Gamma$ is evaluated; the remaining 10% is used as a testing set, containing data which are not explicitly used to train the NN. The training process is continued until the cost function error *on the test set* is minimal, a procedure known as early stopping method. This strategy aims to minimize both the bias and variance errors, so that both under- and over-fitting problems are avoided. The final error of this training strategy will mostly be of an irreducible type, i.e. one that can not be significantly reduced by increasing the complexity of the NN of by running the minimization for a longer time. In fact, the irreducible error faithfully represents the inherent noise of the fitting, due for example to a sparse data set or, more importantly, to the non-existence of a deterministic relationship between inputs and targets. It can only be reduced by the user (not by the machine) by providing more explicit data or more appropriate descriptors in the input layer.

### 2.3. Symmetry issues

In a recent report, Chmiela et al. [49] have developed the sGDML model, a gradient-domain machine learning force-field that makes explicit use of symmetries in order to reduce the size of the training set and also to improve the accuracy and efficiency of calculations performed on high-symmetry molecules. In our current implementation of the neural network, we have not explicitly exploited symmetry in order to improve the efficiency of the energy calculations. It just turns out that the great majority of geometries sampled during a global optimization run on Zn-Mg nanoalloys have no rotational symmetries (i.e. most of them have $C_1$ point group), so computational savings would not be substantial for the particular target system in this paper. For more symmetric systems, the point-group symmetry could be exploited by feeding the input layer of the NN with just one atom from each symmetry orbit, thus reducing the explicit number of atomic energy calculations. In effect, two atoms equivalent by symmetry must have equal contributions to the total energy, and their force vectors must be related by a group operation. This improvement might be implemented in future versions if considered worthwhile.

Notwithstanding this computational efficiency issue, our code satisfies many symmetry properties by construction, and so accounts for all relevant effects associated to symmetry, as the sGDML model does. For example, homogeneity and isotropy of space imply that the total energy of an isolated system must be invariant against global translations and rotations of the system. This is ensured by construction in our code, as the local descriptors depend only on pair distances or triplet angles, both quantities being independent on the coordinate system. Similarly, our NN force-field is conservative by construction, i.e. it will satisfy the conservation of the total energy of an isolated system during its time evolution if used in MD simulations, because the exact consistency relations between energies and forces is analytically ensured (see previous section). Finally, the physical effects of point-group symmetry will be captured also by construction of our local descriptors: in effect, two atoms that are equivalent by symmetry will have exactly the same set of local descriptors, and so the code will assign exactly the same atomic energy to both of them, and forces which are related by the corresponding group operation.

The authors of the sGDML model take advantage of permutation-inversion symmetries to reduce the size of the data set and thus optimize the cost of the training process. In

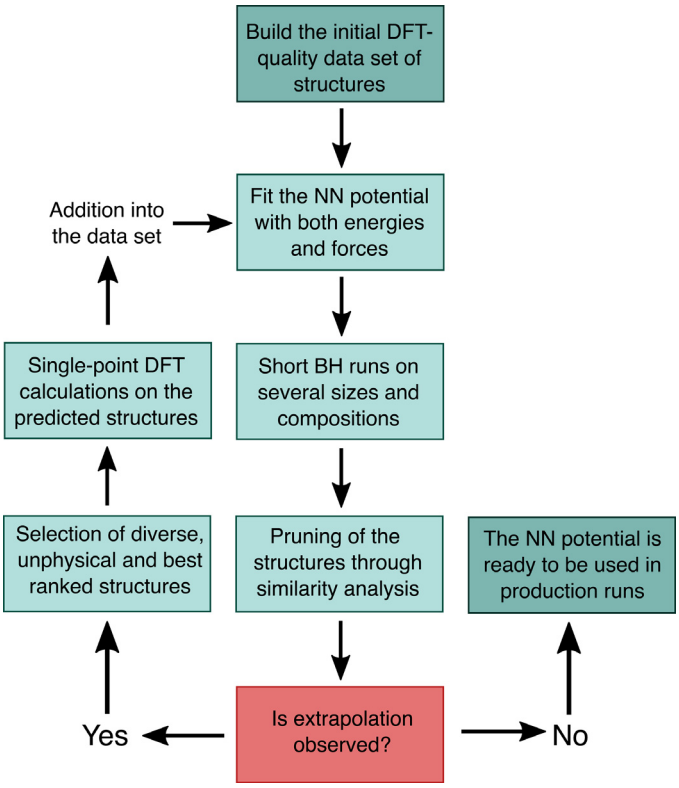


**Fig. 1.** Flow diagram summarizing our training protocol. See the main text for full details about each different step of the process.

effect, permutations of identical atoms define physically equivalent but different local minima on the PES, and there is no point in explicitly sampling two equivalent basins. Similarity of isomers is quantified through their euclidean distance, properly minimized with respect to permutation-inversion operations, which requires the use of bi-partite and multi-partite matching techniques [49]. We have also considered similarity of isomers when building our training data set, with the same goal of defining a training set that does not contain duplicate or too similar structures. We have done this with a less elaborate, home-built, similarity indicator, which is invariant to permutation-inversion operations thus avoiding the complications inherent to the euclidean distance measure. This similarity indicator is described in the ESI.

A flow diagram summarizing the major milestones of our training protocol is offered in Fig. 1.

### 2.4. Application to Zn-Mg nanoalloys

We employ a total of 70 symmetry functions (both radial and angular) in the input layers of the two different Neural Networks that deal with Mg and Zn atoms, respectively (details of these functions are provided in the ESI). This number of symmetry functions was optimized through a careful correlation analysis in order to reduce the number of redundant descriptors, that is, we discarded those which in our initial descriptor set had a large correlation with descriptors already considered.

We have generated two different NN potentials, in order to provide a larger amount of plausible low-energy structures, i.e. to enhance structural diversity. Both Neural Networks have three hidden layers, but different architectures: $70 \times 40 \times 20 \times 10 \times 1$ (amounting to 3881 adjustable parameters) in the first one, and $70 \times 45 \times 30 \times 20 \times 1$ (5216 parameters) in the second one. Those architectures were selected since they provide sufficient accuracy without resulting in over-fitting.

We have implemented our NN potentials into a module of the freely available GMIN code [61]. We use this program to dynamically enlarge the size of our data set during an iterative and automated training process involving Basin Hopping (BH) runs [10,11]. In each iteration, the currently fitted NN force field samples the potential energy surface of the nanoalloy through BH runs, comprising 35000 steps and including both random and swap moves (those that exchange the chemical identity of a randomly chosen Zn-Mg pair of atoms). We then feed the data set with the new structures that the NN predicts, which involves single-point DFT calculations on those new structures, and start a new training on the enlarged dataset. During the initial iterations, i.e. when the NN is not yet sufficiently trained, we found that it easily tends to explore areas of the potential energy surface not covered by the initial data set. In other words, the NN *extrapolates*, thus failing to provide reliable structures. With relatively short BH runs one can easily check the quality of the current NN potential and find new relevant structures to feed the data set. This loop is continued until extrapolation is no longer observed and the NN produces reasonable structures. Notice that global optimization runs are to be preferred over other options such as short molecular dynamics runs, as our ultimate goal is to use the NN force field in a global optimization aimed to locate the global minimum structure of a cluster or nanoalloy, and for that we need global accuracy in the whole PES, rather than extensive sampling of local regions of the PES.

The initial data set was built upon the structures provided in previous works on Zn-Mg nanoalloys [43,44]. It was dynamically enlarged as explained above up to a final size of 49000 nanoparticles in the size range between 4 and 100 atoms, containing both pure Zn and Mg, and (mostly) Zn-Mg nanoalloys of all possible composition ratios and with different chemical orderings. The final Neural Network potentials were fitted to reproduce both energies and forces on this huge, DFT-quality, data set.

With the two resulting NN potentials, we performed a Basin Hopping global optimization search for Zn-Mg nanoalloys of sizes ranging from 6 to 52 atoms and stoichiometries corresponding to $Zn_2Mg$ and $Zn_{11}Mg_2$. We performed 200000 BH steps for each nanoalloy for each of the two different potentials. BH moves comprised either a random change of the atomic coordinates or a swap move, the latter amounting to a 20% of the total BH moves. If no lower energy structure is identified during 10000 consecutive steps, the cluster structure is re-seeded to a random one in order to enhance the sampling. A total amount of around 100 structures (the most stable ones) was saved from both potentials after removing duplicates in the two lists. The resulting list of 100 structures was then relaxed at the Kohn-Sham density functional theory (KS-DFT) level by using the SIESTA code [62]. Exchange-correlation effects were treated within the generalized gradient approximation of Perdew, Burke and Ernzerhof (PBE) [63], and norm conserving pseudopotentials [64,65] including non-linear core corrections [66] were used to represent core-valence interactions. The valence active space includes $3d^{10}4s^2$ electrons for Zn, and $3s^2$ for Mg. The size of the basis set was double-zeta plus two polarization orbitals (DZP2). The clusters were placed in a cubic supercell of 30Å of side. The reliability of this SIESTA setup has been demonstrated in our previous works (the single-point *ab initio* calculations needed during the training stage were performed with the same setup). The structures were relaxed without geometry constraints until the force on each atom was smaller than 0.01 eV/Å. Finally, for the best structure obtained after DFT relaxation, an exhaustive homotopic search was performed with the Neural Network potential in order to better sample the chemical ordering. This was accomplished through a BH search of 200000 steps and with only swap moves. The best configurations found were then relaxed again with SIESTA.

**Table 1**
Mean absolute errors in cohesive energy $E_{coh}$ (in meV/atom) and atomic forces (in eV/Å) on the train and test sets for the two Neural Network potentials.

| | NN with 3881 parameters | | NN with 5216 parameters | |
|---|---|---|---|---|
| | Train set | Test set | Train set | Test set |
| $E_{coh}$ | 10.80 | 11.73 | 8.63 | 9.97 |
| Force | 0.14 | 0.14 | 0.14 | 0.15 |

## 3. Results and discussion

### 3.1. Quality performance of the neural network

Table 1 shows the quality of the fit of our two Neural Network potentials. Specifically, the table displays mean absolute errors (MAE) in cohesive energy (binding energy per atom) and forces in both training and testing sets. The cohesive energy of a nanoalloy with $N$ atoms is defined as $E_{coh}(N) = (E(\mathrm{Mg})N_{\mathrm{Mg}} + E(\mathrm{Zn})N_{\mathrm{Zn}} - E_N)/N$, where $E_N$ is the energy of the nanoalloy formed by $N_{\mathrm{Mg}}$ and $N_{\mathrm{Zn}}$ Mg and Zn atoms, whose atomic energies are $E(\mathrm{Mg})$ and $E(\mathrm{Zn})$, respectively.

The energetic MAE of both NN's is around 10 meV/atom $\approx 1$ kJ/mol, an error size usually considered as representative of *chemical accuracy*. This means that both NN's reproduce PBE energies within an error that is already smaller than the inherent error of PBE functional in reproducing experimental cohesive energies of metals [67], which suggests the NN potentials are as accurate as one may wish them to be (there is no point in requiring a NN fitting error significantly smaller than the inherent error of the *ab initio* reference data) and, in particular, justifies the additivity assumption (Eq. (1)). The MAE of the forces, around 0.15 eV/Å, is also very small as we have checked that forces of that magnitude are induced by atomic displacements of around 0.04 Å, so the geometries predicted by the NN will be very accurate. We additionally notice that the second NN, while containing around 35% more adjustable parameters than the first one, improves the MAE by a mere 1%, so we conclude that both NN's are similarly accurate, and that the error can no longer be substantially reduced by increasing the complexity of the neuronal connections. We believe it is safe to conclude that the remaining error is mostly of an irreducible type. Considering the dynamical adjustment of the training set size employed in our work (see previous section), the irreducible error can not be significantly improved by increasing the size of the training set. It could only be improved by providing additional physical descriptors in the input layer. Right now, the NN is trying to correlate the energy contribution of a given atom with the local geometric and chemical environments of that atom. However, the stability of metal clusters is well known to be additionally influenced by electronic shell effects (associated with the discrete nature of the electronic eigenvalue spectrum of finite systems), which are more dependent on global properties such as the cluster shape, than on local atomic descriptors. It is plausible that including such global descriptors explicitly in the input layer would further diminish the fitting error. Nevertheless, our results show that local descriptors suffice for this particular nanoalloy. As an additional accuracy check, we demonstrate in the ESI that our Neural Network potential clearly outperforms the empirical Coulomb-corrected Gupta potential developed in a previous work [44] in proposing better candidate structures for DFT reoptimization.

### 3.2. Putative global minimum structures of ZnMg nanoalloys with nominal compositions $Zn_2Mg$ and $Zn_{11}Mg_2$

The putative GM structures located in this work are shown in Figs. 2, 3 and 4. In general, their skeletal geometries coincide with those of the pure Zn atomic clusters [7]. $Zn_4Mg_2$ is conformed by 2 tetrahedral units sharing one edge. The GM structure of $Zn_6Mg_3$ is a tri-capped trigonal prism (TTP), and all nanoalloys with $N = 12 - 15$ atoms are obtained by adding atoms to that TTP unit. The first structure with an internal core atom is $Zn_{12}Mg_6$, and is based on a distorted 13-atom decahedron with adatoms capping its square facets, the distortions increasing the coordination number of some adatoms. The GM structure of $Zn_{14}Mg_7$ is based on a $C_3$ twisted pyramid with a dangling atom attached to one of its corners, an exotic structure also found for $Zn_{21}$ [7]. From this size on, the GM structures tend to be more amorphous-like due to the size-mismatch between the two atomic species, although many of them locally display a recognizable tendency towards decahedral packing (with the orientations chosen in the figures, decahedral units can be visually appreciated in clusters with 30, 36, 39 and 51 atoms, for example). Notable exceptions to this general trend are $Zn_{18}Mg_9$ and $Zn_{33}Mg_6$, which are poly-tetrahedral structures

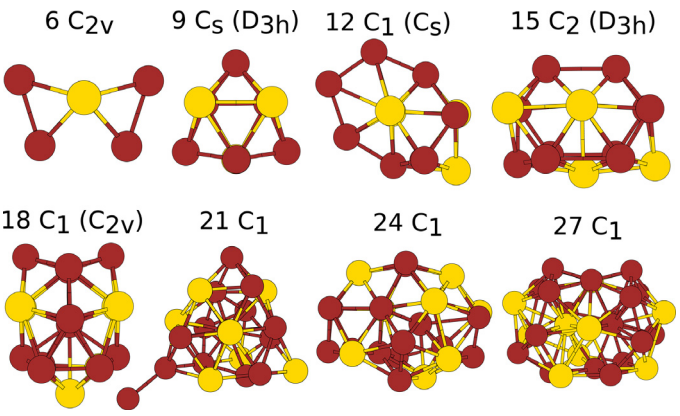


**Fig. 2.** Putative GM structures and approximate point group symmetries of $Zn_2Mg$ nanoalloys with $N = 6 - 27$ atoms. Brown and golden spheres represent Zn and Mg atoms, respectively. For some clusters, we add in brackets the point group of the corresponding homo-atomic cluster in order to better appreciate the symmetry of the atomic skeleton.

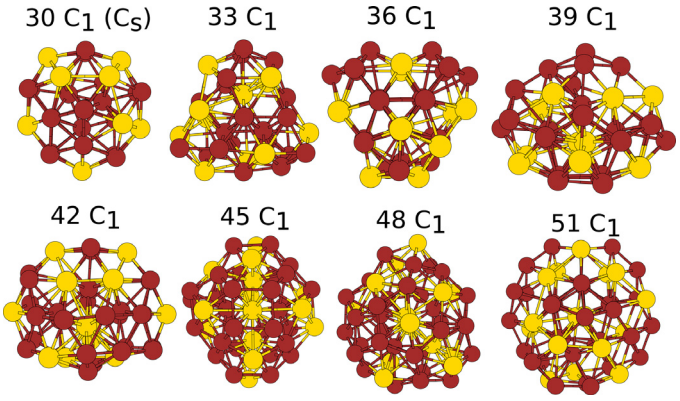


**Fig. 3.** Putative GM structures and approximate point group symmetries of $Zn_2Mg$ nanoalloys with $N = 30 - 51$ atoms. Rest of the caption as in Fig. 2.

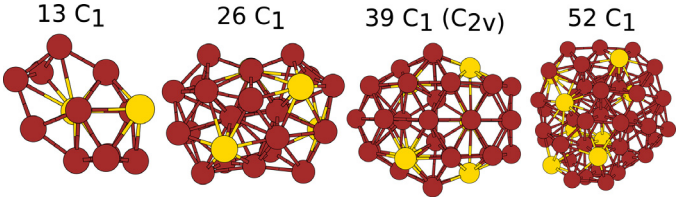


**Fig. 4.** Putative GM structures and approximate point group symmetries of $Zn_{11}Mg_2$ nanoalloys with $N = 13 - 52$ atoms. Rest of the caption as in Fig. 2.

composed of icosahedral and Frank-Kasper polyhedral units. The number of core atoms obviously increases with cluster size: two core atoms first appear at size $N = 26$, and the bigger clusters considered here have 8 core atoms.

Next we describe the chemical ordering trends. There are three important factors that conjointly determine the tendency towards segregation or mixing in Zn-Mg nanoalloys: (1) the bulk cohesive energy of Mg is 12% larger than that of Zn. Assuming that the relative strength of Mg-Mg and Zn-Zn bonds is maintained in the nanoalloys, this factor would tend to maximize the number of Mg-Mg bonds, and so would favor Mg@Zn segregation; (2) however, interatomic distances in bulk Mg are around 30% longer than in bulk Zn. In order to minimize bond strain, the small element tends to segregate to the cluster core, so this factor alone would favor the opposite Zn@Mg segregation. The observation that energetic and steric effects oppose each other already suggests that segregation trends are not clear; (3) finally, charge transfer effects introduce an ionic bonding component that tends to maximize the number of Zn-Mg bonds and thus promotes mixing. In order to quantify the degree of mixing, we have evaluated the following indicator for each of the GM structures:

$$p_{\mathrm{mix}} = \frac{N_{\mathrm{AB}} - N_{\mathrm{AB}}^{m}}{N_{\mathrm{AB}}^{M} - N_{\mathrm{AB}}^{m}}, \tag{8}$$

where $N_{\mathrm{AB}}$ is the number of Zn-Mg bonds. $N_{\mathrm{AB}}^{M}$ and $N_{\mathrm{AB}}^{m}$ are respectively the maximum and minimum values that $N_{\mathrm{AB}}$ can take for that particular *frozen* nuclear skeleton. These two last numbers are obtained from swap-only BH runs without allowing for structural relaxation. The parameter thus defined is normalized between zero and one, with a value $p_{\mathrm{mix}} = 1$ indicating the maximum degree of mixing that a given nuclear skeleton allows; the value $p_{\mathrm{mix}} = 0$, on the contrary, is associated with left-right segregated structures, displaying an interface separating pure zinc from pure Mg sides. The calculation of this indicator demonstrates that most of the GM structures found for the $Zn_2Mg$ composition are indeed maximally mixed, and in any case all of them have $p_{\mathrm{mix}} > 0.9$. Regarding the $Zn_{11}Mg_2$ composition, all GM structures have $p_{\mathrm{mix}} > 0.8$. Thus, concerning chemical order, the dominant building rule for ZnMg nanoalloys is to maximize mixing. This is achieved, as seen in the figures, by preferentially placing the minority component Mg atoms in cluster sites with a higher coordination number, but at the same time avoiding as much as possible direct Mg-Mg contacts. More in detail, we have checked that the shell of ZnMg nanoalloys tends to be well triangulated, each surface site being connected to either 5, 6 or 7 surface sites. With very few exceptions, all of the 5-fold sites are occupied by Zn atoms, while Mg atoms occupy the 6-fold and 7-fold sites. This is a strong trend common to all clusters considered in this paper, although is sometimes incompatible with the global 2:1 composition ratio. For example, we have observed that a Mg atom may occupy a 5-fold site in the shell of $Zn_2Mg$ nanoalloys when occupation of any of the remaining 6-fold or 7-fold sites would imply the creation of a Mg-Mg bond.

The main secondary factor that competes with the maximal mixing rule seems to be steric in nature. In effect, the internal core sites tend to be under a high compressive stress for most of the atomic packings usually observed in small clusters, and occupying them with the bigger Mg atoms would result in a less dense atomic packing because of the size mismatch effect, thus decreasing the cluster stability. Additionally, the coordination volume around internal sites is smaller for Zn-richer compositions. Therefore, the first Mg core atom for the $Zn_{11}Mg_2$ nanoalloys is not observed until size $N = 52$. This is clearly the reason why $Zn_{11}Mg_2$ nanoalloys do not achieve a $p_{\mathrm{mix}} = 1$ value, even if the Mg atoms consistently occupy well separated surface sites with the maximum possible surface coordination: the mixing could only increase if more Mg atoms were at core sites. The same trend operates for the 2:1 composition although is relatively less important there. To summarize, there is a slight Mg-enrichment (beyond the nominal composition) of the surface in Zn-Mg nanoalloys, enforced by the size mismatch effect, that competes with the maximal mixing trend. The two factors together consistently explain the detailed chemical ordering trends in these nanoalloys.

Bulk $MgZn_2$ crystallizes in one of the $AB_2$ Laves phases, which are notorious by a number of reasons. The three simplest Laves phases, occurring for $MgCu_2$ (cubic C15 phase), $MgZn_2$ (hexagonal C14 phase) and $MgNi_2$ (hexagonal C36 phase), are tetrahedrally close-packed structures where the interstices are exclusively tetrahedral. They are all obtained by packing Frank-Kasper coordination polyhedra [68], and differ just in the stacking sequence of those polyhedral units. The coordination number (CN) of the smaller Zn atom is 12, and its coordination polyhedron is an icosahedron formed by 6 Zn and 6 Mg atoms; CN=16 for the larger Mg atom, and its coordination polyhedron is a Z16 Frank-Kasper polyhedron formed by 12 Zn and 4 Mg atoms. Given the importance of Laves phases in many materials processes an in particular in corrosion problems [68], it is interesting to analyze to what extent the small clusters studied in this paper have developed the bulk-like properties of the Laves phases.

To start with, we have explicitly checked that $p_{\mathrm{mix}} = 1$ for both cubic and hexagonal bulk Laves phases, so at least concerning this global tendency towards a maximally mixed chemical ordering, even the smallest $MgZn_2$ nanoalloys are already bulk-like. Another similarity between nano- and bulk alloys is the larger total coordination number of Mg atoms as compared to Zn atoms, as explained in the previous paragraphs. In particular, the total coordination numbers of core atoms are already close to bulk values: on average, we find CN=11 for Zn core atoms and CN=15 for Mg core atoms in most cases, although there are a few examples where the bulk CN values are strictly recovered. The very small clusters here considered are thus slightly less compactly packed as compared to the bulk limit. We also observe that the first coordination layer around internal atoms tends to contain similar numbers of Zn and Mg neighbors around Zn, but a substantially larger number of Zn atoms around Mg. Concerning the coordination polyhedra surrounding internal atoms, these tend to be different to those in the bulk Laves phase, mainly because most of the clusters are still based on distorted decahedral packing, which is different and less dense than the poly-tetrahedral packing realized in the bulk. Nevertheless, at least for the special $Zn_{18}Mg_9$ and $Zn_{33}Mg_6$ clusters, which are based on poly-tetrahedral packing, the coordination polyhedra are indeed distorted versions of their bulk counterparts (see Fig. 5), the distortions being expected due to the surface rounding and relaxation effects. Other distinguishing structural features of the bulk phases are of a longer spatial range and can not be realized in small clusters. For example, there is no possible distinction between C14, C15 and C36 phases as there are not yet enough polyhedral units to define a stacking sequence. Similarly, the compositional layering observed in the bulk is hard to appreciate in these very small systems, although at least $Mg_{15}Zn_{30}$ shows clear signatures of alternating stacked Zn/Mg sheets. In summary, we conclude that there are significant similarities in the short-range structural descriptors of cluster and bulk systems.

### 3.3. Electronic properties of $Zn_2Mg$ nanoalloys

We have evaluated the vertical ionization potential (IP) and electron affinity (EA) of all GM structures through a $\Delta-$SCF calculation, i.e. by explicitly calculating the cation or anion state at the optimal geometry of the neutral cluster and taking the corresponding total energy differences. We have additionally computed the fundamental gap, defined as $E_{\mathrm{GAP}}$=(IP-EA), which is twice the

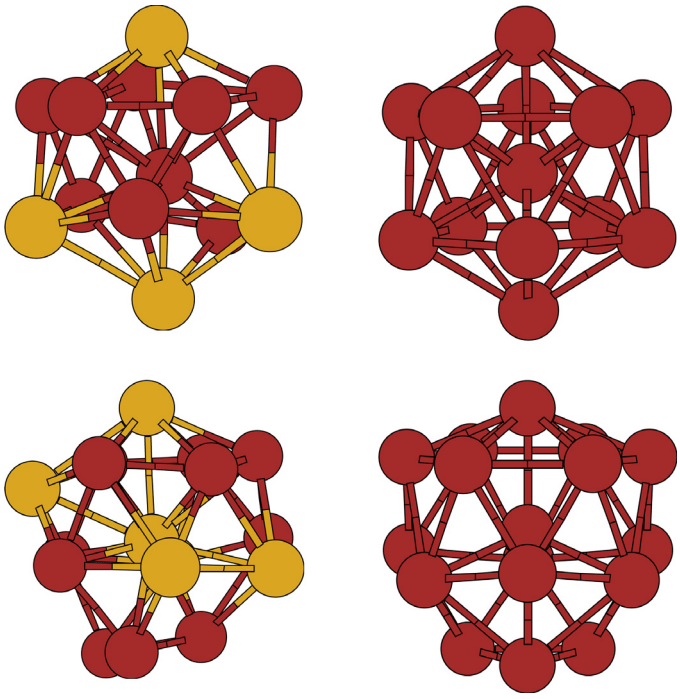

**Fig. 5.** The top row compares the coordination polyhedron of the internal Zn atom in $Zn_{18}Mg_9$ with a perfect Z12 polyhedron, i.e. an icosahedron. The bottom row shows the corresponding comparison between the coordination polyhedron of the internal Mg atom in $Zn_{18}Mg_9$ and a perfect Z16 Frank-Kasper polyhedron.

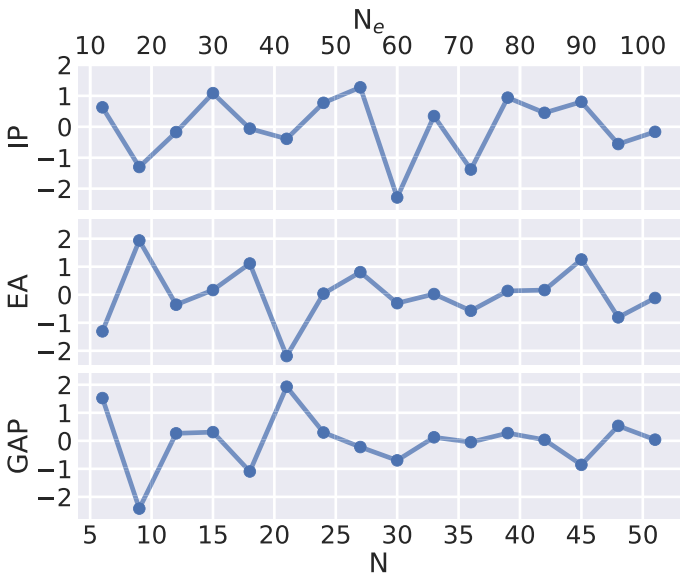


**Fig. 6.** Electronic stability indicators of $Zn_2Mg$ nanoalloys. Standardized vertical ionization energy, electron affinity and fundamental gap as a function of cluster size $N$ (lower scale) or the number of electrons $N_e$ (upper scale) are shown.

chemical hardness. These three quantities are taken as indicators of electronic stability: large gap values identify those clusters that are more stable against both oxidation and reduction processes and, as such, display an enhanced electronic stability; also a large IP or a low EA are indicative of a particularly stable configuration due to the reluctance to releasing or accepting one electron, respectively. According to a simple Jellium picture of delocalised electrons [69,70], it is the electronic stability that drives the thermodynamic stability of the cluster, and clusters displaying an electronic shell closing are thus expected to be particularly stable. The spherical Jellium model predicts these electronic shell closings to occur for metallic clusters with $N_e = 8, 20, 34, 40, 58, 68-70, 92, \cdots$ electrons. Jellium models allowing for ellipsoidal deformations of the confining ionic potential predict additional (sub-shell) closings due to the splitting of angular momentum multiplets in a non-spherical potential. Now, considering that Zn and Mg atoms are both divalent, the number of delocalized electrons increases in steps of 6 along the $Zn_2Mg$ series, and in steps of 26 along the $Zn_{11}Mg_2$ series, so none of the $N_e$ values associated with a spherical electronic shell closure occurs in our cluster sample. Moreover, the non negligible ionic contribution to bonding associated with charge transfer causes Zn-Mg nanoalloys to slightly deviate from a pristine jellium picture, as shown in our previous work [44]. All taken together, it can make it hard to discern marked features in the size evolution of the electronic indicators.

As usual in small metal clusters, the IP displays a globally decreasing trend as a function of size (meanwhile the EA shows an increasing envelope), which is steeper for smaller sizes. The slope of these indicators allows to identify two different size ranges concerning the electronic properties. The ionization potential, for example, steeply decreases from around 6.5 eV to 5 eV in the $N = 6-30$ size interval, and then stabilizes by decreasing at a much slower pace. In our previous studies on pure zinc clusters [6,7], we analyzed the evolution of metallicity and found that while most clusters with $N \geq 25-30$ already display typically metallic bonding, for smaller clusters metallicity is still poorly developed. In this sense the Zn-Mg nanoalloys behave similarly to the pure zinc clusters. In the regions with a globally steep slope, it is difficult to identify local features in those electronic indicators, so for optimal visualization we have decided to remove their global size evolution by fitting the data to a cubic polynomial function, and subtracting from the data series the resulting fit; then, we have shifted the mean value of the trendless data to zero and divided them by the standard deviation of the data set. The resulting "standardized" data are dimensionless and are shown for $Zn_2Mg$ nanoalloys in Fig. 6, while the raw data are provided in Fig. S2 in the ESI.

Although none of the clusters has the number of electrons required for a spherical shell closing, the results can be shown to accommodate quite well to jellium expectations by focusing on those cluster sizes that bracket an electron shell closing. The EA, to start with, is consistently higher than average just before (and lower than average just after) the main electronic shell closings, which explains all of the sudden drops in that curve. For example, the EA is very high for $Zn_{12}Mg_6$ and low for $Zn_{14}Mg_7$, these two clusters bracketing the shell closing occurring for $N_e = 40$ electrons. Similarly, one would expect substantial IP drops after each electron shell closing, but those drops only occur in clusters with a well developed metallicity, specifically at cluster pairs bracketing the $N_e = 58, 70, 92$ shell closings. Meanwhile, the IP drop at $N_e = 40$ is very weak, and upon crossing the $N_e = 20$ shell closing the IP even increases. This peculiar behavior has been explained in our previous works on pure zinc clusters [6,7]: it is related to the presence of low-coordinated adatoms and to the coexistence of insulating and metallic bonding contributions within a single cluster, features that obviously depart from a jellium picture of delocalized electrons. $Zn_{14}Mg_7$, for example, displays a low-coordinated adatom attached to the corner of a 20-atom twisted pyramid. The pyramid itself is a superatom with 40 valence electrons, but the two additional electrons in $Zn_{14}Mg_7$ form a lone pair on the adatom rather than contributing to the jellium sea of delocalized electrons [69,70]. This way, the cluster avoids the occupation of very unstable delocalized orbitals and keeps a high IP value beyond the electronic shell closing.

The fundamental gap results from the delicate balance between the IP and EA values. The gap should be locally maximum at an

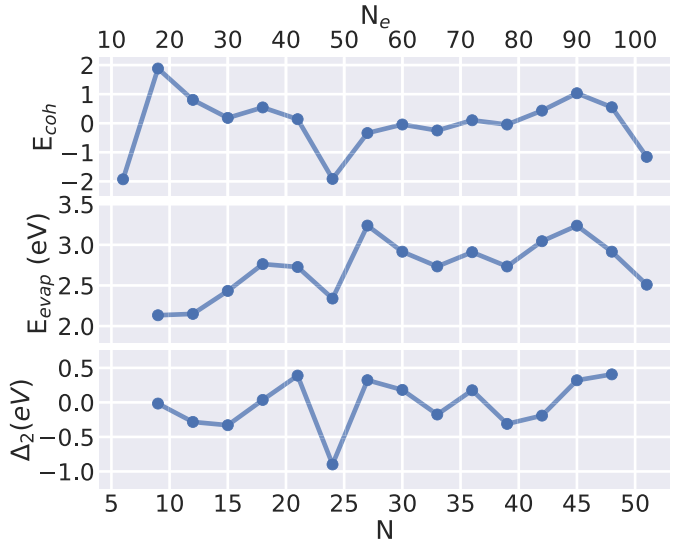


**Fig. 7.** Size dependent stabilities of $Zn_2Mg$ nanoalloys as a function of the total number of atoms $N$ (lower scale) or the total number of electrons $N_e$ (upper scale). The upper graph shows the cohesive energies as standardized dimensionless quantities (see Fig. S3 in the ESI for the absolute cohesive energy values). The middle graph shows the energy cost of evaporating (or dissociating) a $Zn_2Mg$ trimer. Finally, the lower graph displays the second-order energy differences.

electronic shell closing, but its behavior *around* the shell closing is not clear in advance. In the $N \geq 25$ metallic region, the drop in the EA upon crossing a shell closing tends to damp the effect of the corresponding drop in the IP, so the fundamental gap is comparatively featureless in that region. Around $N_e = 58$, for example, the drop in the IP is stronger than the drop in the EA, so a weak drop is observed in the gap; around $N_e = 92$, on the contrary, the drop in the EA is stronger so we observe an increase in the gap. In the small size regime with $N \leq 25$, the electron localization effects on the IP, described in the previous paragraph, explain the steepest increases in the gap, occurring precisely at the $N_e = 20$ and $N_e = 40$ shell closings.

There is a secondary magic number, predicted by ellipsoidal jellium models and observed in many metal clusters, that is exactly realized in the $Zn_2Mg$ series. This is the 15-atom nanoalloy containing 30 electrons, which according to the deformable jellium models should be a marked magic number induced by a prolate shape. In effect, the structure of $Zn_{10}Mg_5$ contains two stacked TTP units resulting in a strong prolate deformation, and it displays a local maximum in both the IP and the gap.

In conclusion, the main variations in the electronic indicators of $Zn_2Mg$ nanoalloys conform to general expectations about systems with delocalized electrons, even if none of the clusters has the right number of electrons to produce a spherical shell closing. Some care has to be taken for a few clusters due to electron localization effects (in lone pair orbitals), as this affects the proper counting of delocalized electrons in the system.

### 3.4. Stabilities of $Zn_2Mg$ nanoalloys

We analyze the size dependence of cluster stabilities for $Zn_2Mg$ nanoalloys through 3 different indicators, shown in Fig. 7: the cohesive energy $E_{coh}$ (defined previously) provides an absolute measure for the cluster global stability, as it quantifies the total internal energy content stored in the chemical bonds; the evaporation energy $E_{evap}$ is here defined as the energy required to dissociate a $Zn_2Mg$ formula unit, so it quantifies the stability with respect to a particular fragmentation channel; finally, the second-order energy difference is defined as $\Delta_2(N) = E_{N-3} + E_{N+3} - 2E_N = E_{evap}(N) - E_{evap}(N+1)$. Both $E_{evap}$ and $\Delta_2$ provide more "local" stability measures in comparing the energy of a cluster of size $N$ with its adjacent neighbors at $N-3$ and $N+3$. In particular, $\Delta_2$ estimates the difference between the evaporation rates of two neighboring cluster sizes in an evaporative ensemble, so this indicator correlates with the cluster populations determined by mass spectrometry on free-standing cluster beams. There are no experimental studies on $Zn_2Mg$ nanoalloys as far as we know, and obviously there is no evidence that evaporation of a $Zn_2Mg$ trimer is the dominant fragmentation channel, yet we have found it useful to consider these indicators to discuss cluster stability at a theoretical level. Our $\Delta_2(N)$ values are expected to correlate with true abundances in an evaporative ensemble of $Zn_2Mg$ nanoalloys assuming that trimer evaporation is the dominant channel. Notice that the three indicators provide related but different information: in particular, the more stable clusters (as determined by the cohesive energy) do not need to coincide with the more abundant sizes (as determined by $\Delta_2$).

As with the electronic properties discussed in the previous section, the analysis of stabilities is slightly complicated by the fact that the $N_e$ scale is quite sparse along the $Zn_2Mg$ series, and also due to the re-entrance of localized molecular orbital states right after the main electronic shell closings. As shown in our previous works on zinc clusters [6,7], occupation of those localized orbitals amounts to a temporary departure from the usual jellium filling pattern, which slightly delays the opening of a new electronic shell. The main effect on stabilities is that clusters with $N_e + 2$ electrons (with $N_e$ a spherical shell closing) tend to be more stable than clusters with $N_e - 2$ electrons. With these caveats in mind, the stabilities in Fig. 7 can be seen to correlate quite well with electronic shell closing effects. The clusters with a higher-than-average cohesive energy are precisely those closer to the spherical shell closings. In effect, the standardized cohesive energy displays local maxima at (or substantial drops after) $N_e = 18, 36, 42, 60, 72$ and 90 electrons, all of them two electrons away from an exact electronic shell closing. The stability drop is specially pronounced after $N_e = 40, 90$, producing deep stability minima for clusters with 48 and 102 electrons.

Some of the clusters with an enhanced stability would also display an enhanced population in an evaporative ensemble where dissociation of a $Zn_2Mg$ trimer is the dominant fragmentation channel. As an example, the cluster with 72 valence electrons persists as a local maximum in the evaporation energy and $\Delta_2$ indicators. For some other clusters, however, a high stability does not correlate with an enhanced abundance. The cluster with 36 electrons, for example, is more stable than the cluster with 42 electrons but significantly less abundant; electron localization on the dangling atom of $Zn_{14}Mg_7$ equips the cluster with a high evaporation energy even if it has two additional electrons on top of an electronic shell closing, producing a marked maximum in $\Delta_2$ for $N_e = 42$ electrons. Analogously, the population of clusters with 90 and 96 electrons is comparable due to the very low stability at $N_e = 102$ electrons. As a final example, the maximum in the cohesive energy for $N_e = 60$ shifts to $N_e = 54$ in the $\Delta_2$ indicator: here, it is the very low stability of the cluster with $N_e = 48$ electrons that results in a very high evaporation energy for $N_e = 54$, so that the cascade of trimer evaporation events would be stalled there maximizing the population of $Zn_{18}Mg_9$. Electronic shell effects are thus the dominant factor determining the stabilities of $Zn_2Mg$ nanoalloys. In the ESI we further analyze possible relationships between stability and structural properties, and demonstrate that the more stable clusters are also more compactly packed.

Although the discussion of $E_{evap}$ and $\Delta_2$ is in itself interesting, we emphasize that it is restricted to one particular fragmentation channel, enforced by the consideration of a single fixed stoichiometry in our work. For example, even if $Zn_{14}Mg_7$ displays a high re-

sistance to dissociate a $Zn_2Mg$ trimer, we expect it to evaporate the Zn dangling atom quite easily, so that cluster would not be specially abundant in an experimental sample at high temperature. An exhaustive consideration of all relevant fragmentation channels would be required for a reliable prediction of particularly abundant sizes in an evaporative ensemble. The cohesive energy results, however, are sound as they do not refer to any particular fragmentation path.

## 4. Conclusions

We have reported a Neural Network potential and its training protocol. The potential is particularly suited to explore the complex high-dimensional potential energy landscape of metallic nanostructures with up to hundreds of atoms, and is capable of reproducing the *ab initio* values of cluster energies and forces within *chemical accuracy*. The automated dynamical update of the data set, afforded with Basin Hopping global optimization searches performed during the training stage, makes our approach ready to use and very efficient at a relatively low computational cost. The potential can be easily implemented into standard optimization or molecular dynamics codes. We have implemented our Neural Network approach into a module of the freely available GMIN code [61] to perform Basin Hopping searches, and applied it for the first time to Zn-Mg nanoalloys ranging from 6 to 52 atoms, focusing on two particular nominal compositions, $MgZn_2$ and $Mg_2Zn_{11}$, which are known to be optimal in corrosion applications as protective coatings. The ZnMg nanoalloys pose a real challenge for any structural seeker, due to previously reported unexpected trends such as the coexistence of metallic and insulating bonds, or the segregation versus mixing tendencies. Our Neural Network potential has been found to outperform, in terms of accuracy and efficiency, the empirical Coulomb-corrected Gupta potential developed in a previous work [44] (which already constituted a step forward over standard potentials) in proposing better candidate structures for DFT reoptimization.

In general, the skeletal atomic structures of the investigated ZnMg nanoalloys are similar to those of pure Zn clusters [7]. As concerns the chemical ordering, the nanoalloys clearly tend to maximize mixing, but with a slight natural segregation of Mg atoms towards the cluster surface, this being the sole factor preventing the nanoalloys to achieve a complete mixing. This shows that Mg atoms will be largely but evenly distributed over the cluster surface, thus playing a major role in determining its reactivity properties. On the other hand, the short-range atomic ordering in the nanoalloys shares similarities with the bulk Laves phases: they both exhibit a preference for maximal mixing, very similar coordination numbers, and even the same coordination polyhedra around Zn and Mg atoms in some cases.

The analysis of the electronic structure and stability suggests that $MgZn_2$ nanoalloys largely conform to a jellium picture of delocalized electrons. The electronic shell closings promote the formation of stronger and shorter bonds, so those magic sizes are expected to be more resistant against penetration of corroding agents. On the other hand, the cluster-specific finding concerning the natural tendency of Mg to segregate to the surface could imply that the precise amount of Mg atoms needed in Zn clusters to improve their anticorrosive properties is actually lower than in the macroscopic regime. Bearing in mind also that the Mg atoms tend to occupy the most coordinated surface sites, even a small concentration of Mg atoms would be mostly if not completely located on the cluster surface, and precisely in those local regions which are expected to be structurally more resistant against distortions due to the strong Zn-Mg bonds. Given the higher reactivity of Mg (as compared to Zn) against external re-agents such as oxygen, the alloying will preferentially guide the corroding agents towards the Mg surface sites, which would later fixate that corroding agent. This way, the favorable effect of Mg alloying for corrosion protection may be to fix the external re-agent precisely at those sites where the surface is more resistant against the penetration of oxygen, thus protecting the cluster core. Avoiding penetration during the early stages of corrosion is crucial, as the addition of further oxygen will eventually produce a passivating oxide crust around the cluster surface.

The results obtained are therefore very promising. We will explore in future works the capabilities of Neural Network potentials in describing more complex interactions such as the attack of corroding molecules on the Zn-Mg system, within our long-term project on corrosion. That way we will be able to analyze the detailed structural and electronic properties of the oxide crust formed on the nanoalloy surface, specifically its resistance against the filtration of corroding ions towards the core. We strongly believe that our current neural network protocol may open new prospects for a variety of problems in Materials Science which are difficult to tackle with conventional potentials; we are also working in generating a friendly version of our neural network code that we hope will be available soon.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

We gratefully acknowledge the support of the Spanish Ministry of Economy and Competitiveness (Project PGC2018-093745-B-I00).

## Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.actamat.2021.117341.

## References

[1] D. Schooss, M.N. Blom, J.H. Parks, B. von Issendorff, H. Haberland, M.M. Kappes, The structures of $Ag_{55}^{+}$ and $Ag_{55}^{-}$: trapped ion electron diffraction and density functional theory, Nano Lett. 5 (10) (2005) 1972–1977.
[2] P. Gruene, D.M. Rayner, B. Redlich, A.F.G. van der Meer, J.T. Lyon, G. Meijer, A. Fielicke, Structures of neutral $Au_7$, $Au_{19}$, and $Au_{20}$ clusters in the gas phase, Science 321 (5889) (2008) 674–676.
[3] K. Majer, M. Lei, C. Hock, B. von Issendorff, A. Aguado, Structural and electronic properties of oxidized sodium clusters: a combined photoelectron and density functional study, J. Chem. Phys. 131 (20) (2009) 204313.
[4] M. Yang, F. Yang, K.A. Jackson, J. Jellinek, Probing the structural evolution of $Cu_N^-$, N= 9–20, through a comparison of computed electron removal energies and experimental photoelectron spectra, J. Chem. Phys. 132 (6) (2010) 064306.
[5] L. Ma, B. von Issendorff, A. Aguado, Photoelectron spectroscopy of cold aluminum cluster anions: comparison with density functional theory results, J. Chem. Phys. 132 (10) (2010) 104303.
[6] A. Aguado, A. Vega, A. Lebon, B. von Issendorff, Insulating or metallic: coexistence of different electronic phases in zinc clusters, Angew. Chem. Int. Ed. 54 (7) (2015) 2111–2115.
[7] A. Aguado, A. Vega, A. Lebon, B. von Issendorff, Are zinc clusters really amorphous? A detailed protocol for locating global minimum structures of clusters, Nanoscale 10 (40) (2018) 19162–19181.
[8] D.M. Deaven, K.-M. Ho, Molecular geometry optimization with a genetic algorithm, Phys. Rev. Lett. 75 (2) (1995) 288.
[9] R.L. Johnston, Evolving better nanoparticles: genetic algorithms for optimising cluster geometries, Dalton Trans. (22) (2003) 4193–4207.
[10] D.J. Wales, J.P.K. Doye, Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms, J. Phys. Chem. A 101 (28) (1997) 5111–5116.
[11] D.J. Wales, H.A. Scheraga, Global optimization of clusters, crystals, and biomolecules, Science 285 (5432) (1999) 1368–1372.
[12] M.T. Oakley, R.L. Johnston, D.J. Wales, Symmetrisation schemes for global optimisation of atomic clusters, Phys. Chem. Chem. Phys. 15 (11) (2013) 3965–3976.
[13] F. Calvo, D. Schebarchov, D.J. Wales, Grand and semigrand canonical basin-hopping, J. Chem. Theory Comput. 12 (2) (2016) 902–909.

[14] R. Ferrando, A. Fortunelli, R.L. Johnston, Searching for the optimum structures of alloy nanoclusters, Phys. Chem. Chem. Phys. 10 (5) (2008) 640–649.
[15] S. Wang, Z. Wu, S. Dai, D.-e. Jiang, Deep learning accelerated determination of hydride locations in metal nanoclusters, Angew. Chem. Int. Ed. 60 (22) (2021) 12289–12292.
[16] C. Zeni, K. Rossi, A. Glielmo, F. Baletto, On machine learning force fields for metallic nanoparticles, Adv. Phys. X 4 (1) (2019) 1654919.
[17] C. Zeni, K. Rossi, A. Glielmo, Á. Fekete, N. Gaston, F. Baletto, A. De Vita, Building machine learning force fields for nanoclusters, J. Chem. Phys. 148 (24) (2018) 241739.
[18] S. Chiriki, S. Jindal, S.S. Bulusu, Neural network potentials for dynamics and thermodynamics of gold nanoparticles, J. Chem. Phys. 146 (8) (2017) 084314.
[19] H. Zhai, A.N. Alexandrova, Ensemble-average representation of Pt clusters in conditions of catalysis accessed through GPU accelerated deep neural network fitting global optimization, J. Chem. Theory Comput. 12 (12) (2016) 6213–6226.
[20] R. Ouyang, Y. Xie, D.-e. Jiang, Global minimization of gold clusters by combining neural network potentials and the basin-hopping method, Nanoscale 7 (36) (2015) 14817–14821.
[21] T. Morawietz, J. Behler, A density-functional theory-based neural network potential for water clusters including van der Waals corrections, J. Phys. Chem. A 117 (32) (2013) 7356–7366.
[22] W.S. McCulloch, W. Pitts, A logical calculus of the ideas immanent in nervous activity, Bull. Math. Biophys. 5 (4) (1943) 115–133.
[23] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proc. IEEE 86 (11) (1998) 2278–2324.
[24] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.
[25] J. Schmidhuber, Deep learning in neural networks: an overview, Neural Netw. 61 (2015) 85–117.
[26] D.F. Specht, A general regression neural network, IEEE Trans. Neural Net. 2 (6) (1991) 568–576.
[27] C.M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 1995.
[28] G. Montavon, G. Orr, K.-R. Müller, Neural Networks: Tricks of the Trade, vol. 7700, springer, 2012.
[29] K. Hornik, Approximation capabilities of multilayer feedforward networks, Neural Netw. 4 (2) (1991) 251–257.
[30] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Netw. 2 (5) (1989) 359–366.
[31] S. Hajinazar, J. Shao, A.N. Kolmogorov, Stratified construction of neural network based interatomic models for multicomponent materials, Phys. Rev. B 95 (1) (2017) 014114.
[32] K. Lee, D. Yoo, W. Jeong, S. Han, SIMPLE-NN: an efficient package for training and executing neural-network interatomic potentials, Comp. Phys. Comm. 242 (2019) 95–103.
[33] N. Artrith, A. Urban, An implementation of artificial neural-network potentials for atomistic materials simulations: performance for $TiO_2$, Comput. Mater. Sci. 114 (2016) 135–150.
[34] T. Xie, J.C. Grossman, Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties, Phys. Rev. Lett. 120 (14) (2018) 145301.
[35] N. Artrith, T. Morawietz, J. Behler, High-dimensional neural-network potentials for multicomponent systems: applications to zinc oxide, Phys. Rev. B 83 (15) (2011) 153101.
[36] J.S. Smith, O. Isayev, A.E. Roitberg, ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost, Chem. Sci. 8 (4) (2017) 3192–3203.
[37] K.T. Schütt, H.E. Sauceda, P.-J. Kindermans, A. Tkatchenko, K.-R. Müller, SchNet–a deep learning architecture for molecules and materials, J. Chem. Phys. 148 (24) (2018) 241722.
[38] S. Hajinazar, E.D. Sandoval, A.J. Cullo, A.N. Kolmogorov, Multitribe evolutionary search for stable Cu–Pd–Ag nanoparticles using neural network models, Phys. Chem. Chem. Phys. 21 (17) (2019) 8729–8742.
[39] A. Thorn, J. Rojas-Nunez, S. Hajinazar, S.E. Baltazar, A.N. Kolmogorov, Toward ab initio ground states of gold clusters via neural network modeling, J. Phys. Chem. C 123 (50) (2019) 30088–30098.
[40] N.C. Hosking, M.A. Ström, P.H. Shipway, C.D. Rudd, Corrosion resistance of zinc–magnesium coated steel, Corros. Sci. 49 (9) (2007) 3669–3695.
[41] T. Prosek, A. Nazarov, U. Bexell, D. Thierry, J. Serak, Corrosion mechanism of model zinc–magnesium alloys in atmospheric conditions, Corros. Sci. 50 (8) (2008) 2216–2231.
[42] E. Diler, S. Rioual, B. Lescop, D. Thierry, B. Rouvellou, Chemistry of corrosion products of Zn and MgZn pure phases under atmospheric conditions, Corros. Sci. 65 (2012) 178–186.
[43] A. Lebon, A. Aguado, A. Vega, Nanoscale reactivity of $Zn_xMg_{20-x}$ investigated by structural and electronic indicators, Corros. Sci. 124 (2017) 35–45.
[44] P. Álvarez-Zapatero, A. Vega, A. Aguado, Incorporating charge transfer effects into a metallic empirical potential for accurate structure determination in $(ZnMg)_N$ nanoalloys, Nanoscale 12 (39) (2020) 20432–20448.
[45] J. Behler, M. Parrinello, Generalized neural-network representation of high-dimensional potential-energy surfaces, Phys. Rev. Lett. 98 (14) (2007) 146401.
[46] J. Behler, Neural network potential-energy surfaces in chemistry: a tool for large-scale simulations, Phys. Chem. Chem. Phys. 13 (40) (2011) 17930–17955.
[47] J. Behler, Constructing high-dimensional neural network potentials: a tutorial review, Int. J. Quantum Chem. 115 (16) (2015) 1032–1050.
[48] S. Chmiela, A. Tkatchenko, H.E. Sauceda, I. Poltavsky, K.T. Schütt, K.-R. Müller, Machine learning of accurate energy-conserving molecular force fields, Sci. Adv. 3 (5) (2017) e1603015.
[49] S. Chmiela, H.E. Sauceda, K.-R. Müller, A. Tkatchenko, Towards exact molecular dynamics simulations with machine-learned force fields, Nat. Comm. 9 (1) (2018) 1–10.
[50] H.E. Sauceda, L.E. Gálvez-González, S. Chmiela, L.O. Paz-Borbón, K.-R. Müller, A. Tkatchenko, BIGDML: towards exact machine learning force fields for materials, arXiv (2021).
[51] A.P. Bartók, G. Csányi, Gaussian approximation potentials: a brief tutorial introduction, Int. J. Quantum Chem. 115 (16) (2015) 1051–1057.
[52] P. Ramachandran, B. Zoph, Q.V. Le, Searching for activation functions, arXiv preprint arXiv:1710.05941 (2017).
[53] K. He, X. Zhang, S. Ren, J. Sun, Delving deep into rectifiers: surpassing human-level performance on imagenet classification, in: Proc. IEEE ICCV, 2015, pp. 1026–1034.
[54] D.-A. Clevert, T. Unterthiner, S. Hochreiter, Fast and accurate deep network learning by exponential linear units (ELUs), arXiv preprint arXiv:1511.07289 (2015).
[55] T. Dozat, Incorporating nesterov momentum into adam, in: ICLR Workshop, vol. 1, 2016, pp. 2013–2016.
[56] S. Ruder, An overview of gradient descent optimization algorithms, arXiv preprint arXiv:1609.04747 (2016).
[57] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).
[58] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, Nature 323 (6088) (1986) 533–536.
[59] H. Martens, T. Naes, Multivariate Calibration, John Wiley & Sons, 1992.
[60] M.L. Astion, M.H. Wener, R.G. Thomas, G.G. Hunder, D.A. Bloch, Overtraining in neural networks that interpret clinical data, Clin. Chem. 39 (9) (1993) 1998–2004.
[61] D.J. Wales, T.V. Bogdan, GMIN: a program for finding global minima and calculating thermodynamic properties from basin-sampling, URL http://www-wales.ch.cam.ac.uk/GMIN(2013).
[62] J.M. Soler, E. Artacho, J.D. Gale, A. García, J. Junquera, P. Ordejón, D. Sánchez–Portal, The SIESTA method for ab initio order-N materials simulation, J. Phys. Cond. Matt. 14 (11) (2002) 2745.
[63] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77 (18) (1996) 3865.
[64] D.R. Hamann, M. Schlüter, C. Chiang, Norm-conserving pseudopotentials, Phys. Rev. Lett. 43 (20) (1979) 1494.
[65] L. Kleinman, D.M. Bylander, Efficacious form for model pseudopotentials, Phys. Rev. Lett. 48 (20) (1982) 1425.
[66] S.G. Louie, S. Froyen, M.L. Cohen, Nonlinear ionic pseudopotentials in spin-density-functional calculations, Phys. Rev. B 26 (4) (1982) 1738.
[67] P. Janthon, S. Luo, S.M. Kozlov, F. Vines, J. Limtrakul, D.G. Truhlar, F. Illas, Bulk properties of transition metals: a challenge for the design of universal density functionals, Phys. Chem. Chem. Phys. 10 (9) (2014) 3832–3839.
[68] F. Stein, A. Leineweber, Laves phases: a review of their functional and structural applications and an improved fundamental understanding of stability and properties, J. Mat. Sci. 56 (9) (2021) 5321–5427.
[69] M. Brack, The physics of simple metal clusters: self-consistent jellium model and semiclassical approaches, Rev. Mod. Phys. 65 (3) (1993) 677.
[70] W.A. De Heer, The physics of simple metal clusters: experimental aspects and simple models, Rev. Mod. Phys. 65 (3) (1993) 611.