Semi-supervised generative approach to chemical disorder: application to point-defect formation in uranium-plutonium mixed oxides

Maciej Karcz, Luca Messina, Eiji Kawasaki, Serenah-Miora Rajaonson, Didier Bathellier, Maylise Nastar, Thomas Schuler, Emeric Bourasseau

Semi-supervised generative approach to chemical disorder: application to point-defect formation in uranium-plutonium mixed oxides

Maciej J. Karcz,[1,2,*] Luca Messina,[1,†] Eiji Kawasaki,[2] Serenah Rajaonson,[1] Didier Bathellier,[1] Maylise Nastar,[3] Thomas Schuler,[3] and Emeric Bourasseau[1]

[1] *CEA, DES, IRESNE, DEC, Cadarache, F-13108 Saint-Paul-Lez-Durance, France*
[2] *Université Paris-Saclay, CEA, LIST, F-91120, Palaiseau France*
[3] *Université Paris-Saclay, CEA, Service de Recherche en Corrosion et Comportement des Matériaux, SRMP, F-91191 Gif-sur-Yvette, France*

(Dated: July 24, 2023)

Chemical disorder has a major impact on the characterization of the atomic-scale properties of highly complex chemical compounds, such as the properties of point defects. Due to the vast amount of possible atomic configurations, the study of such properties becomes intractable if treated with direct sampling. In this work, we propose an alternative approach, in which samples are selected based on the local atomic composition around the defect, and the defect formation energy is obtained as a function of this local composition with a reduced computational cost. We apply this approach to $(U, Pu)O_2$ nuclear fuels. The formation-energy distribution is computed using machine-learning generative methods, and used to investigate the impact of chemical disorder and the range of influence of local composition on the defect properties. The predicted distributions are then used to calculate the concentration of thermal defects. This approach allows for the first time for the computation of the latter property with a physically meaningful exploration of the configuration space, and opens the way to a more efficient determination of physico-chemical properties in other chemically-disordered compounds such as high-entropy alloys.

I. INTRODUCTION

Multi-component solid solutions have raised much interest due to the great versatility of their applications and the possibilities of optimizing their properties. By altering the chemical composition, it is possible to tune the desired properties and increase, e.g., their irradiation resistance [1, 2], or to modify their elastic or electronic [3–5] properties. However, these compounds can be characterized by different degrees of chemical disorder [6]. It is the case for instance of high-entropy alloys (HEA) [7–9] or mixed-actinide oxides used as nuclear fuels [10]. Because of the chemical disorder, determining their physical properties at the atomic scale can be challenging. Since different atomic species occupy the same crystal lattice, they can be distributed on the lattice sites in many different ways. The large number of possible configurations makes it computationally expensive to exhaustively explore the configuration space for computing properties that depend on this distribution. As a result, characterizing these properties is a challenging task.

Point-defect properties are an example of properties that are strongly dependent on the atomic distribution. Two main approaches exist in the literature to address this challenge. The first one consists of using special quasirandom structures (SQS) [11–13]. This allows for the selection of the most disordered structures, which are expected to be the most probable ones in ideally disordered solutions, i.e., solutions with negligible mixing enthalpy. This was the case for instance of previous works aimed at computing the vacancy formation and migration energy in AlHfScTiZr HEAs [13], or the vacancy formation energy in NIFeMnCr HEAs [12]. However, if the solution is not perfectly ideal, the most disordered structures are not necessarily the most probable ones. Furthermore, even in the case of perfectly disordered solutions, SQS provides only few structures that are not necessarily representative of the whole configuration space. This can be sufficient to compute properties at higher temperature, where the energy difference between different atomic configurations is limited, and the system is less sensitive to the specific arrangement of atoms. Conversely, at low temperature, only a few specific low-energy configurations contribute significantly to the physical properties of the system, and they do not necessarily correspond to perfectly disordered SQS supercells. This hinders the versatility of the SQS method.

The second approach is to select configurations via Markov Chain Monte Carlo (MCMC) techniques. For instance, in the work of Takoukam-Takoundjou *et al.* [11], MCMC was applied to sample a large number of configurations and compute the average thermodynamic properties of chemically-disordered $(U, Pu)O_2$. Due to the large size of the configuration space, the ergodicity of the Markov process of the MCMC calculations can be ensured only with the exploration of a very large amount of configurations. Depending on the method used to compute the forces and energies of the system [14], this can require a heavy amount of computational resources. Takoukam-Takoundjou's work was performed with an interatomic potential [15], but it would have been impossible to carry out such an extensive sampling with more accurate *ab initio* methods such as density functional theory (DFT) [16].

* maciej.karcz@cea.fr
† luca.messina@cea.fr

A different approach was attempted by Bathellier *et al.* [17] to study the defect formation energy in $(U, Pu)O_2$ using an interatomic potential [15]. The idea is to carry out a systematic exploration of the configuration space, by limiting its size to a small atomic environment around the defect, and then progressively increasing its size. Since the number of configurations grows exponentially with size, such a method becomes impractical if the range of influence of chemical disorder around the defect is larger than the size of the local environment that can be treated within a reasonable computational cost. Indeed, Bathellier *et al.* performed a systematic exploration up to the second cationic nearest neighbors (nn) around the defect (18 cations), while showing at the same time that the local environment affects the defect formation energy up to the third nn (42 cations). However, systematic exploration up to the 3nn (2^{42} possible configurations) with the same interatomic potential would require approximately 367 billion hours of CPU time on an average supercomputing facility. Because of that, the systematic approach in [17] was limited to the 2nn shell, and the effect of the 3nn shell could not be determined.

The aim of our study is to propose an alternative approach, which allows for a physically significant exploration of the configuration space while reducing the number of required configurations and thus minimizing the computational cost. It consists of two parts. First, we reduce the size of the configuration space by focusing only on the atoms in the closest vicinity of a defect. Many studies conclude in fact that the chemical configuration in the proximity of a vacancy has a strong impact on the formation energy calculations, and it is reasonable to assume that this impact fades out with distance [17–22]. Then, we sample configurations from this reduced space to approximate the distribution of the targeted defect property with a machine-learning (ML) model. This distribution can be used to extract the physical information about the impact and the range of influence of the chemical disorder on the defect property. We show that this approach requires the sampling of far fewer configurations than MCMC, while yielding a set of configurations that is more physically representative than a few SQS structures. In addition, we show that the size of the local environment can be easily extended to the limit where it has no more influence on the target properties, with little added computational cost.

ML methods of generative type [23] are well suited for the prediction of probability distributions. There are several different types of generative methods. For example, generative adversarial networks (GANs) [24] allow for the construction of probability distributions by framing the objective as a supervised learning problem with two competing sub-models. Other examples are variational autoencoders (VAE) [25, 26] or deep belief networks (DBN) [27], both allowing for the generation of complex distributions. A model with an excessive amount of adjustable parameters can often demand a substantial amount of training data. However, using large and complex architectures is not always necessary, as we show in this work. In our approach, we consider the defect property (e.g., the formation energy) to be mainly influenced by the atoms closest to the defect, and treat the impact of further atoms as a random noise that we assume is normally distributed. Therefore, we adopt a semi-supervised approach to model the formation-energy distribution using a Mixture Density Network (MDN) [28], which treats the distribution as a Gaussian random variable [29]. Further details on this approach are provided in Section II D.

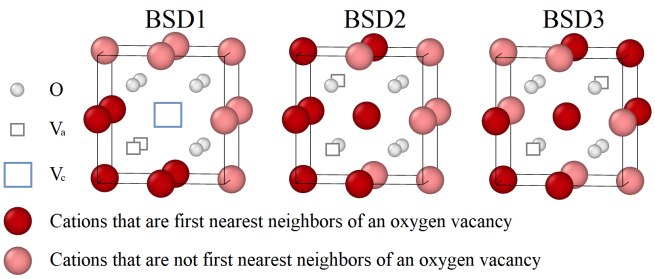

FIG. 1. Representation of three types of bound Schottky defects in the fluorite structure of $(U, Pu)O_2$, with the oxygen or cation vacancies marked as V_a or V_c respectively

We demonstrate the benefits of this approach by studying the formation energy of bound Schottky defects (BSD) properties in $(U, Pu)O_2$, namely the calculation of their formation energy and equilibrium concentration as a function of temperature. A schematic representation of different types of BSD defects in the $(U, Pu)O_2$ fluorite structure is shown in Fig. 1. These defects are a predominant radiation damage feature and a favorable site for fission gas trapping in UO_2, as shown by Positron Annihilation Spectroscopy (PAS) measurements [30, 31] coupled with DFT calculations [32, 33]. This indicates that bound Schottky defects are among the most stable defects in uranium oxides. However, there is very scarce data about the properties of BSD or other point defects in $(U, Pu)O_2$ in the literature [11, 17, 34]. We show that MDN can be effectively applied to such a system in a very robust and cost-efficient way, allowing for the determination of disorder-dependent properties and the influence of larger local environments with relatively little computational effort, compared to the previous study [17]. Moreover, the same approach could be useful from the perspective of studying other disordered systems, for instance for the investigation of defect properties in HEAs.

The paper is organized as follows. The mathematical framework to compute the concentration of thermal defects in disordered compounds is presented at the beginning of Sec. II. In Sec. II C, we show how we compute defect concentration from a formation-energy distribution, and how we predict this distribution with an MDN. Descriptor notation and the MDN framework are introduced in Sec. II D. Databases used for training the model and the architecture of the latter are described in Sec. III. Finally, the results concerning defect concentration and the range of influence of chemical disorder are presented

in Sec. IV and discussed in Sec. V.

II. MATHEMATICAL FRAMEWORK

A. Concentration of thermal defects

The concentration of vacancy-type defects can be computed as the ensamble average of the exponential of the Gibbs formation free energy $G_d(T)$:

$$\begin{aligned} C_d(T) &= \left\langle \exp\left(\frac{-G_d(T)}{k_\mathrm{B}T}\right)\right\rangle \\ &= \left\langle \exp\left(\frac{-(H_d - TS_d)}{k_\mathrm{B}T}\right)\right\rangle, \end{aligned} \tag{1}$$

which can be re-written as:

$$C_d(T) = \left\langle \exp\left(\frac{S_d}{k_\mathrm{B}}\right)\exp\left(\frac{-H_d}{k_\mathrm{B}T}\right)\right\rangle. \tag{2}$$

Here, H_d is the formation enthalpy, S_d the formation entropy, T the temperature, d the defect type, and k_B the Boltzmann constant. In the harmonic approximation, the entropy term in Eq. (2) is independent of temperature, and the vibrational contribution to S_d can be computed from the phonon spectrum of the system. Since S_d gives only a second-order contribution to $C_d(T)$, we will neglect it in this work, and focus on the approximation of enthalpy H_d. However, the approach to chemical disorder presented in this work can be equally applied to the calculation of S_d.

Let us define an atomic configuration $x_c \in X$ as a given atomic supercell, where $X \in \mathbb{R}^{3N}$ is the space of atomic configurations, and c denotes the type of cation that was removed from the supercell to form the Schottky defect. From now on, we impose $H_d = E_d^\mathrm{f}(x_c)$ and refer to the formation enthalpy as formation energy. The average defect concentration can be written as:

$$\begin{aligned} C_d(T) &= \left\langle \exp\left(-\frac{E_d^\mathrm{f}(x_c)}{k_\mathrm{B}T}\right)\right\rangle \\ &= \sum_{i=1}^{N} w(x_c^i)\exp\left(-\frac{E_d^\mathrm{f}(x_c^i)}{k_\mathrm{B}T}\right). \end{aligned} \tag{3}$$

$E_d^\mathrm{f}(x_c)$ is the formation energy of a defect of type d in the configuration x_c, N denotes the number of configurations in the configuration space X, and x_c^i indicates iterating over all configurations in a set of size N. Here $w(x_c)$ is the normalized weight that corresponds to the Gibbs-Boltzmann distribution and depends on the configuration energy $E(x_c)$ as follows:

$$w(x_c) = \frac{\exp\left(-\frac{E(x_c)}{k_\mathrm{B}T}\right)}{Z}, \tag{4}$$

where Z is the partition function, calculated over all of the configurations N in the configuration space X:

$$Z = \sum_{j=1}^{N} \exp\left(-\frac{E\left(x_c^j\right)}{k_\mathrm{B}T}\right). \tag{5}$$

As this work focuses on the study of bound Schottky defects, from this point on 'd' will be replaced with 'BSD' in the following formulas.

B. Formation energy of bound Schottky defects

The BSD formation energy in concentrated disordered compounds can be calculated as [35, 36]:

$$E_\mathrm{BSD}^\mathrm{f}(x_c) = E_\mathrm{BSD}(x_c) - E(x_c) + \mu_c + 2\mu_\mathrm{O} - k_\mathrm{B}T\log(y_c), \tag{6}$$

where c can be either U or Pu, μ is the chemical potential of U, Pu, or O, and y_c is the concentration of species c in the supercell, so $y_\mathrm{U} = 1 - y_\mathrm{Pu}$. $E(x_c)$ is the energy of a supercell without defects, where the cation vacancy has been replaced by a cation of type c. $E_\mathrm{BSD}(x_c)$ is the energy of the same supercell with the cation c and two oxygen atoms removed.

If we consider the case of an ideal solution, where local order can be neglected, the chemical potentials are given by [17, 35]:

$$\begin{aligned} \mu_\mathrm{U}^0 + 2\mu_\mathrm{O}^0 &= \frac{E(\mathrm{UO_2})}{N_\mathrm{C}} + k_\mathrm{B}T\log(y_\mathrm{U}), \\ \mu_\mathrm{Pu}^0 + 2\mu_\mathrm{O}^0 &= \frac{E(\mathrm{PuO_2})}{N_\mathrm{C}} + k_\mathrm{B}T\log(y_\mathrm{Pu}). \end{aligned} \tag{7}$$

$E(\mathrm{UO_2})$ and $E(\mathrm{PuO_2})$ are the energies of the pure $\mathrm{UO_2}$ and $\mathrm{PuO_2}$ supercells, respectively. μ_U^0, μ_Pu^0 and μ_O^0 are the chemical potentials of U, Pu, and O in the perfectly disordered solution. By N_C is denoted the number of cations in the supercell. This formula is valid under the assumption that the effect of defects or short-range order on the chemical potentials is negligible. This is equivalent to treating the system as an ideal solution as has been the case in previous works [11, 37]. The formation energy can be expressed in a simplified form [17]:

$$E_\mathrm{BSD}^{\mathrm{f},0}(x_\mathrm{U}) = E_\mathrm{BSD}(x_\mathrm{U}) - E(x_\mathrm{U}) + \frac{E(\mathrm{UO_2})}{N_\mathrm{C}}, \tag{8}$$

$$E_\mathrm{BSD}^{\mathrm{f},0}(x_\mathrm{Pu}) = E_\mathrm{BSD}(x_\mathrm{Pu}) - E(x_\mathrm{Pu}) + \frac{E(\mathrm{PuO_2})}{N_\mathrm{C}}. \tag{9}$$

In principle, the difference between Eq. (8) and Eq. (9) comes from the interactions between the cation (U or Pu) and its local environment. However, in the case of ideal solid solutions, choosing either one of the above formulas

$E^{\mathrm{f},0}_{\mathrm{BSD}}(x_{\mathrm{U}})$ or $E^{\mathrm{f},0}_{\mathrm{BSD}}(x_{\mathrm{Pu}})$ is equally valid and should give the same defect concentration $C_{\mathrm{BSD}}(T)$, given by Eq. (3).

Since the chemical potentials are written under the assumption that local order can be neglected, the difference between defect concentration computed with Eq. (3) using $E_{\mathrm{BSD}}(x_{\mathrm{U}})$ or $E_{\mathrm{BSD}}(x_{\mathrm{Pu}})$ as a reference can give an indication on the level of local ordering in the system. In Appendix A, we show that this difference is relatively low, which indicates that local order has a limited effect in $(\mathrm{U},\mathrm{Pu})\mathrm{O}_2$ according to the used interatomic potential (the Cooper-Rushton-Grimes potential or CRG [15]). Therefore, in this work, all calculations are performed under the assumption of an ideal solution, using Eq. (8) to calculate the BSD formation energy.

C. Calculation of defect concentration from the formation energy distribution

The concentration of defects can be calculated in a direct way using Eq. 3, i.e., by computing the defect formation energy in all configurations in the configuration space. This is, however, extremely computationally expensive in the case of disordered compounds. In order to reduce this computational cost, we aim to compute the defect concentration by approximating the underlying formation energy distribution.

Equation (3) can be rewritten as:

$$C_{\mathrm{BSD}}(T) = \sum_{i=1}^{N} w(x_c^i) f(x_c^i), \tag{10}$$

where $f(x_c) = \exp\left(-\frac{E^{\mathrm{f}}_{\mathrm{BSD}}(x_c)}{k_B T}\right)$. Under the ideal solid-solution approximation, we can assume that $w(x_c)$ is constant and equal to $\frac{1}{N}$, similarly to what was done in Bathellier's work [17]. In other words, we assume that every configuration has the same probability of occurrence, so we can write:

$$C_{\mathrm{BSD}}(T) = \frac{1}{N}\sum_{i=1}^{N} f(x_c^i). \tag{11}$$

If we denote the probability distribution of configurations by $p(x_c)$, we can approximate the defect concentration in the following way:

$$C_{\mathrm{BSD}}(T) = \mathbf{E}_{x_c\sim p(x_c)}[f(x_c)] \simeq \frac{1}{M}\sum_{i=1}^{M} f(x_c^{(i)}), \tag{12}$$

where every configuration x_c has the same probability $p(x_c)$, and M is the number of samples used for the approximation, so that $M \neq N$ and we need in particular $M \ll N$. However, the number of samples M needed to obtain an accurate estimate of $C_{\mathrm{BSD}}(T)$ is in principle very high. Computing the formation energy $E^{\mathrm{f}}_{\mathrm{BSD}}(x_c)$ in each configuration x_c, for instance by DFT or interatomic potentials, is a very computationally demanding task.

Since the defect concentration depends exclusively on the formation-energy distribution, we assume that it can be computed by sampling the latter one. We therefore re-write Eq. 11 to express $C_{\mathrm{BSD}}(T)$ as a function of the formation-energy distribution. Using the fact that there is exactly one formation energy value for each configuration, we can write:

$$\begin{aligned} C_{\mathrm{BSD}}(T) &= \frac{1}{N}\sum_{i=1}^{N} f(x_c^i) \\ &= \frac{1}{N}\sum_{j=1}^{N'} f_E(E^{\mathrm{f},j}) \sum_{i=1}^{N} \delta(E^{\mathrm{f},j} - E^{\mathrm{f}}_{\mathrm{BSD}}(x_c^i)) \\ &= \sum_{j=1}^{N'} \left(\frac{1}{N}\sum_{i=1}^{N} \delta(E^{\mathrm{f},j} - E^{\mathrm{f}}_{\mathrm{BSD}}(x_c^i))\right) f_E(E^{\mathrm{f},j}) \\ &= \sum_{j=1}^{N'} g(E^{\mathrm{f},j}) f_E(E^{\mathrm{f},j}), \end{aligned} \tag{13}$$

where E^{f} is a value of the function $E^{\mathrm{f}}_{\mathrm{BSD}}(x_c)$ for a given configuration x_c, and $f_E(E^{\mathrm{f}}) = \exp\left(-\frac{E^{\mathrm{f}}}{k_B T}\right)$. While every configuration has one formation energy only, one energy can correspond to multiple configurations. Therefore, $g(E^{\mathrm{f}})$ expresses the density of states that describes the number of occurrences of a given energy E^{f}, and N' is the number of discrete values of energy E^{f}, so $N \geqslant N'$. Since $g(E^{\mathrm{f}}) \geq 0\ \forall j$ and $\sum_j g(E^{\mathrm{f},j}) = 1$, we can write:

$$C_{\mathrm{BSD}}(T) = \mathbf{E}_{E^{\mathrm{f}}\sim p(E^{\mathrm{f}})}[f_E(E^{\mathrm{f}})] \simeq \frac{1}{M}\sum_{i=1}^{M} f_E(E^{\mathrm{f},(i)}), \tag{14}$$

where $p(E^{\mathrm{f}}) = g(E^{\mathrm{f}})$, so:

$$p(E^{\mathrm{f}}) = \frac{1}{N}\sum_{i=1}^{N} \delta(E^{\mathrm{f}} - E^{\mathrm{f}}_{\mathrm{BSD}}(x_c^i)). \tag{15}$$

Instead of sampling the probability distribution of configurations $p(x_c)$, $C_{\mathrm{BSD}}(T)$ can be obtained by sampling the probability distribution of formation energies $p(E^{\mathrm{f}})$. Since the actual $p(E^{\mathrm{f}})$ is unknown, we rely on a ML model to estimate it as accurately as possible. The advantage is that sampling from this estimated distribution of $p(E^{\mathrm{f}})$ does not require additional energy calculations. In other words, the computational cost to obtain $C_{\mathrm{BSD}}(T)$ is shifted from the direct formation-energy calculation of many samples to the training of a ML model. The objective is thus to obtain accurate ML model predictions using minimal data.

D. Semi-supervised estimation of formation-energy probability distribution

The probability distribution $p(E^{\mathrm{f}})$ expresses the distribution of formation energy over all possible configurations. In this work, the latter will be expressed with descriptors used in [17]. In this framework, each configuration is described by its descriptor k, later also called "mode", representing the atomic environment around the BSD defect. Different modes can be viewed in Fig. (2), where each mode (or each k) is represented by the pair of $(N^{\mathrm{c}}_{\mathrm{Pu}}, N^{\mathrm{a}}_{\mathrm{Pu}})$. $N^{\mathrm{c}}_{\mathrm{Pu}} \in \mathbb{N}$ is the number of Pu atoms (from 0 to 12) in the first sphere of influence near the cation vacancy, and $N^{\mathrm{a}}_{\mathrm{Pu}} \in \mathbb{N}$ the number of Pu atoms (from 0 to 6 for BSD2 and BSD3, from 0 to 5 for BSD1) among them that are located near oxygen vacancies. Since $N^{\mathrm{c}}_{\mathrm{Pu}}$ considers all atoms in the first sphere of influence (see Sec. III A), Pu atoms that are close to oxygen vacancies $N^{\mathrm{a}}_{\mathrm{Pu}}$ are also counted in $N^{\mathrm{c}}_{\mathrm{Pu}}$, so that $N^{\mathrm{c}}_{\mathrm{Pu}} \geq N^{\mathrm{a}}_{\mathrm{Pu}}$.

We can rewrite $p(E^{\mathrm{f}})$ as:

$$p(E^{\mathrm{f}}) \approx \sum_i p(E^{\mathrm{f}}|k^i)p(k^i), \tag{16}$$

where the approximation comes from the used local descriptor k. Here $p(E^{\mathrm{f}}|k)$ is the conditional probability of energy E^{f} given descriptor k, and $p(k)$ is its probability. Given the previously undertaken assumption that every configuration has the same probability, as in Eq. (11), we can compute $p(k)$ analytically by counting the total number of configurations included in each k:

$$p(k) = \frac{L_k}{P}, \tag{17}$$

where L_k is the number of configurations that can represented by descriptor k, and P is the total number of configurations corresponding to all possible permutations of U and Pu in the first sphere of influence (1nn):

$$L_k = \binom{N^{\mathrm{c},k}_{\mathrm{Pu}}}{N^{\mathrm{a},k}_{\mathrm{Pu}}} \binom{N^{\mathrm{c},k}_{\mathrm{Pu}}}{N^{\mathrm{a}}_{\mathrm{BSD}} - N^{\mathrm{a},k}_{\mathrm{Pu}}} {y_{\mathrm{Pu}}}^{n_{\mathrm{Pu}}} {y_{\mathrm{U}}}^{n_{\mathrm{U}}}, \tag{18}$$

$$P = \sum_{i=1}^{K_{\mathrm{BSD}}} L_{k^i}. \tag{19}$$

$N^{\mathrm{a}}_{\mathrm{BSD}}$ is the number of possible cationic positions around the anionic vacancies and it depends on the BSD type ($N^{\mathrm{a}}_{\mathrm{BSD1}} = 5; N^{\mathrm{a}}_{\mathrm{BSD2}}, N^{\mathrm{a}}_{\mathrm{BSD3}} = 6$). n_{Pu} (n_{U}) is the total number of Pu (U) atoms in the whole supercell, respectively, K_{BSD} is the number of possible descriptors that depends on the BSD type (BSD1 = 48; BSD2, BSD3 = 49), and k^i indicates iterating over all descriptors in a set of size K_{BSD}.

The estimation of the probability distribution $p(E^{\mathrm{f}})$, defined in Eq. (16), can be approached in a semi-supervised way. Thanks to the undertaken assumption that local order between the atoms can be neglected, the unsupervised part $p(k)$ of Eq. (16) is calculated analytically. However, the likelihood of the supervised part $p(E^{\mathrm{f}}|k)$ is not known and has to be addressed differently. As it is not possible to calculate $p(E^{\mathrm{f}})$ for the whole spectrum of configurations due to the high computational cost, we thus aim at computing the approximation $p_\theta(E^{\mathrm{f}})$ of $p(E^{\mathrm{f}})$:

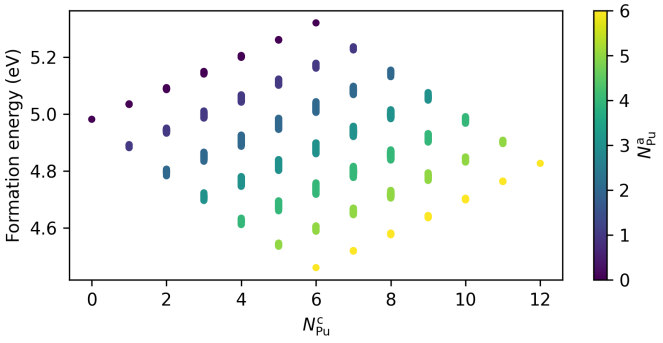

FIG. 2. BSD3 formation energy [17] in $(\mathrm{U},\mathrm{Pu})\mathrm{O}_2$ with 50% Pu content, for all possible atomic configurations in the first sphere of influence around the defect (see Sec. III A). Each point is described by the number of Pu near cation ($N^{\mathrm{c}}_{\mathrm{Pu}}$) and anion ($N^{\mathrm{a}}_{\mathrm{Pu}}$) vacancies.

$$p_\theta(E^{\mathrm{f}}) = \sum_i p_\theta(E^{\mathrm{f}}|k^i)p(k^i). \tag{20}$$

Therefore, the task of estimating $p(E^{\mathrm{f}})$ comes down to the approximation $p_\theta(E^{\mathrm{f}}|k)$ of $p(E^{\mathrm{f}}|k)$ for each descriptor k. As we are expecting the impact of atoms on the formation energy calculations to fade with distance and that the closest atoms have the most important influence, we can decompose E^{f} into two main parts:

$$E^{\mathrm{f}} = E^{\mathrm{f}}(N^{\mathrm{c}}_{\mathrm{Pu}}, N^{\mathrm{a}}_{\mathrm{Pu}}) + \varepsilon = E^{\mathrm{f}}(k) + \varepsilon. \tag{21}$$

By $E^{\mathrm{f}}(k)$ is denoted the impact on the formation energy calculations of E^{f} that comes from the 12 1nn atoms around the defect. Here we assume ε to be a random normally distributed noise coming from the more distant neighbors. In this example, one can consider $E^{\mathrm{f}}(k)$ as a factor that mostly influences the mean, and ε the variance of the calculated energies. This can be observed in Fig. (2), where the mean corresponds to the center of each mode and the variance influences its spread.

Based on the assumption given in Eq. (21), we approximate $p(E^{\mathrm{f}}|k)$ using a Mixture Density Network. As the MDN predicts a distribution that is a Gaussian mixture, then the approximation $p_\theta(E^{\mathrm{f}}|k)$ of $p(E^{\mathrm{f}}|k)$ can be expressed as:

$$p_\theta(E^{\mathrm{f}}|k) = \sum_j \pi_j(k,\theta)\mathcal{N}(E^{\mathrm{f}}|\mu_j(k,\theta), \sigma_j^2(k,\theta)). \tag{22}$$

TABLE I. Summary of the number of configurations in the 1-4nn formation energy databases of different types of BSD defects. The different amount of produced configurations for BSD1 (1440) and BSD2 or BSD3 (1470) comes from the fact that the number of modes is different between BSD1 (48) and BSD2/BSD3 (49). This is the consequence of different possible placement of oxygen atoms around a BSD1, which can have a maximum of 5 neighboring Pu atoms, with respect to the 6 possible placements around a BSD2 or a BSD3 (see Fig. 1).

Database	1nn	2nn	3nn	4nn
BSD1	4096	1440	1440	1440
BSD2	4096	1470	1470	1470
BSD3	4096	22668	1470	1470

Here, j is the number of Gaussian distributions used in the mixture, $\pi_j(k)$ is the mixing coefficient, while $\mu_j(k)$ and $\sigma_j^2(k)$ correspond to the mean and variance of each component in the mixture, respectively. The component density $\mathcal{N}$ can be described as follows:

$$\mathcal{N}(E^{\mathrm{f}}|\mu,\sigma^2) = \frac{\exp\frac{-(E^{\mathrm{f}}-\mu)^2}{2\sigma^2}}{\sqrt{2\pi\sigma^2}}. \tag{23}$$

To measure the difference between $p_\theta(E^{\mathrm{f}})$ and $p(E^{\mathrm{f}})$, a standard approach based on the Kullback-Leibler divergence (KL) [38–40] was applied. A more detailed explanation is included in Appendix B.

III. DATABASES AND MODEL ARCHITECTURE

A. Description of databases

To compute the conditional probability distribution $p_\theta(E^{\mathrm{f}}|k)$ in Eq. (22), several MDN models are trained on specifically prepared databases. The outcome of the MDN can provide relevant physical information, such as the impact and range of influence of chemical disorder on defect properties.

Atomic configurations are built as 2592 atoms supercells (864 cations and 1728 anions) of $(\mathrm{U,Pu})\mathrm{O}_2$ with 50% Pu concentration. This corresponds to a $6\times6\times6$ replication of a primitive cell of a fluorite structure. All atomic configurations are built with the same reference supercell and differ from each other only in their atomic compositions in the closest proximity of the defect. In this way, we are limiting the size of the configuration space only to the possible atomic configurations on the $x$nn sphere, where x varies depending on the sphere radius. The reference supercell contains a random distribution of Pu and U in a perfect fluorite structure without defects. The BSD defect is then placed in the center of the supercell, and the composition of the $x$nn sites around the latter is altered to produce sets of locally-changing configurations. The radius x of the local environment is progressively increased from 1nn to 4nn atoms, containing respectively 12, 18, 42, and 54 atoms in total.

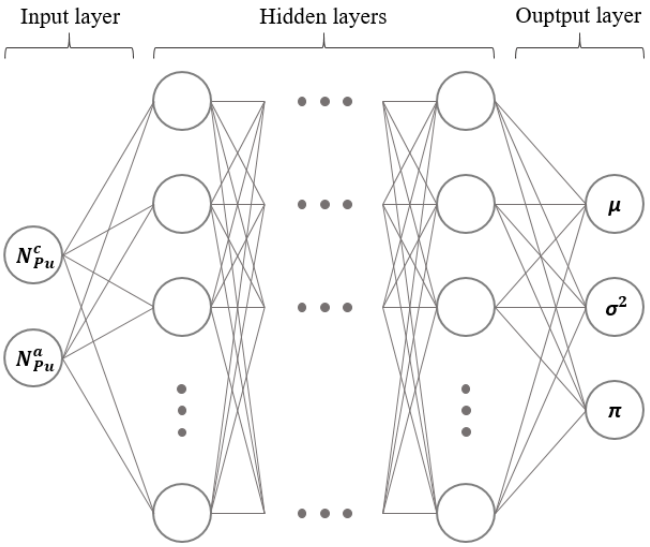

FIG. 3. General architecture of the used MDN models. $N^{\mathrm{c}}_{\mathrm{Pu}}$ is the Pu content in the first sphere of influence around the BSD defect, $N^{\mathrm{a}}_{\mathrm{Pu}}$ is the Pu content near oxygen vacancies, whereas μ, σ^2 and π represent the mean, variance and mixing coefficient of every element of the mixture, respectively.

We build a configuration database for each local-environment radius, as shown in Table I. All the 1nn databases and the 2nn-BSD3 one are taken from [17]. The other databases are built by selecting 30 1nn configurations for each descriptor $k = (N^{\mathrm{c}}_{\mathrm{Pu}}, N^{\mathrm{a}}_{\mathrm{Pu}})$. The U/Pu distribution for the other atoms in the local environment (from 2nn to 4nn) is then determined randomly. In this way, the calculation of the formation energy is based on the main influence coming from 1nn atoms, and the noise induced by 2-4nn atoms, accordingly with Eq. (21).

The BSD formation energies in each configuration are obtained via classical molecular statics (or 0-K energy minimization) with the LAMMPS code [41] and the Cooper-Rushton-Grimes (CRG) empirical potential [15].

B. MDN architecture and training

A general structure of the used MDN model is shown in Fig. 3. The architecture is optimized so to achieve a trade-off between training cost and model accuracy. This leads to an architecture containing 2 inputs in an input layer, 10 neurons in one hidden layer, and 3 outputs in the output layer, corresponding to the mean μ, variance σ^2, and mixing coefficient π of a single Gaussian, i.e., 63 parameters in total. We observed that adding more gaussians to the output distribution ($j > 1$ in Eq. (22)) or more hidden layers for the predictions for a single descriptor k does not result in significant performance im-

provements. Also, smaller models are preferred, as they require fewer data for training. For the hidden layer, a hyperbolic tangent activation function is used.

Each of the training datasets described in the previous section is used to train a different MDN model, resulting in several training sessions for each of the BSD defects. The procedure including database preparation, the training of the MDN model, and the $C_{\mathrm{BSD}}(T)$ estimation described in the next section is summarized in Procedure 1.

Procedure 1: $C_{\mathrm{BSD}}(T)$ estimation with MDN

Data: $\mathcal{D} = \{(k^i, E^{\mathrm{f},i})\}_{i=1}^{L}$
Find the parameters θ^* by training $p_\theta(E^{\mathrm{f}})$ on database $\mathcal{D}$, e.g., as in Appendix B;
for each unique k in $\mathcal{D}$ **do**
  Evaluate $[\mu, \sigma^2, \pi]$ from trained MDN model;
  Calculate $p(k)$ as in Eq. (17);
  $\boldsymbol{V_{\mathrm{mix}}} \leftarrow [\mu, \sigma^2, \pi * p(k)]$;
end
Sample M energies E^{f} from Gaussian mixture $\boldsymbol{V_{\mathrm{mix}}}$;
Approximate $C_{\mathrm{BSD}}(T)$ as in Eq. (14);

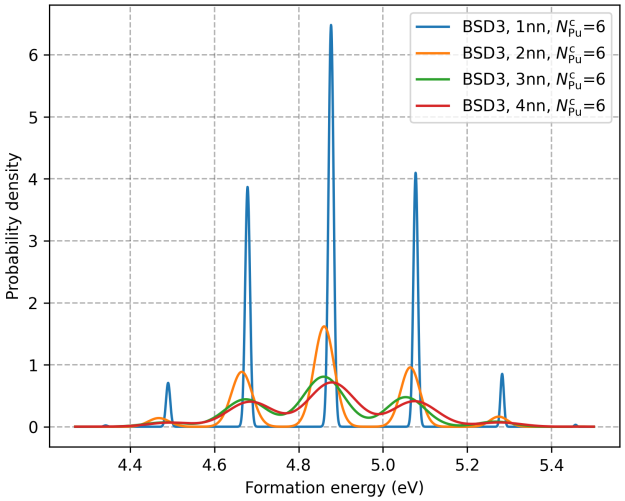

FIG. 4. Probability distribution of the BSD3 formation energy in $(\mathrm{U,Pu})\mathrm{O}_2$ with 50% Pu, for local atomic configurations with $N^c_{\mathrm{Pu}} = 6$ (6 Pu atoms near the cation vacancy), as predicted by the MDN models trained on 980 configurations from 1-4nn databases.

To train the MDN model, the training dataset is first prepared. Configurations are described by their descriptor $k = (N^c_{\mathrm{Pu}}, N^{\mathrm{a}}_{\mathrm{Pu}})$ and the associated BSD formation energy E^{f}. Then, the MDN model is trained to predict a single Gaussian distribution of E^{f} for each descriptor k in a supervised way, as explained in Appendix B. Next, to obtain the final mixture of Gaussian distributions $\boldsymbol{V_{\mathrm{mix}}}$ of formation energies, the MDN model is tasked to evaluate the parameters of Gaussian distributions (μ, σ^2, π) for each unique k in the initial database. Additionally, the probability $p(k)$ of each descriptor k is computed analytically. $\boldsymbol{V_{\mathrm{mix}}}$ is then expressed as a collection of Gaussian distributions, one for each k, where each element is given by:

$$V^k_{\mathrm{mix}} = \pi(k)p(k)\mathcal{N}(\mu(k), \sigma^2(k)). \tag{24}$$

In this work, each element V^k_{mix} consists of only one Gaussian distribution, therefore $\pi = 1$, and the amplitudes of each element of the mixture $\boldsymbol{V_{\mathrm{mix}}}$ are equal to $p(k)$. For example, for BSD3 databases, $\boldsymbol{V_{\mathrm{mix}}}$ would consist of 49 of such vectors, i.e. 49 Gaussian distributions. In our specific case, we found that increasing the number of Gaussian distributions for each descriptor k, such as using 2 or 3 distributions (which would mean 98 or 147 Gaussian distributions in $\boldsymbol{V_{\mathrm{mix}}}$ respectively), did not lead to improved accuracy in our predictions.

Based on the calculations of the negative log-likelihood, detailed in Appendix C, we determine that approximately 200 configurations are required as a minimum for MDN training. As we increase the number of configurations in the training, the rate of increase of the MDN accuracy gradually diminishes, until the improvement becomes negligible above $\approx$ 1000 configurations. Even though the predictions of the MDN models are more accurate when trained on larger datasets, this does not necessarily entail much improvement in the $C_{\mathrm{BSD}}(T)$ computations. As we show in the next section, roughly 200 configurations can be sufficient for calculations of ensemble averages, as is the case of the $C_{\mathrm{BSD}}(T)$. However, higher precision is necessary for the examination of the range of influence of the local environment on the BSD formation energy. For the results shown in the next section, training datasets are built by drawing 4 or 20 samples at random, from each mode, which corresponds to 192 and 960 samples for BSD1 and 196 and 980 samples for BSD2 and BSD3, respectively.

IV. RESULTS

The MDN is applied to predict the probability distributions of formation energy for the three types of BSD defects, with the models including the 1-4nn local environment. The predictions for the BSD3 formation energy for configurations with $N^c_{\mathrm{Pu}} = 6$ is shown in Fig. 4. For instance, the blue 1nn curve corresponds to the probability distribution of the formation energies predicted by the 1nn model, where only the variance on the first nearest-neighbor shell is taken into account. The sharp peaks correspond to the "modes" that can be observed in Fig. 2 for $N^c_{\mathrm{Pu}} = 6$. The sharpness of the peaks depends on the size of the considered sphere of influence. When the size of the local environment is increased, additional noise from 2-4nn atoms is included, resulting in a decreasing height and wider spread of the visible peaks as in the orange, green, and red curves. The modes can still be visible in the 2-4nn curves, but due to the in-

TABLE II. Concentration of BSD defects $C_{\mathrm{BSD}}(T)$ per cationic site for selected temperatures in $(\mathrm{U,Pu})\mathrm{O}_2$ with 50% Pu, and effective formation energy $E^{\mathrm{f}}_{\mathrm{eff}}$ calculated as the negative slope of $C_{\mathrm{BSD}}(T)$ in logarithmic scale (Eq. (25)). The name of the database indicates: the considered sphere of influence, the calculation mode (namely, the direct calculation (DC) or the MDN approach), the type of BSD defect, and the number of configurations in the (training) database. The values of the MDN predictions are given as a mean ± 2 times standard deviation, calculated from the predictions of 100 MDN models trained on different data subsets.

Database	500K	1000K	1500K	$E^{\mathrm{f}}_{\mathrm{eff}}$[eV]
1nn, DC, BSD3, 4096	2.51×10^{-48}	6.84×10^{-25}	6.25×10^{-17}	4.70
1nn, MDN, BSD3, 980	$(2.50 \pm 0.47) \times 10^{-48}$	$(6.82 \pm 0.25) \times 10^{-25}$	$(6.24 \pm 0.12) \times 10^{-17}$	4.70 ± 0.01
1nn, MDN, BSD3, 196	$(2.52 \pm 0.44) \times 10^{-48}$	$(6.83 \pm 0.32) \times 10^{-25}$	$(6.25 \pm 0.16) \times 10^{-17}$	4.70 ± 0.01
3nn, MDN, BSD3, 196	$(4.28 \pm 1.58) \times 10^{-48}$	$(8.31 \pm 0.87) \times 10^{-25}$	$(7.04 \pm 0.45) \times 10^{-17}$	4.68 ± 0.02
3nn, MDN, BSD2, 196	$(2.88 \pm 0.88) \times 10^{-51}$	$(2.64 \pm 0.28) \times 10^{-26}$	$(7.43 \pm 0.50) \times 10^{-18}$	5.00 ± 0.01
3nn, MDN, BSD1, 192	$(2.06 \pm 0.75) \times 10^{-59}$	$(1.74 \pm 0.20) \times 10^{-30}$	$(1.12 \pm 0.08) \times 10^{-20}$	5.80 ± 0.02

troduced noise, their mean values are slightly shifted in comparison to the 1nn distribution.

Including larger local environment have a decreasingly lower impact on the formation energy calculation. To obtain an insight on the relative contribution of 1-4nn atoms to the formation energy computations, the Kullback-Leibler (KL) divergence in Eq. (B2) [38] can be used to compare the differences between the distributions. We obtain $\mathrm{KL}(1\mathrm{nn}||2\mathrm{nn}) \approx 0.805$, $\mathrm{KL}(2\mathrm{nn}||3\mathrm{nn}) \approx 0.272$, and $\mathrm{KL}(3\mathrm{nn}||4\mathrm{nn}) \approx 0.048$. We can observe a significant difference between the 1nn and 2nn distributions, and a lower but still relevant between the 2nn and 3nn ones. However, the difference is considerably smaller between the 3nn and 4nn distributions. Hence we can conclude that the boundary beyond which further atoms start to have negligible influence on $E^{\mathrm{f}}_{\mathrm{BSD}}$ is the 3nn cationic sphere of influence, corresponding to a radius of 7 Å around the defect in the $(\mathrm{U,Pu})\mathrm{O}_2$ fluorite structure. This is consistent with what was found in Bathellier's work with a different approach [17]. The same conclusions apply to the other types of BSD defects.

Next, the obtained probability distribution is used to compute the BSD concentration with a Monte Carlo sampling as in Eq. (14), for a range of temperatures between 500K and 2000K, as shown in Fig. 5. For this computations specifically, we use the 1nn database of Bathellier *et al.* [17] of BSD3 formation energy, consisting of 4096 configurations. This is because the available 1nn database is complete (consisting of all possible configurations of U and Pu atoms in the first sphere of influence), so it is possible to compare defect concentrations calculated from the MDN predictions with the direct calculation (DC) of $C_{\mathrm{BSD}}(T)$ with Eq. (11). It is worth reminding, however, that using only 1nn atoms is a weak approximation of $C_{\mathrm{BSD}}(T)$, and is used only for the evaluation of the performance of the 1nn MDN model. The orange and green curves in Fig. 5 represent the Monte-Carlo based calculation of $C_{\mathrm{BSD}}(T)$, done by sampling the probability distribution predicted by the MDN model trained on 196 and 980 random configurations from the 1nn database, respectively. Based on the $C_{\mathrm{BSD}}(T)$ calculations, an effective formation energy $E^{\mathrm{f}}_{\mathrm{eff}}$ can be obtained as the negative slope of each curve:

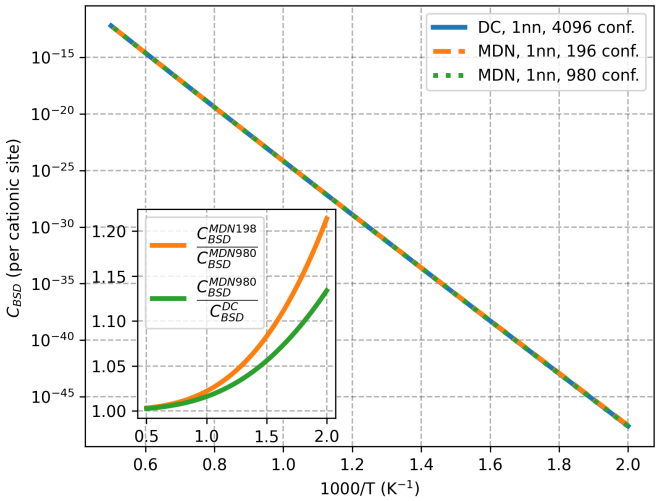

FIG. 5. Calculated BSD3 concentration $C_{\mathrm{BSD3}}(T)$ in $(\mathrm{U,Pu})\mathrm{O}_2$ with 50% Pu. Comparison between the direct calculation, based on the whole 1nn database of 4096 configurations in the work of Bathellier *et al.* [17], and the one obtained by Monte-Carlo sampling of the formation energy distribution predicted by two MDN models, trained on 196 and 980 configurations picked from the 1nn database. The subplot presents the ratio between the results of different models in the whole range of temperatures.

$$C_{\mathrm{BSD}}(T) = \exp\left(-\frac{E^{\mathrm{f}}_{\mathrm{eff}}}{k_{\mathrm{B}}T}\right), \qquad E^{\mathrm{f}}_{\mathrm{eff}} = -k_{\mathrm{B}}T\log\left(C_{\mathrm{BSD}}(T)\right). \tag{25}$$

To estimate the variance of the MDN model parameters and its effect on the predicted defect concentration, 100 models are trained on different training datasets, formed by picking configurations at random from the 1nn database. The resulting $E^{\mathrm{f}}_{\mathrm{eff}}$ and $C_{\mathrm{BSD}}(T)$ calculations for selected temperatures are shown in Table II.

As can be seen in Fig. 5, MDN provides a sufficiently accurate prediction of $C_{\mathrm{BSD}}(T)$ even with only 196 training configurations. Larger training databases improve

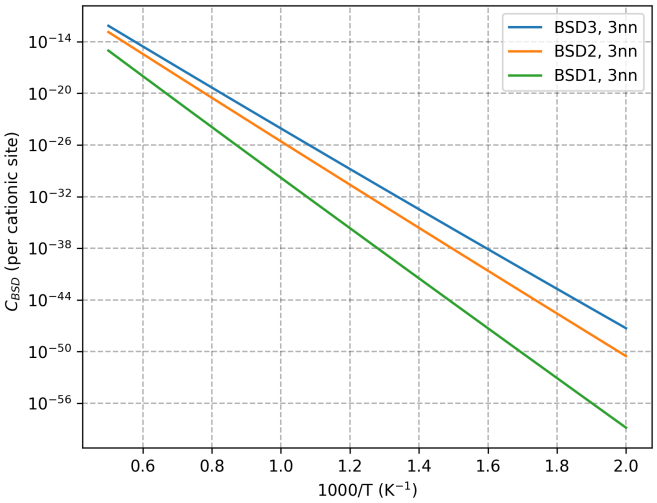

FIG. 6. Comparison between the concentration $C_{\mathrm{BSD}}(T)$ of different BSD defects in $(\mathrm{U,Pu})\mathrm{O}_2$ with 50% Pu, computed by Monte-Carlo sampling of the formation-energy distribution provided by three MDN models, trained on the 3nn databases.

the model accuracy, but concerning the BSD equilibrium concentration and their effective formation energies $E^{\mathrm{f}}_{\mathrm{eff}}$, this results in a small difference of about ≈ 10 meV. It is important to note that 10 meV is well below the accuracy level that one can expect from empirical interatomic potentials such as CRG, and not far from what can be normally achieved with electronic-structure calculations.

Finally, the defect concentration is computed for the three types of BSD defects, as shown in Fig. 6. In this case, each graph is obtained by Monte-Carlo sampling of the energy distribution provided by the MDN model trained on ≈ 200 configurations from the 3nn databases. Similarly as before, 100 models were trained to estimate the variance of the MDN model parameters. The results for selected temperatures are presented in Table II.

The differences between the curves in Fig. 6 come from the difference in BSD stability predicted by the CRG potential. According to the calculations done in this work and in agreement with Bathellier *et al.* [17], the BSD3 is the most stable among the BSD defects with the lowest formation energy. This results in the higher concentration of BSD3 defects that can be observed in Fig. 6 (blue curve). The same conclusions about the stability of the BSD defects were obtained by Balboa *et al.* [42], where the BSD formation energies were computed using the CRG potential by an arithmetic average of the formation energies obtained in seven different configurations. On the contrary, Cheik Njifon [43] found by means of DFT+U calculations that the BSD2 is the most stable BSD defect. However, D. Bathellier showed that this mismatch is due to the small DFT supercell size (96 atoms) used in Cheik Njifon's work and that a match between DFT and CRG is found when using 324-atom DFT supercells. Additionally, it is important to note that, in Cheik Njifon's work, only one SQS supercell was used, so the results are not as reliable as the ones yielded by a proper sampling of the configuration space.

It would be interesting to compare the BSD defect concentrations obtained in this work with the MDN method, with experimental measurements in $(\mathrm{U,Pu})\mathrm{O}_2$. However, to the authors' knowledge, there is at present no such study in the literature. For a comprehensive comparison with such experiments, the formation entropy would have to be included. This could be done by exploring the configuration space with the same MDN-based method. From the perspective of extending our methodology to local-atomic dependent properties of other multicomponent materials such as HEAs, the application is relatively straightforward. The primary modification required is to extend the definition of descriptor k to include the additional chemical species involved. In the case of more complex configuration spaces, it might also be necessary to increase the number of Gaussian distributions predicted for each unique descriptor. This would increase the model capability to express the more complex relationships between constituents.

V. DISCUSSIONS AND CONCLUSIONS

This work showed that the semi-supervised application of the Mixture Density Network (MDN) is a robust approach to evaluate and investigate local-atomic dependent properties in chemically disordered compounds such as mixed-oxides $(\mathrm{U,Pu})\mathrm{O}_2$ nuclear fuels. In our study, based on Cooper-Rushton-Grimes (CRG) potential [15], we showed that by sampling configurations according to their local environment, we can train the MDN model to obtain an accurate prediction of the formation energy distribution with a very limited amount of configurations, thus greatly reducing the computational costs. That allowed us to provide a first-of-a-kind calculation of the equilibrium concentration of Bound Schottky Defect (BSD) defects in $(\mathrm{U,Pu})\mathrm{O}_2$, with a guided sampling of the configuration space that goes well beyond the extent that had been reached so far. We showed in Fig. 5 that it was possible to consider all the relevant local neighbors around the defect (up to 4nn cations) and reach a sufficient accuracy in the estimation of the formation-energy probability density function with a relatively low amount of configurations (about 200). As a comparison, 22668 configurations were required to characterize the BSD3 formation energy in Bathellier's work [17], just to reach the 2nn sphere of influence. Assuming, as a rough estimate, 5 minutes of computational time per energy minimization of a 2592-atom supercell with the CRG potential, that implies 17 hours of CPU time for the MDN approach, as opposed to 1889 hours in the previous study [17]. This makes the MDN approach much more flexible and transferable, allowing one to study, for instance, other types of defects, or repeat the study with other interatomic potentials. Comparing the MDN approach with MCMC methods, the advantage is even more clear:

the Monte-Carlo study of Takoukam-Takoundjou *et al.* [11] on the thermodynamic properties in $(U,Pu)O_2$ required in comparison 50 million configurations.

Furthermore, our approach allowed us to address the issue of the size of the local environment to be considered to obtain an accurate defect formation-energy estimation. Previous studies were limited by the high computational cost required to go beyond the 2nn sphere of influence. Thanks to the chosen simple local descriptor of atomic configurations, we showed that the training cost of MDN models remained the same, while allowing us to obtain the same accuracy independently from the size of the considered sphere of influence. As a result, MDN succeeded in measuring the range of influence of larger local environments on the formation energy distribution. Furthermore, we observed in Fig. 4 that the added influence of N-body interactions between the atoms beyond the 1nn sphere led to an increase of the variance of the formation energy distributions predicted by the MDN models up to the 3nn sphere of influence. Beyond that, the distribution was almost unaffected. This marks the 3nn cationic sphere as a limit beyond which further atoms start to have negligible influence on $E^{\mathrm{f}}_{\mathrm{BSD}}$ calculations. This conclusion confirms quantitatively what was shown in the previous study in a more qualitative way [17]. By the same approach, one could easily include the effect of further spheres, which previously was computationally unreachable.

It is worth reminding that in this work we used the simple local-environment descriptors introduced by Bathellier [17]. The use of more advanced descriptors [44–47] could potentially improve the model performance and decrease the computation cost even further. Furthermore, the model considers the system as an ideal solution, and requires the preliminary calculation of a training database. This database was built based on previous physical knowledge (gathered in Bathellier's work) concerning the interactions between U/Pu atoms and O vacancies. Namely, the existance of the "modes" shown in Fig. 2 was known, and the production of the training database was optimized based on this knowledge. This can be improved with alternative, fully unsupervised approaches. One of the possibilities that will be explored in future work is the use of generative machine-learning techniques capable of generating new configurations and learning the hidden distribution without the requirement for any initial data. Such an approach would also allow us to go beyond the ideal-solution assumption.

The presented results show that the evaluation of local properties influenced by chemical disorder can be efficiently approached through the semi-supervised estimation of their probability distribution functions in the configuration space. The reduced amount of calculations needed to obtain satisfactorily accurate estimations with the MDN models makes it possible to envisage the direct use of DFT to produce the training databases.

This work takes an important step forward in the study of atomic-scale properties of disordered compounds, a field that has been strongly affected so far by computational-time limitations and a lack of efficient sampling strategies. The same methodology can be applied to study other defects and other properties that depend on the local atomic configuration, such as formation entropies, migration energies, attempt frequencies [48], and so on. The presented approach can also be useful to address the properties of high-entropy alloys (HEA) [8], where the same problem of characterization of the vast configuration space can be often encountered.

VI. ACKNOWLEDGMENTS

In this work, high-performance resources from the TGCC (Très Grand Centre de Calcul) of CEA were used for the calculations. This research contributes to the CEA SIMU/PICI2 project. It received funding from the CEA FOCUS program "Expérimentation Numérique et Jumeau Numérique" (EJN). The authors wish to express their gratitude to J. Bouchet and M. Freyss for the fruitful discussions.

Appendix A: Impact of local order

The used formulas for the formation energy computation in Eq. (8) and Eq. (9) rely on the assumption that local order can be neglected. If this is true, then both formulas must yield the same defect concentration $C_{\mathrm{BSD}}(T)$. In order to verify the assumption, we compare the BSD concentration obtained by direct calculation on all data in two 1nn datasets from [17]: namely, the datasets with configurations where a U (Pu) atom was removed from the original supercell to form a BSD3 defect. As it can be seen in Fig. 8, the resulting defect concentrations are close to one another. This means that the effect of local order in this system (as predicted by the CRG potential) is indeed small.

Appendix B: Formation energy probability distribution function

Let us define a dataset $\mathcal{D}$ as a collection of L energies, $L \leq N$, where N is the total number of configurations in a configuration space:

$$\begin{aligned} \mathcal{D} &= \{E^{\mathrm{f},i}\}_{i=1}^{L}, \\ E^{\mathrm{f}} &\sim p(E^{\mathrm{f}}). \end{aligned} \tag{B1}$$

The formation energy of a defect can be described with a probability density function $p(E^{\mathrm{f}})$ as defined in Eq. (15). Since it is not possible to calculate $p(E^{\mathrm{f}})$ for the whole spectrum of configurations, we aim at computing the approximation $p_\theta(E^{\mathrm{f}})$ of the true distribution $p(E^{\mathrm{f}})$, as in Eq. (20). To measure the similarity between the

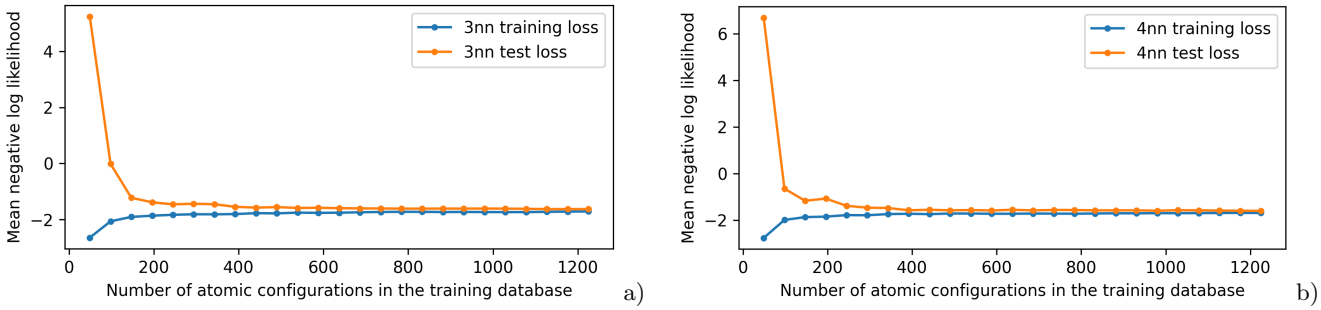

FIG. 7. MDN accuracy computed as the mean negative log-likelihood as a function of the number of configurations included in the training database (BSD3 in $(U,Pu)O_2$ with 50% Pu). Figures a) and b) show the models trained on the 3nn and 4nn databases, respectively.

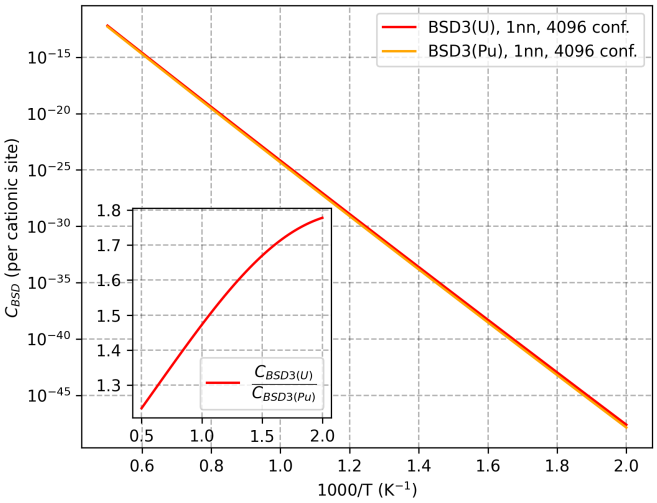

FIG. 8. Verification of the difference between the BSD3 concentration obtained by using either Eq. (8) or Eq. (9). The proximity of the two curves indicates that the effect of local order is small according to the CRG potential. The subplot presents the ratio between both $C_{\mathrm{BSD}}(T)$ calculations for the whole range of temperatures.

two distributions, the Kullback-Leibler divergence can be used:

$$KL(p||p_\theta) = \sum_j p(E^{\mathrm{f},j}) \log\left(\frac{p(E^{\mathrm{f},j})}{p_\theta(E^{\mathrm{f},j})}\right), \tag{B2}$$

over all possible formation energies E^{f}. The objective is to find the set of parameters θ^* that minimizes the KL divergence:

$$\theta^* = \arg_\theta \min[KL(p||p_\theta)]. \tag{B3}$$

We can rewrite the KL divergence as an expected value as follows:

$$\begin{aligned} KL(p||p_\theta) &= \sum_i p(E^{\mathrm{f},i}) \log\left(\frac{p(E^{\mathrm{f},i})}{p_\theta(E^{\mathrm{f},i})}\right) \\ &= \mathbf{E}_{E^{\mathrm{f}}\sim p(E^{\mathrm{f}})}[-\log(p_\theta(E^{\mathrm{f}}))] \\ &- \mathbf{E}_{E^{\mathrm{f}}\sim p(E^{\mathrm{f}})}[-\log(p(E^{\mathrm{f}}))] \\ &= \mathbf{E}_{E^{\mathrm{f}}\sim p(E^{\mathrm{f}})}[-\log(p_\theta(E^{\mathrm{f}}))] - H_p \ , \end{aligned} \tag{B4}$$

where H_p is the Shannon entropy. Since H_p is not dependent on θ, the minimal value of the KL divergence can be written as:

$$\theta^* = \arg_\theta \min[\mathbf{E}_{E^{\mathrm{f}}\sim p(E^{\mathrm{f}})}[-\log(p_\theta(E^{\mathrm{f}}))]]. \tag{B5}$$

Therefore, computing θ^* comes down to computing the minimal negative log-likelihood (using, e.g., the gradient descent techniques):

$$\begin{aligned} \theta^* &= \arg_\theta \min \mathcal{L}(\theta, \mathcal{D}), \\ \mathcal{L}(\theta, \mathcal{D}) &= -\frac{1}{L}\sum_{i=1}^{L} \log p_\theta(E^{\mathrm{f},(i)}), \end{aligned} \tag{B6}$$

We therefore minimize the loss function in Eq. (B6) to find the most optimal set of parameters θ^* that yields an approximation $p_{\theta^*}(E^{\mathrm{f}})$ of the probability distribution. The calculation of $C_{\mathrm{BSD}}(T)$ can then be approximated similarly as in Eq. (14), but by sampling $p_{\theta^*}(E^{\mathrm{f}})$:

$$C_{\mathrm{BSD}}(T) \approx \mathbf{E}_{E^{\mathrm{f}}\sim p_{\theta^*}(E^{\mathrm{f}})}[f_E(E^{\mathrm{f}})] \simeq \frac{1}{M}\sum_{i=1}^{M} f_E(E^{\mathrm{f},(i)}). \tag{B7}$$

Appendix C: Verification of the MDN architecture

In Appendix B, we have shown that by minimizing the loss function in Eq. (B6) we can increase the accuracy

of the MDN model and it depends on the size of the training database. Therefore, we evaluate the MDN performance by computing the mean negative log-likelihood as a function of the number of configurations used for the model training. The results for the BSD3 3nn and 4nn databases in $(U, Pu)O_2$ with 50% Pu are shown in Fig. 7. Similar results are obtained for all BSD defects. It can be seen that the minimal number of configurations needed for MDN training is ≈ 200. A further increase in the size of the training dataset does not significantly improve the accuracy.

[1] O. El-Atwani, N. Li, M. Li, A. Devaraj, J. Baldwin, M. M. Schneider, D. Sobieraj, J. S. Wróbel, D. Nguyen-Manh, S. A. Maloy, *et al.*, Outstanding radiation resistance of tungsten-based high-entropy alloys, Science advances **5**, eaav2002 (2019).

[2] S. Xia, X. Yang, T. Yang, S. Liu, and Y. Zhang, Irradiation resistance in Al x CoCrFeNi high entropy alloys, Jom **67**, 2340 (2015).

[3] J. Kim, H. Kwon, J.-H. Kim, K.-M. Roh, D. Shin, and H. D. Jang, Elastic and electronic properties of partially ordered and disordered Zr (C1- xNx) solid solution compounds: a first principles calculation study, Journal of Alloys and Compounds **619**, 788 (2015).

[4] B. Ghebouli, M. Ghebouli, M. Fatmi, T. Chihi, Z. Heiba, and S. Boucetta, Structural, Elastic, and Electronic Properties of CuClxBr (1-x) Compounds under Pressure, Chinese Journal of Physics **51**, 738 (2013).

[5] K. Ciesielski, L. C. Gomes, G. A. Rome, E. A. Bensen, J. M. Adamczyk, D. Kaczorowski, E. Ertekin, and E. S. Toberer, Structural defects in compounds Zn X Sb (X= Cr, Mn, Fe): Origin of disorder and its relationship with electronic properties, Physical Review Materials **6**, 063602 (2022).

[6] O. Madelung, Disorder, in *Introduction to Solid-State Theory* (Springer Berlin Heidelberg, Berlin, Heidelberg, 1978) pp. 435–472.

[7] A. Lin-Vines, J. Wilson, A. Fraile, L. J. Evitts, M. Rushton, J. Astbury, W. Lee, and S. Middleburgh, Defect behaviour in the MoNbTaVW high entropy alloy (HEA), Results in Materials **15**, 100320 (2022).

[8] E. J. Pickering, A. W. Carruthers, P. J. Barron, S. C. Middleburgh, D. E. Armstrong, and A. S. Gandy, High-Entropy Alloys for Advanced Nuclear Applications, Entropy **23**, 98 (2021).

[9] E. P. George, D. Raabe, and R. O. Ritchie, High-entropy alloys, Nature Reviews Materials **4**, 515 (2019).

[10] M. Beauvy, G. Berthoud, M. Defranceschi, G. Ducros, Y. Guérin, Y. Limoge, C. Madic, G. Santarini, J. Seiler, P. Sollogoub, *et al.*, Nuclear fuels (2009).

[11] C. Takoukam-Takoundjou, E. Bourasseau, and V. Lachet, *Study of thermodynamic properties of $U_{1-y}Pu_yO_2$ MOX fuel using classical molecular Monte Carlo simulations*, Journal of Nuclear Materials 534 , 1 (2020).

[12] C. Li, J. Yin, K. Odbadrakh, B. C. Sales, S. J. Zinkle, G. M. Stocks, and B. D. Wirth, First principle study of magnetism and vacancy energetics in a near equimolar NiFeMnCr high entropy alloy, Journal of Applied Physics **125**, 155103 (2019).

[13] X. Zhang, S. V. Divinski, and B. Grabowski, Ab initio prediction of vacancy energetics in HCP Al-Hf-Sc-Ti-Zr high entropy alloys and the subsystems, Acta Materialia **227**, 117677 (2022).

[14] W. K. T. Cho and Y. Y. Liu, Sampling from complicated and unknown distributions: Monte Carlo and Markov Chain Monte Carlo methods for redistricting, Physica A: Statistical Mechanics and its Applications **506**, 170 (2018).

[15] M. Cooper, M. Rushton, and R. Grimes, A many-body potential approach to modelling the thermomechanical properties of actinide oxides, Journal of Physics: Condensed Matter **26**, 105401 (2014).

[16] N. A. Benedek, A. L.-S. Chua, C. Elsässer, A. P. Sutton, and M. W. Finnis, Interatomic potentials for strontium titanate: An assessment of their transferability and comparison with density functional theory, Physical Review B **78**, 064110 (2008).

[17] D. Bathellier, L. Messina, M. Freyss, M. Bertolus, T. Schuler, M. Nastar, P. Olsson, and E. Bourasseau, Effect of cationic chemical disorder on defect formation energies in uranium-plutonium mixed oxides, Journal of Applied Physics **132**, 175103 (2022).

[18] A. Manzoor, Y. Zhang, and D. S. Aidhy, Factors affecting the vacancy formation energy in Fe70Ni10Cr20 random concentrated alloy, Computational Materials Science **198**, 110669 (2021).

[19] S. Zhao, G. M. Stocks, and Y. Zhang, Defect energetics of concentrated solid-solution alloys from ab initio calculations: Ni 0.5 Co 0.5, Ni 0.5 Fe 0.5, Ni 0.8 Fe 0.2 and Ni 0.8 Cr 0.2, Physical Chemistry Chemical Physics **18**, 24043 (2016).

[20] X. Zhang and M. H. Sluiter, Ab initio prediction of vacancy properties in concentrated alloys: The case of fcc Cu-Ni, Physical Review B **91**, 174107 (2015).

[21] J. Piochaud, T. Klaver, G. Adjanor, P. Olsson, C. Domain, and C. Becquart, First-principles study of point defects in an fcc Fe-10Ni-20Cr model alloy, Physical review B **89**, 024101 (2014).

[22] A. Van der Ven and G. Ceder, Vacancies in ordered and disordered binary alloys treated with the cluster expansion, Physical Review B **71**, 054102 (2005).

[23] G. Harshvardhan, M. K. Gourisaria, M. Pandey, and S. S. Rautaray, A comprehensive survey and analysis of generative models in machine learning, Computer Science Review **38**, 100285 (2020).

[24] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, Generative adversarial nets, Advances in neural information processing systems **27** (2014).

[25] C. Doersch, Tutorial on variational autoencoders, arXiv preprint arXiv:1606.05908 (2016).

[26] J. Baima, A. M. Goryaeva, T. D. Swinburne, J.-B. Maillet, M. Nastar, and M.-C. Marinica, Capabilities and limits of autoencoders for extracting collective variables in atomistic materials science, Physical Chemistry Chemical Physics **24**, 23152 (2022).

[27] G. E. Hinton, Deep belief networks, Scholarpedia **4**, 5947 (2009).

[28] C. M. Bishop, Mixture density networks, Aston University (1994).

[29] S. G. Kwak and J. H. Kim, Central limit theorem: the cornerstone of modern statistics, Korean journal of anesthesiology **70**, 144 (2017).

[30] J. Wiktor, M.-F. Barthe, G. Jomard, M. Torrent, M. Freyss, and M. Bertolus, Coupled experimental and DFT+U investigation of positron lifetimes in UO_2, Physical Review B **90**, 184101 (2014).

[31] M. S. Gautam, Static versus energy-dependent nucleus-nucleus potential for description of sub-barrier fusion dynamics of reactions, Chinese Physics C **39**, 114102 (2015).

[32] E. Vathonne, D. A. Andersson, M. Freyss, R. Perriot, M. W. Cooper, C. R. Stanek, and M. Bertolus, Determination of krypton diffusion coefficients in uranium dioxide using atomic scale calculations, Inorganic Chemistry **56**, 125 (2017).

[33] A. E. Thompson and C. Wolverton, First-principles study of noble gas impurities and defects in UO_2, Physical Review B **84**, 134111 (2011).

[34] I. Cheik-Njifon, M. Bertolus, R. Hayn, and M. Freyss, Electronic structure investigation of the bulk properties of uranium–plutonium mixed oxides $(U, Pu)O_2$, Inorganic chemistry **57**, 10974 (2018).

[35] F. Soisson and M. Nastar, Atomistic simulations of diffusive phase transformations with non-conservative point defects, MRS Communications **12**, 1015 (2022).

[36] K. Li, C.-C. Fu, M. Nastar, F. Soisson, and M. Y. Lavrentiev, Magnetochemical effects on phase stability and vacancy formation in fcc fe-ni alloys, Physical Review B **106**, 024106 (2022).

[37] J.-F. Vigier, P. M. Martin, L. Martel, D. Prieur, A. C. Scheinost, and J. Somers, Structural investigation of $(U_{0.7}Pu_{0.3})O_{2-x}$ mixed oxides, Inorganic Chemistry **54**, 5358 (2015).

[38] J. M. Joyce, Kullback-Leibler divergence, in *International encyclopedia of statistical science* (Springer, 2011) pp. 720–722.

[39] A. Ihalage and Y. Hao, Analogical discovery of disordered perovskite oxides by crystal structure information hidden in unsupervised material fingerprints, npj Computational Materials **7**, 1 (2021).

[40] S. L. Brunton and J. N. Kutz, *Data-driven science and engineering: Machine learning, dynamical systems, and control* (Cambridge University Press, 2022).

[41] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, *et al.*, Lammps-a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales, Computer Physics Communications **271**, 108171 (2022).

[42] H. Balboa, L. Van Brutzel, A. Chartier, and Y. Le Bouar, Damage characterization of $(U, Pu)O_2$ under irradiation by molecular dynamics simulations, Journal of Nuclear Materials **512**, 440 (2018).

[43] I. Cheik Njifon, *Modélisation des modifications structurales, électroniques et thermodynamiques induites par les défauts ponctuels dans les oxydes mixtes à base d'actinides* $(U, Pu)O_2$, Ph.D. thesis, Aix-Marseille (2018).

[44] A. M. G. M.-C. Marinica, Milady - machine learning dynamics, `https://ai-atoms.github.io/milady-docs/` (2022).

[45] A. P. Bartók, R. Kondor, and G. Csányi, On representing chemical environments, Physical Review B **87**, 184115 (2013).

[46] F. Musil, A. Grisafi, A. P. Bartók, C. Ortner, G. Csányi, and M. Ceriotti, Physics-inspired structural representations for molecules and materials, Chemical Reviews **121**, 9759 (2021).

[47] A. M. Goryaeva, J. Dérès, C. Lapointe, P. Grigorev, T. D. Swinburne, J. R. Kermode, L. Ventelon, J. Baima, and M.-C. Marinica, Efficient and transferable machine learning potentials for the simulation of crystal defects in bcc fe and w, Physical Review Materials **5**, 103803 (2021).

[48] H. Mehrer, Diffusion in solids: fundamentals, methods, materials, diffusion-controlled processes (Springer Science & Business Media, 2007) pp. 55–103.