Atomistic kinetic Monte Carlo simulation of precipitation and segregation in metals for nuclear applications, using a novel methodology based on artificial neural networks

ULB

Université Libre de Bruxelles
Faculté des sciences
Physique des Solides Irradiés et
des Nanostructures (PSIN)

Atomistic kinetic Monte Carlo simulation of precipitation and segregation in metals for nuclear applications, using a novel methodology based on artificial neural networks

Thèse présentée par Nicolas Castin
en vue de l'obtention du titre de docteur en sciences.

SCK·CEN
STUDIECENTRUM VOOR KERNENERGIE
CENTRE D'ÉTUDE DE L'ÉNERGIE NUCLÉAIRE

Studiecentrum voor Kernenergie
Centre d'étude de l'énergie nucléaire
Nuclear Materials Science institute (NMS)
Structural Materials group (SMA)

Directeur de la thèse:
Monsieur le professeur Marc Hou

Mentor au SCK•CEN:
Dr. Lorenzo Malerba

Année académique 2010 – 2011

Pour toi maman,

Para Ines, Mateo y Milagros

Jury members

President : Prof. Dr. Daniele Carati
Université Libre de Bruxelles,
Physique Théorique et Mathématique (PTM),
Brussels, Belgium.

Promoter : Prof. Dr. Marc Hou
Université Libre de Bruxelles,
Laboratoire des solides irradiés et Nanostructure (PSIN),
Brussels, Belgium.

Mentor : Dr. Lorenzo Malerba
Studiecentrum voor Kernenergie,
Nuclear Materials Science institute (NMS),
Mol, Belgium.

Prof. Dr. Jean-Paul Ryckaert
Université Libre de Bruxelles,
Laboratoire de Physique de la Matière Molle (LPMM),
Brussels, Belgium.

Prof. Dr. Gianluca Bontempi
Université Libre de Bruxelles,
Centre de BioInformatique et BioModélisation (CINBIOS),
Brussels, Belgium.

Prof. Charlotte Becquart
Université des Sciences et Technologies de Lille (USTL),
Ecole Nationale Supérieure de Chimie de Lille(ENSCL),
Laboratoire de Métallurgie Physique et Génie des Matériaux,
Lille, France.

Dr. Frédéric Soisson
Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA),
Service de Recherches de Métallurgie Physique (DMN-SRMP),
Gif-sur-Yvette, France.

Dr. Christophe Ortiz
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT)
Laboratorio Nacional de Fusión por Confinamiento Magnético,
Unidad de Materiales,
Madrid, Spain.

Table des matières

Acknowledgements

program at that time, and directed me to Lorenzo when I was searching for a master thesis subject and mentor.

Now, I logically come to acknowledge my working group. There is first of all Lorenzo Malerba, my mentor. Could I say anything else that I probably wouldn't be here without him? Lorenzo has been not far from the ideal thesis mentor, for a number of reasons: the project he proposed me for my thesis was unquestionably an original, unique, useful and challenging one; he only rarely told me anything about what I should do, unless I came to him and asked; he could help me to publish my work in appropriate journals, teaching me how to write scientific articles properly; he was always supporting me "above his head", with, the least I can say, only success. I also have to devote special thanks to my direct colleagues, Giovanni Bonny and Dmitry Terentyev, who have unquestionably been masters and guides for me as well. I always thought it was a great luck to work in a cooperative environment, with quality people you can trust, and from whom there's much to learn. I feel today like a real scientist, and this is thanks to them.

A bit more sentimental and private acknowledgements

I feel lucky today to have worked in an international environment, surrounded by people from all corners of the world. This was also a part of the adventure. The atmosphere in the office has always been full of friendship, and very entertaining during our legendary coffee breaks, including so many special persons: él senor Boludaso (la marseillaise à l'oeil de verre), Kevin (Huhu a dada), Taiti Monique, Gunter, Frank Duvel, Amélie (hou la la), Sacha, Boris "Fischer", Dima (so what ?), Papito (self-pumping Fricadel potential guy), 2nd class Canonier Ville, Mr. cleanser Penneman, super Ingenieure Yveke (allons enfants de la paaaatriiiie...), super Luigi-Mario, The Boer-Rémi-Anno "fuel elements" group, when he was here Akira the samurai,... Their contribution to my daily good humour has been invaluable.

Un mot spécial pour Frédéric, mon meilleur ami au bureau, le seul avec qui je peux rire de tout (et de tout le monde aussi mais ça il ne faut pas le dire). Aussi, c'est beaucoup grâce à lui si j'ai des amis en dehors du bureau dans les environs de Mol, ce qui me permet presque d'y vivre une vie normale: MariaVittoria, Monia, Paolo, Monique,...

Tout le monde remercie papa et maman dans un tel document. Je ne suis donc pas en train de révolutionner le monde des acknowledgements. Évidemment, mes parents ont payé pour mes études, et sont fiers de moi, je n'ai pas besoin de l'écrire ici pour vous l'apprendre. Mais ce qui est plus encore important à mes yeux, c'est que mes parents m'ont toujours soutenu. Je peux en dire autant de mon frère Benoit (Bweuhèhèhè), son amie Gina, et mon cousin Simon. Je ne parle pas ici du travail

de thèse en lui-même, mais plutôt de toutes les questions idéologiques associées au fait de travailler pour le secteur nucléaire, surtout dans le contexte actuel. Dans le doute, entouré d'ignorants et de simples d'esprit, ils ont toujours été une lumière qui m'a guidé dans la bonne direction. Mon seul regret dans cette histoire est que mon grand père ne soit plus de ce monde pour assister à ma carrière (il est décédé avant la fin de mes études), car tout ce que je suis aujourd'hui, professionnellement, n'est qu'une dérivée à long terme de tout ce qu'il nous a appris, à mon père d'abord, puis à moi-même indirectement.

Finalmente, quiero escribir unas lineas sobre mis amigos de Argentina. Sin sorpresa la primer persona en quien pienso es mi novia Ines, y sus hijos Mateo y Mili. No es el lugar para decirles cuanto los amo, y ya lo saben de todas maneras. Hay muchas otras personas que quisiera agradecer, por su acogida, y por los buenos momentos que vivimos durante las tres veces que estuve en el país. Pienso a Julián, Roberto, Viviana, Mariana, el hombre Luis, Ana, Juan, Julián, Alejandro,... Finalmente, quiero reservar la última frase especialmente para Mónica, porque es un bien presagio de terminar la introducción con una luz de amistad, de espiritualidad, y de valentía.

Thanks to all of you,
Bedankt allemaal,
Merci à tous,
Gracias à todos,

Nicolas Castin

Abstract

The operation and design of nuclear reactors is continuously concerned with safety. The lifetime of these facilities is limited by changes in the mechanical behaviour of their metallic components (mainly the reactor pressure vessel but also the internals), either accelerated or induced by neutron irradiation. To achieve an accurate quantitative prediction of these changes, depending on the materials compositions and irradiation conditions, is therefore of prime importance in the field of nuclear material science. A formidable experience has been gathered in the last few decades for the case of generation II reactors, thanks to years of safe exploitation of these reactors, and also thanks to the battery of experiments that was performed. This experience, however, cannot directly be enjoyed for the design of new generations of reactors, because the conditions in which materials are expected to be used are sensibly different, and, in fact, more severe by means of temperature, cumulated neutron fluence, and, finally, because of the different chemical environment. In consequence, modelling is nowadays regarded as a vital complement to experimental approaches, for the purpose of achieving a better understanding of the physical and chemical processes that take place in metallic materials under neutron irradiation.

Modelling the effects of neutron irradiation in steels is inherently a multiscale problem. The starting point is the atomic collision cascades, initiated by impinging high energy neutrons, that introduce mobile point-defects in the material. Several different physical models, acting at increasing time- and length-scales, must be developed to properly account for all processes that lead to changes in the macroscopic properties of the material, because of the presence of these mobile point-defects. Furthermore, links are needed between the different physical models involved, because the outputs obtained at a given level are used as inputs for the next models. In this thesis, such a link is performed between the atomic level and coarse-grained models, by developing a novel atomistic kinetic Monte Carlo (AKMC) algorithm, where the material is described as a set of atoms sitting in regular lattice sites. The simulated process is naturally decomposed in elementary thermally activated migration events of point-defects, vacancies or self-interstitial atoms (SIA), that are in competition with each others according to their respective

frequencies of occurrence. The latter are calculated on the basis of the migration energies, that are calculated with little approximations using a non-approximate method taking into account all the effects of long-range static atomic relaxation and chemical interactions. The proposed AKMC algorithm is therefore a physical model, entirely based on a given interatomic potential that is used in the most complete possible way, without the need to define arbitrary parameters that should be, for example, fitted to experimental data. To speed-up the simulation by several orders of magnitude, avoiding the systematic rigorous calculation of migration energies, artificial neural networks (ANN) are trained to predict them, according to the point-defects local atomic environments.

The developed ANN-based AKMC model is successfully applied to the simulation of thermal annealing experiments (where a single vacancy is introduced in the simulation box), for the main purpose of validation by comparison of the obtained results with experimental data. Very satisfying comparison is achieved for two model alloys of interest for nuclear materials science. In both cases, the evolution with annealing time of the average clusters radius and the clusters density are in excellent agreement with experimental data found in literature, differently from results previously obtained by other authors. Next, the algorithm is successfully generalized to allow for the introduction of any number of vacancies, which is one of the two necessary ingredients for the simulation of neutron irradiation damage in metals. This allows for the simulation of long and complex processes, for example the calculation and tabulation of diffusion coefficients and lifetimes of Cu-vacancy clusters, that are necessary inputs for coarse-grained models. Finally, convincing evidence is brought that the ANN-based AKMC algorithm can be generalized to the introduction of SIAs as well, thereby opening the way for simulation of complete irradiation cycles in the future.

Résumé

La sécurité des installations nucléaires est constamment un souci majeur lors de leur exploitation, mais aussi lors de la conception de nouveaux réacteurs. Leurs durées de vie est limitée à cause des changements de comportement mécanique de leurs composants métalliques (principalement la cuve du réacteur mais aussi ses composants internes), qui sont accélérés ou induits par l'irradiation de neutrons. Une prédiction quantitative précise de ces changements, en fonction de la composition des matériaux et des conditions d'irradiation, est par conséquent un objectif de première importance pour la science des matériaux nucléaires. Une vaste expérience a été obtenue concernant les réacteurs de seconde génération, grâce aux nombreuses années d'exploitation de ces réacteurs, mais aussi grâce à la batterie de travaux expérimentaux qui ont été conduits. Cette expérience, cependant, n'est pas directement exploitable pour les réacteurs de nouvelle génération, parce que les conditions dans lesquelles les matériaux seront vraisemblablement utilisés sont sensiblement différentes, et, en réalité, plus sévères en termes de température et dose neutronique cumulée, et également parce que l'environnement chimique est différent. En conséquence, la modélisation est de nos jours considérée comme un complément vital aux approches expérimentales, avec l'objectif d'apporter une meilleure compréhension des processus physiques et chimiques qui se produisent dans les matériaux métalliques sous irradiation de neutrons.

La modélisation des effets de l'irradiation de neutrons dans les aciers est par nature un problème multi-échelle. Le point de départ est la simulation des cascades de collisions atomiques initiées par les neutrons à hautes énergies qui pénètrent dans le matériau, créant ainsi des défauts ponctuels mobiles. Différents modèles physiques, considérant des échelles de temps et de longueur croissantes, doivent être développés afin de convenablement tenir en compte de tous les différents processus qui provoquent des changements de comportement macroscopique, à cause de la présence de ces défauts ponctuels mobiles. En outre, des liens entre les différents modèles doivent être créés, parce que les prédictions de chacun d'entre eux doivent servir de paramètres d'entrée pour les modèles qui travaillent aux échelles supérieures. Dans cette thèse, un tel lien est créé entre le niveau atomique et les modèles à gros-grains, en développant un nouvel algorithme Monte-Carlo

cinétique atomistique (MCCA), où le matériau est décrit comme une collection d'atomes occupant des sites cristallographiques réguliers. Le processus simulé est dès lors naturellement décomposé en séries d'évènements élémentaires activés thermiquement, correspondant à la migration des défauts ponctuels (lacunes ou interstitiels) vers des positions de proches voisins, qui sont en permanence en compétition en fonction de leurs fréquences d'occurrences respectives. Ces dernières sont calculées en fonction des énergies de migrations, qui sont elles-mêmes calculées avec peu d'approximations par une méthode qui prend en compte tous les effets de la relaxation statique et des interactions chimiques à longue portée. Le nouvel algorithme MCCA est par conséquent un modèle physique, entièrement basé sur un potentiel inter-atomique approprié qui est utilisé de la manière la plus complète possible, sans définir de paramètres empiriques qui devraient être, par exemple, fittés depuis des données expérimentales. Finalement, l'algorithme est accéléré de plusieurs ordres de grandeur en utilisant des réseaux de neurones artificiels (RNA), entraînés à prédire les énergies de migrations des défauts ponctuels en fonction de leur environnement atomique local.

Le nouvel algorithme MCCA est utilisé avec succès pour simuler des expériences de recuits (pour lesquels une seule lacune doit être introduite dans la boîte), afin de valider le modèle grâce à une comparaison directe de ses prédictions avec des résultats expérimentaux trouvés dans la littérature. Une comparaison très satisfaisante est accomplie pour deux alliages modèles importants pour la science des matériaux nucléaires. Dans les deux cas, l'évolution avec le temps de recuit du rayon moyen des précipités formés, ainsi que de leur densité, est en très bonne adéquation avec les mesures expérimentales trouvées dans la littérature, contrairement à ce que d'autres auteurs avaient jusqu'à présent réussi. Ensuite, l'algorithme est généralisé avec succès afin de permettre l'introduction d'un grand nombre de lacunes, ce qui est un des deux ingrédients nécessaires pour la simulation des effets de l'irradiation de neutrons dans les métaux. Cet accomplissement permet la simulation de processus longs et complexes, par exemple le calcul de coefficients de diffusions et temps de vies d'amats de cuivre-lacunes, qui sont des paramètres d'entrée nécessaires pour des modèles de simulation à gros-grains. Finalement, des preuves convaincantes sont apportées que l'algorithme MCCA peut être, dans un futur proche, généralisé d'avantage et permettre la prise en compte des interstitiels, ouvrant ainsi la voie vers la simulation de cycles complets d'irradiation.

List of abbreviations

ANN	Artificial Neural Network
AKMC	Atomistic kinetic Monte Carlo
bcc	Body centred cubic
CG	Conjugate Gradients
CPU time	Central Processing Unit time, i.e. computation (wall) time
DBD	Delta-Bar-Delta
DFT	Density Functional Theory
dpa	displacement per atom
GIACA	Gradually Improving Accuracy Constructive Algorithm
IAP	Interatomic potential
KW	Kang-Weinberg
LAE	Local Atomic Environment
LM	Levenberg-Marquardt
MD	Molecular Dynamics
MMC	Metropolis Monte Carlo
NEB	Nudged Elastic Band
OKMC	Object Kinetic Monte Carlo
RPROP	Resilient Propagation
SIA	Self-Interstitial Atom
TAE	Thermal annealing experiment
WSC	Wigner-Seitz cell

1 Motivation and framework of the thesis

Electricity is a vector for the transportation and distribution of energy that presents a number of practical advantages. It allows for the production of large quantities of energy in centralized areas, with the highest possible efficiency, and transport and dispatch within a large community of users, finding the correct balance between efficiency (achieved with high voltage power lines) and safety (transformed to low voltage that is more appropriate for distribution in particular houses for example). Once received, electricity can be transformed in many different forms of useful effects, ranging for example from a mixer (rotation of an axis), air dryer (production and blow of hot air), television or computer (electronics), transportation (electric car or bus),...The production of large quantities of electricity is unquestionably an appreciable added value in the development of our society, not only from the point of view of the common people, but also from the point of view of industry.

The possible primary resources for the production of electricity are manifold: fossil fuels, coal, sun, wind, water, and, finally, the controlled chained nuclear reactions from fissile or fissionable materials. Concerning the last one, the question of the safety of the installations (nuclear reactors) and of the possible consequences of severe accidents has always been, and is especially now, a source of debate and concern. However, the production of electricity without gas, petrol or coal has been, and in the future will even more be, a necessary condition to ensure a progressive then definite transition from the era of fossil fuels, also at shorter term to reduce the emissions of green-house-gases (201). As a matter of fact, nuclear energy has been recommended by groups of experts as a strategic decision when advising governments. Concerning Belgium, see for example (193; 195; 199). At the moment, nuclear energy is included as part of the Strategic Energy Technology plan (SET-plan) of the European Union towards cost effective low-carbon energy production (206).

On the side of possible achievements in the development of profitable fusion technologies, the increase of production of nuclear energy from the reaction of fission, or at least the maintenance of the current levels, can be achieved by two different means:

- The lifetime extension of existing nuclear installations, mainly second generation light-water-reactors (LWR), i.e. either pressurized-water-reactors or boiling-water-reactors, usually designed for a 40-years lifetime. The necessary condition to achieve this is the demonstration to safety authorities that an extension of years of services can be envisaged while keeping the same level of safety. This mainly concerns the pressurized vessel that must be able to withstand abnormal changes of temperature and pressure in either service or accidental conditions. Lifetime limiting factors are mostly hardening and embrittlement of the vessel, but also stress corrosion cracking of in-core components.

- The construction of new reactors, of new generation: III or III+ in the immediate future, IV in a more distant future. In these new generation reactors: (i) Safety will be improved, for example accounting for the core melt-down accident (this is the original feature of the EPR design (196)), or via a systematic use of passive safety systems. (ii) The proportion of use of fissile or fissionable materials as compared to second generation reactors by higher burnup in GenIII and III+ reactors, and by building fast breeder reactors in GenIV. (iii) The warm coolant produced by the core, either liquid or gaseous, must be produced at higher temperature and higher pressure (if gaseous), for reasons of thermodynamical efficiency of production of electricity. As an order of magnitude, in LWR reactors, steam is produced at about 150 bars and 300°C, allowing for an efficiency of hardly 30%.

In either cases, as well as in the case of fusion applications, the main factor limiting the lifetime of the structural materials is their exposure to prolonged fluxes of neutrons. As discussed in the next section, these cause a sometimes severe degradation of the mechanical properties of materials, thereby putting at risk the integrity of the components. It is clear, therefore, that a complete understanding of the effects of neutron irradiation is required for either the design of new components, and/or the prediction of the evolution of the behaviour of given materials under given irradiation conditions and in a given environment. A large experience in this field has quite naturally been gathered in the case of LWRs, after decades of safe exploitation and by conducting a battery of experimental studies. This formidable experience, however, is insufficient if fourth generation reactors are involved, for the simple reason that the operation conditions to which materials are expected to be subjected are sensibly different: (i) the service temperature is expected to be much higher; (ii) the cumulated irradiation dose is expected to be up to three

orders of magnitude larger; (iii) the cooling fluids must be different than water, as a necessary condition against thermalisation of the core neutrons in fast reactors. The construction of sufficiently large databases of experimental data in order to cover these conditions is a formidable time- and money-consuming challenge that cannot be realistically achieved within the desired time-frame for the construction of new generation reactors. For this reason, the development of models, nowadays often based based on computer simulations, of the effects of irradiation in steels is an activity of research that is growing in importance, recognized and recommended by the scientific community as a vital complement to experimental approaches (108; 207).

The present thesis is situated in this framework, and adds a small but valuable contribution to the development of a long-term project of research, which includes aspects of fundamental but also technological research, involving many partners in the world (especially in Europe), that aims at providing a better physical understanding of the effects of neutron irradiation in metals. The rest of this chapter is organized as follows. In section 1.1, the effects of neutron irradiation in steels are briefly described, with the purpose to situate the objectives of the thesis in the global picture. Last, in section 1.2, the multiscale modelling approach used in this thesis is presented, and the exact objectives of the thesis are stated.

1.1 Effects of neutron irradiation in steels

The interested reader can find complete information about the effects of neutron irradiation in steels in books, for example in Ref (69; 118). In this section, some aspects of this complex problem are evoked, to fix the ideas and situate the framework of the thesis.

The origin of all the effects of neutron irradiation in steels is naturally to be found at the atomic level. Impinging high energy neutrons (> 1 MeV), after penetrating the material, enter in collision with atoms that are ejected from their position, thereby initiating atomic collision cascades. These cascades take place like lightnings, lasting very short times (only a few picoseconds). Many thousands of atoms are involved, but most of them actually almost immediately come back to a regular crystal lattice position. Some atoms, however, do not: debris of the cascade are point-defects, constituted by an equal number of vacancies - that correspond to the absence of atoms in volumes where there used to be one atom before irradiation - and self-interstitial atoms (SIA) - that correspond to the inclusion of a second atom in volumes where there used to be only one. The pair constituted by one vacancy and its corresponding SIA is generally called a Frenkel pair. Many of these Frenkel pairs recombine in a very short time (about 100 ps), but the other ones have opportunities to get away from each other, with successions of thermally activated

jumps. SIA or vacancies, alone, migrate in three dimensions, SIAs being sensibly faster because of the lower activation energies (47; 60; 61; 102; 169). When they form clusters, SIAs end-up in parallel orientations and migrate mostly in one dimension (102; 134; 165; 188), though possibly influenced by the chemical environment (102), whereas clusters of vacancies keep a relatively slow and 3-dimensional motion. Both point-defects are capable, in general, of transporting mass, thereby either enhancing or provoking segregation and precipitation of insoluble atoms (10; 127; 150; 154; 181–183).

In fission light-water reactors, the first metallic materials to be irradiated are the clads of the nuclear fuel, that mainly play the role of containment for the fission products, either gaseous or not. Clads in LWRs are made of zirconium alloys. Next, structural materials, whose main objective is to maintain a given geometry of the fuel elements in the core, allowing for an appropriate flow of the coolant, are irradiated typically up to doses over one displacement-per-atom (1 dpa) per year. They are for the most austenitic stainless steels that may undergo, as a combined consequence of irradiation and contact with water, irradiation assisted stress corrosion cracking, radiation induced creep and swelling may become an issue for these materials. Finally, the reactor pressure vessel steel (RPV) itself is a bainitic steel irradiated at typically 275°C to 300°C. The neutron dose accumulated is of limited extent (typically 0.1 dpa after 40 years of normal operation). However, it is proven that these steels undergo as a consequence hardening and embrittlement, phenomena for which the formation of copper-rich precipitates, also enriched in Ni, Mn and Si (7; 31; 85; 136; 137), and sometimes copper-free as well as other nanofeatures, under neutron irradiation, is widely accepted to be the main cause (7; 31; 40; 52; 62; 116; 120; 125; 126; 144; 147), as a consequence of their acting as obstacles to the motion of dislocations. The RPV is the key component to decide the lifetime of the reactor, because it cannot be replaced. As a consequence, much effort is devoted to predicting how hardening and embrittlement evolve with irradiation dose, depending on environmental conditions and composition and microstructure of the steel. Any model for the prediction of RPV steel hardening versus radiation dose, which is the basic requirement for the RPV lifetime assessment, therefore needs to be able to account as correctly as possible for the build-up of, at least, Cu precipitates, ideally taking into account in general also the presence of many other chemical elements.

Working conditions are, as already mentioned, much harder for new generation nuclear reactor materials, for either fission or fusion applications. Not only the neutron fluxes are expected to be higher (and in some cases also the energy of impinging neutron), but the service temperature is expected to be in the range between 500°C and 1000°C, and, finally, the cumulated dose can go as high as 150 or even 200 dpa. In this range of temperature and doses, the most problematic effect that is expected for these materials is irradiation-induced swelling, as well as both thermal and irradiation-induced creep. In addition, microchemical changes in terms

of segregation or appearance of new phases will influence the mechanical properties of these materials. Swelling itself, by being correlated with the formation of large voids and bubbles, has non-hardening embrittlement consequences. Most of these nanostructural changes are believed to be largely determined by the formation and properties (especially mobility) of SIA, while the kinetics of void and precipitate formation will also be influenced by the mobility of vacancies and their clusters. This highlights the importance, for understanding and modelling purposes, of studying and quantifying the behaviour of point-defects and their clusters, in interaction with chemical species, at the atomic level.

1.2 Multi-scale modelling: a fundamental and long-term perspective of research in the field of nuclear materials

It is clear from what has been discussed in the previous section that modelling the effects of neutron irradiation in steels, from the atomic collision cascades to the change of mechanical properties is an inherently multiscale problem: many effects have to be taken into account, involving numbers of atoms that vary by several orders of magnitude, during time-scales that also vary by many orders of magnitudes. These cannot possibly be encompassed with the appropriate level of accuracy by a single physical model. Instead, a chain of models, based on computer simulations but that also include experimental aspects, must be constructed from the sub-atomic to the macroscopic scale, as depicted in Fig. 1.1.

Due to their importance, multi-scale modelling activities are supported by several Euratom projects (197), where the present thesis has been involved (except for the last one):

- FP7-PERFORM60 (205) and its predecessor FP6-PERFECT (108; 204), support multi-scale modelling and irradiation experiments on model alloys for the purpose of lifetime extension of current LWRs and to qualitatively understand and quantitatively predict the damage mechanisms governing the degradation of the reactor pressure vessel and internal structures.

- EFDA (European Fusion Development Agreement) (194) and FP7- GETMAT (56; 200) support multi-scale modelling and irradiation experiments on model alloys for the purpose of the design of structural materials for advanced nuclear applications, namely fusion and Generation IV fission reactors, respectively.

- FP7-FBridge (198) supports multi-scale modelling to develop advanced fuel systems to be used in Generation IV fission reactors.
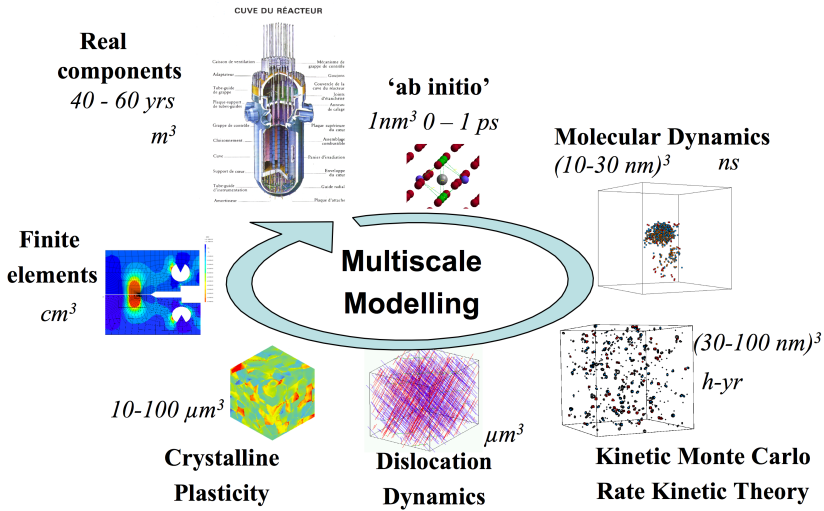
Figure 1.1: Schematic representation of the multi-scale modelling scheme of the effects of neutron irradiation in metals. Models ranging from the subatomic level ("ab initio" on the figure) to the macroscopic scale ("Finite elements" on the figure) are interconnected, receiving as inputs informations from the previous level and providing data used as inputs by the next level.

The development of a fully integrated suite of models encompassing different length and time scales is a long-term approach, that can be regarded as fundamental research that combines many different disciplines of physics and also of computing science. However, the multiscale modelling approach is the only one the can be applied in order to address the problem of irradiation effects based on physical considerations. Intermediate results obtained along the path of building fully integrated platforms of models are often extremely useful in practice to streamline experimental activities and to provide qualitative interpretations to involved experimental results. A complete description of the multiscale modelling scheme is outside the scope of the present document. In this section, we focuss on the scales near the atomic one as depicted in Fig. 1.2, because it is the level of interest for the thesis:

- At the electronic level, matter is modelled as ions and electrons. This is the range of ab initio methods, for example in the Density functional theory (DFT) formulation, where the equations of Schrödinger are solved focusing on the electronic fundamental state.

- At the atomic level, atoms are considered as entities that only have a chemical nature: the latter is defined by electrons and nuclei, but no details about
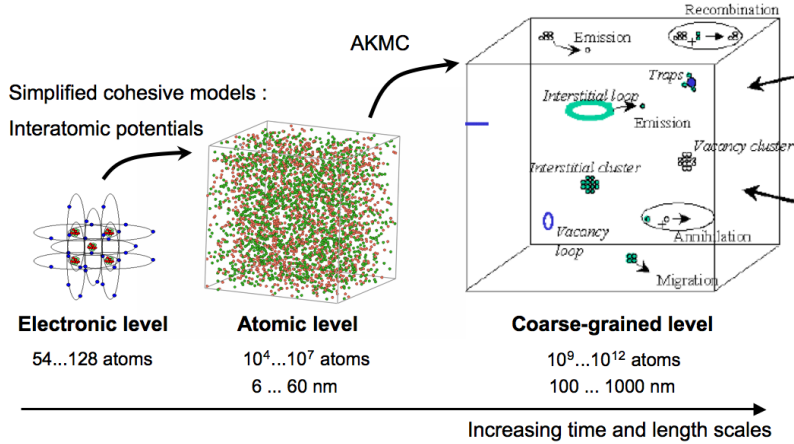
Figure 1.2: Schematic representation of the multi-scale modelling scheme of the effects of neutron irradiation in metals.

these are explicitly included in the models. This is the domain of molecular Dynamics (MD) simulations, where the Newtonian equations of motion are solved for every individual atom of the system, based on the knowledge of the classical interatomic potential that governs the interaction between atoms of same or different species.

- At the coarse-grained level, such as for example in object kinetic Monte Carlo (OKMC) models (Fig. 1.2), the atomic level description is made implicit, and the simulation considers objects (formed by atoms or point-defects) that are mobile and can interact with each other, such as clusters of point-defects and/or solute atoms, dislocations,... Continuum models based on similar levels of abstraction also exist, such as those based on rate equations.

Bridges are necessary between the different levels of approximations represented in Fig. 1.2:

- The atomic level requires a cohesive model in order to describe the system from the energy point of view. This can in principle directly be done with the resolution of the Schrödinger equations, but this is far too complex to be undertaken in practice. For this reason, interatomic potentials are developed instead, as will be discussed further in chapter 2.

- The coarse-grained level requires a large number of parameters as inputs, in order to quantify the motion and interactions of the objects included in the

7

simulation. These numbers, only very rarely derivable from experiments, can be obtained from simulations at the atomic level. However, even remaining at the atomic length-scale, the time-scale encompassed by e.g. molecular dynamics techniques remains too limited to allow these quantities to be properly assessed. Thus, atomic-level models capable of extending the time-frame of MD simulations are needed. This is where and why atomistic kinetic Monte Carlo (AKMC) models are used, as discussed later in chapter 2.

The effect of Neutron irradiation is introduced in most models in terms of collections of Frenkel pairs and point-defect clusters that appear at the correct rate, at either the atomic or the coarse-grained level. For this purpose, displacements cascades induced by impinging neutrons can be modelled by molecular dynamics simulations (8; 9), or also in the framework of the binary collision approximation (152; 159). The debris of these cascades are used as "damage source terms" in other models.

The nanostructural evolution induced by irradiation is a very complex phenomenon that involves interactions among many different actors: point-defects and different chemical species, forming clusters and re-distributing themselves according to often complex and mutually interfering mechanisms. The development of detailed physical models including all these ingredients is therefore an objective that can only be reached by proceeding step by step. A nanostructural and microchemical evolution problem which is somewhat simpler to model is that of thermal ageing. When a material is kept for a long time at high temperature, the equilibrium concentration of vacancies increases and the diffusivity of chemical species is accordingly enhanced. In addition, specific new phases may become stable. This leads to changes in the distribution of chemical species that are to a certain extent similar to those induced by irradiation. As a matter of act, in many instances irradiation merely acts as a way to accelerate thermal diffusion processes, allowing them to occur also at temperatures at which, without supersaturation of point-defects, no microchemical evolution would be detected within reasonable observation times. One important example of a thermal ageing process accelerated by irradiation is Cu precipitation in RPV steels. Another one is the formation of Cr-rich precipitates in high-Cr ferritic/martensitic steels. Because of this close connection between thermal ageing and irradiation effects, thermal ageing experiments are very often performed before irradiation experiments, in order to gather information of thermodynamic and kinetic type on the physical processes of interest. Often, thermal ageing experiments are performed in parallel with irradiation experiments, in order to assess which part of the degradation has to be ascribed to irradiation, and what is merely the effect of prolonged exposure to high temperature. Likewise, it is common and recommended practice that the development of physical models for irradiation processes are developed by addressing first thermal ageing problems. It is indeed unlikely that a model will describe correctly irradiation processes, if it does not describe correctly thermal ageing processes. In addition, since the latter involve

less mechanisms, they represent an ideal intermediate step for the development of models, for which experimental results are available in large quantity for a step-by-step validation. This approach is used in the present thesis as well.

1.3 Objective and organisation of the thesis

In this thesis, a bridge between atomic-level and coarse-grained models is proposed, via the development of a novel and innovative AKMC algorithm. The objective is to model the thermally activated migration of point-defects with the least approximation possible, thereby ultimately allowing for accurate modelling of the effects of neutron irradiation in metals. This is achieved with the use of a virtually approximation-free method for the calculation of the corresponding migration energies, taking into account all the effects of long-range atomic relaxation and long-range chemical interactions. An important requirement, however, is to reduce as much as possible the complexity of the model, in order to allow also the simulation of long and complex processes, where millions or even billions of migration events are involved, within an affordable CPU time. For this reason, the proposed AKMC algorithm relies on the use of artificial neural networks (ANN), that, after being trained on the basis of relevant examples of energy data, can provide fast and on-the-fly predictions of the required migration energies. The obtained algorithm is, eventually, a physical model that requires as only input a relevant interatomic potential, which is thereby exploited in the most complete possible way. The simulated processes are thus decomposed in series of elementary point-defects migration events that are described in a fundamental way, without involving empirical parameters that must for example be fitted to experimental data.

In chapter 2, the theoretical background of AKMC models is summarized, and the different existing approaches are presented to fix the ideas and discuss the objectives that the new algorithm has to aim at. The latter is then completely described at the end of the chapter. Next, chapter 3 gives a brief overview of the artificial neural network technique, for the design of numerical predicting tools used during the thesis. The new ANN-based AKMC algorithm is then applied, in chapter 4, to the simulation of thermal ageing experiments. Comparison with experimentally obtained data (found in the literature) allows full validation of the developed physical model to be achieved. Finally, the AKMC method is, in chapter 5, generalized to allow for the introduction in the simulated volume of (ideally) any number of point-defects, i.e. the basis are set to simulate irradiation processes. In addition, the model is used to estimate diffusion coefficients and lifetimes of complex clusters containing chemical species and point-defects: these data cannot, in any conceivable way, be deduced from experiments, while being of fundamental importance to parametrise coarse-grained models of irradiation processes.

2 Constructing a bridge between the atomic level and coarser-grained models

In the general multi-scale modelling scene, a link between atomic and coarse-grained levels must be established. Models such as object kinetic Monte Carlo (OKMC) or rate equations need a quantity of input parameters in order to properly describe the mobility of various objects (e.g. clusters involving atoms of different chemical species, as well as vacancies or self-interstitials) and their mutual interaction. These parameters are generally inaccessible to experiments. One possible approach is to obtain these input parameters from simulations conducted at the atomic level. Many mechanisms such as migration, dissociation and interactions can, in this way, be qualitatively and quantitatively studied with little approximations, given as input a cohesive energy model.

In this chapter, we describe a widespread model that can be used to provide the link between atomic- and coarse-grained levels, often named atomistic kinetic Monte Carlo (AKMC). The fundamentals of this approach, that will be described later, can be briefly summarized as follows:

- The system being studied is described at the atomic level, and evolves via the migration of point-defects: vacancies and/or self-interstitial atoms (SIAs). Only very few approximations or assumptions are therefore made, in principle, because the processes being studied are naturally decomposed in series of elementary events that are permanently in competition with each other. In the algorithm, events are not chosen following a predefined sequence, but

according to the corresponding migration frequencies, that vary with the local energy landscape.

- The major drawback of this generality lies in the computing time required. As the system is described at the atomic level, one can easily imagine that any of the processes studied will be the result of a huge number of elementary migration events.

The development of promising AKMC models is thus essentially connected to the problem of finding the correct balance between two antagonistic effects. On the one hand, because they directly drive the evolution of the system, the migration frequencies of point-defects should be calculated as rigorously as possible, taking into account all effects of chemical interactions, static relaxation, and also ideally dynamical contributions (such as vibrational entropy). Unfortunately, a rigorous calculation is inevitably synonymous with time-consuming operations, and the global speed of the simulation is in consequence tremendously reduced. On the other hand, oppositely, the migration frequencies should be calculated in the most simple possible way, or at least in the fastest possible way, for reasons of computing time limitations. The objective of this thesis is to develop a new AKMC algorithm that achieves a good compromise between these two antagonistic requirements.

The chapter is organized in the following way. In section 2.1, methods for modelling at the atomic level are briefly described to fix the ideas, and provide a comprehensive explanation of the limitations of these approaches. Next, section 2.2 introduces AKMC algorithms, and tells how these models can create a link between the atomic level and coarser-grained models. The new algorithm developed in the thesis is fully described in section 2.3.

Without loss of generality, the chapter focusses the discussions on the case of bulk materials, with a perfect bcc crystallographic structure, except for the presence of a single vacancy, because AKMC methods have been mainly used in this framework in the past. The application to the simulation of thermal ageing and irradiation damage is described later in chapters 4 and 5.

2.1 Atomic-level modelling

"Certainly no subject is making more progress on so many fronts than biology, and if we were to name the most powerful assumption of all, which leads one on and on in an attempt to understand life, it is that all things are made of atoms, and that everything that living things do can be understood in terms of the jigglings and wigglings of atoms."

Richard Feynmann, Lectures on Physics, vol. 1, p. 3-6 (1963)
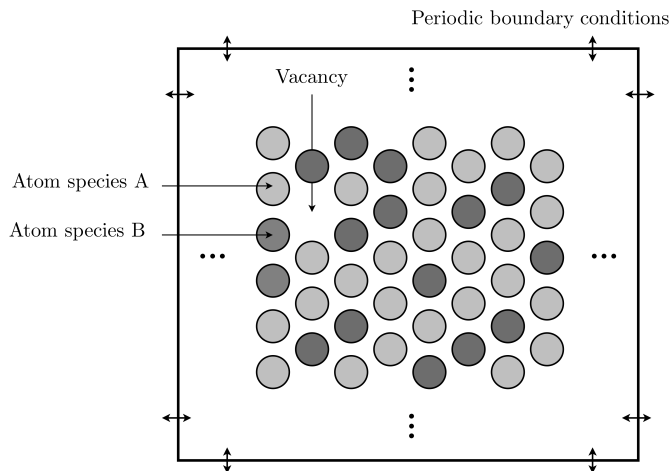
Figure 2.1: Two-dimensional example of an atomic system. Atoms are considered as a spherical and elementary entity, assigned with a given chemical species. In this example, the atoms are arranged in space according to an equilibrium crystallographic structure. A crystallographic site, however, is vacant. Periodic boundary conditions are applied to emulate the infinity of the crystal. Different colours denote different chemical species.

Atomistic simulations describe the material being studied as a set of atoms, located in a three-dimensional and periodic box (or other boundary conditions if necessary), as depicted in Fig. 2.1. In this example, the material is a bcc single-crystal without crystallographic defects, except for the presence of a vacancy. The objective of multiscale modelling is to describe phenomena with macroscopic consequences (for example disappearance of the protective oxide layer and therefore corrosion) in terms of the atomic-level phenomenon that induced this (segregation of chemical species at the surface) and also of the physical mechanism producing it, i.e. the migration of the vacancy, that induces mass transport in the material. For this purpose, the interactions between the atoms of the system, that govern the diffusion of the vacancy, must be modelled.

Ideally, all atomic interactions can be described exactly by solving the equation of Schrödinger for the material, thereby avoiding the use of empirically or arbitrarily chosen parameters, but instead relying directly on the first principles of physics. The resolution of these equations, which in principle include a huge amount of unknowns (three for every electron in the framework of the Born-Oppenheimer approximation), is impossible. They are therefore in practice solved by reducing drastically the amount of variables. A tremendously easier way to do so is the use of the density functional theory (DFT), laid out by Hohenberg and Kohn in 1964
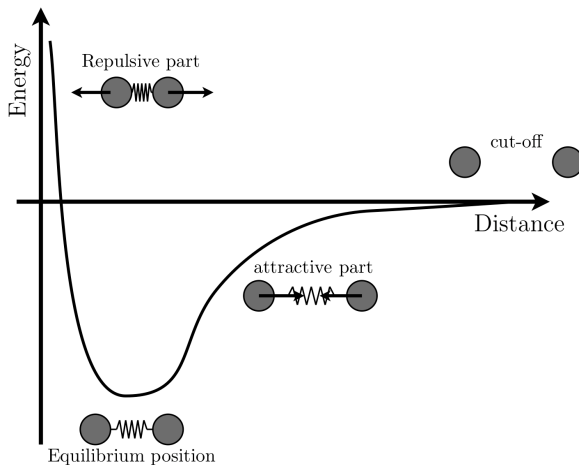
Figure 2.2: Evolution with the distance between two atoms of the energy contribution of their interaction to the total energy of the system. For short distances, the energy is large, and atomic forces are repulsive. Forces are attractive for longer distances than the equilibrium position, but vanish beyond the cut-off range.

(76). To summarize, they demonstrated that an exact solution of the Schrödinger equation can be found searching for an electronic density, that has only three spatial variables, in replacement of the position of all electrons of the system (3N variables). This theory, however, is limited to ground states, and cannot address excited states of matter. If this limitation is not a problem for the physical issue at hand, then DFT enormously reduces the complexity of many first principles calculations, and is consequently very popular nowadays. The interested reader can find more information about DFT in Ref. (166).

Unfortunately, DFT methods, in spite of their relative simplicity compared to the 3N variables problem, remain very complex and heavy to conduct in practice, for reasons of both required memory and, especially, required computing time. Even with supercomputers available nowadays, DFT methods can, in the best case, only handle a few thousands of atoms at the very most. The development of simplified cohesive models to describe atomic interactions, taking a higher level of abstraction where electrons are not included, is thus a necessary work to extend the range of action of atomistic simulations. In this thesis, interatomic potentials (IAP) are used for this purpose. In these potentials, the energy associated with the interaction of atoms varies in magnitude with their respective chemical natures and with their mutual distance only. Fig. 2.2 depicts the classical evolution of that energy function, for an isolated pair of atoms, with the mutual distance. Atomic forces are calculated

as the first order derivative of the energy. Forces are repulsive if the atoms are too close to each other, and oppositely are attractive is their are far from each other. They vanish if the distance is larger than a certain cut-off, denoted henceforth as r_{co}. The energy function depicted in Fig. 2.2 changes quantitatively if chemical species are changed, but remains qualitatively the same.

Several forms of IAPs exist. They can, first of all, be classified in different families, namely central force potentials, local concentration dependent potentials, angle-dependent potentials, and also cluster expansions. The families, also, count several different models. The work in this thesis is based on central force potentials, in particular in the embedded atom method (EAM) formalism. For more information about IAP, the interested reader is referred to G. Bonny's PhD dissertation (27). For the present thesis, it is useful to keep the following points in mind:

- IAPs are not intended to be a perfect substitute to ab initio methods. Results of modelling obtained using an IAP are only the reflection of the reality that stems out of the potential. In that respect, IAP are in any case designed to fulfil certain objectives, for example a good description of the interaction between point-defects, correct formation energies for some defects,... But, and it is in fact unfortunate, IAP can probably not reflect the complete complexity of atomic interactions, and eventually can never lead to a rigorously correct prediction of nature under all conditions. Nevertheless, if well fitted they can be sufficiently accurate for a range of properties and are certainly capable of providing at least correct trends as functions of local concentrations or other variables. It should also be noted that to date, and certainly in the future as well, they represent the only affordable way of simulating systems containing up to tens of millions of atoms, a goal very far to be reachable in DFT. Finally, DFT itself contains approximations that may affect the quantitative reliability of the results.

- IAP can be fitted on the basis of data from various origins: ab initio data, of course, but also experimental data, for example to achieve the reproduction of diffusion coefficients that can be measured experimentally. Also, innovative techniques have been recently developed to design potentials that lead to a correct reproduction of experimental phase diagrams (27; 140).

Given the IAP, the evolution with time of the atomic system depicted in Fig. 2.1 can be modelled with the resolution of Newton's equations for motion. Without temperature, the atoms are timesteply displaced to reach a final equilibrium position that is the closest from the initial state of the system. This operation can be named *static relaxation*, because only the atomic forces stemming out of the potential energy vanish, and the effects of external contributions such as temperature are disregarded. In this thesis, this operation is performed using the conjugate gradients method (141).
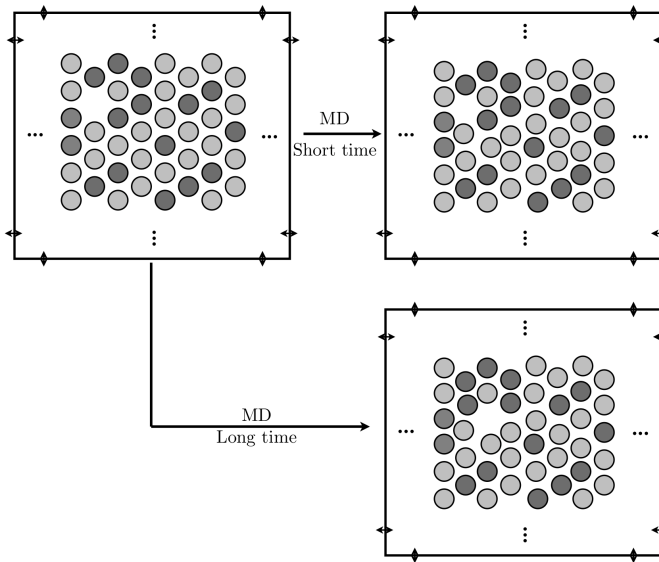
Figure 2.3: Simple example of MD simulation for a monocrystal with one vacancy. After short simulation times, all atoms remain in the vicinity of their static equilibrium position. After longer times, one atom manages to escape, and find another stable position that was vacant. The vacancy has in turn migrated.

Temperature can be introduced in the system by stochastically increasing the kinetic energy of the atoms, with a random direction and a magnitude that respect a Boltzmann distribution. The simulation is then called *Molecular Dynamics* (MD). In this case, the atomic system do not stay in a static equilibrium, because atoms are continuously moving, as depicted in Fig. 2.3. The motion of the atoms, even tough stochastic, does not provoke significant changes in the system, at least at the beginning of the simulation: the free motion is constrained by the increase of the total energy, and forces derived from the potential tend to pull all atoms back to their respective static equilibrium positions. If the simulation is continued for a (much) longer time, some atoms have more and more serious occasions to attempt escaping from their present static equilibrium positions. Those situated close to the vacant site, for instance, have a chance to escape, and find a new equilibrium position somewhere in the vacant site, as depicted in Fig. 2.3. When this happens, the vacancy is displaced to another position, and other atoms have the chance, later in time, to achieve a similar migration. For convenience, the phenomenon here described is called vacancy migration, even if in reality a vacancy is not an object properly speaking, but rather the absence of an object. Vacancies are nevertheless

regarded as objects in crystallography. A succession of simple migration events, such as the vacancy migration depicted in Fig. 2.3, have the effect of transporting atoms, and eventually rearrange the atomic system. The simulation of such processes, driven by thermally activated migration of point-defects, is part of the ingredients necessary to simulate, in the general multiscale modelling scheme, the effects of neutron irradiation in steels, as well as thermal ageing processes where irradiation is not involved. It is clear, unfortunately, that this cannot practically be achieved with MD, for the obvious reason that a huge number of timesteps are necessary. Most of these time steps correspond to stochastic events that are relatively uninteresting to consider: only the atoms are oscillating under the effect of temperature, but the system remains, essentially, identical. For this reason, a more promising approach is to describe the events taking place in the material with a higher level of abstraction, that should, ideally, be a correct approximation of the complete process simulated with MD. The atomistic kinetic Monte Carlo approach (AKMC), described in the next section, is in this respect a possibility.

2.2 Atomistic kinetic Monte Carlo

"I am convinced God does not play dice"
Albert Einstein, in a private letter to Max Born, 4 December 1926.

"Einstein, don't tell God what he should do !"
Niels Bohr

The Atomistic kinetic Monte Carlo (AKMC) model is a widespread tool to study diffusion-controlled microstructural and microchemical evolution in alloys during thermal ageing and under irradiation: see e.g. (13; 14; 21; 23; 41; 91; 97; 158; 162–164; 176–178; 180). It creates a link between the atomic level and coarse-grained models, because the most important features of the atomic-level are retained, while the time and length scales reachable by the model are largely extended by several orders of magnitude thanks to the MC approach.

The theoretical foundations of the method have been known for a long time, having been rigorously laid out by Fichthorn and Weinberg in 1991 (57), who denoted it as dynamical Monte Carlo. The residence time algorithm, which is at the core of the method, was developed even before, in a first average version by Young and Elcock (190) and in a more precise way by Bortz et al. (28).

This section is intended to provide a comprehensive overview of the AKMC method, and how it has been used in the past. It is organized in the following way. In section 2.2.2, the algorithm of the AKMC method is described and discussed. Without loss of generality, discussions focus on the simple case already used as example

in the previous sections, i.e. a bulk material with solute atoms and one single vacancy as mobile point-defect. The reason is that AKMC methods have been used previously most of the time exclusively in this framework. Next, in section 2.2.3, existing methods for the calculation of the vacancy migration frequencies are listed and discussed. The objective of the thesis, by developing a new approach, are thereby introduced. Lastly, section 2.3.3 discusses how AKMC algorithms can be parallelized.

2.2.1 General ideas

As anticipated in section 2.1, the main objective of AKMC is to describe the evolution of an atomic system, driven by the thermally activated migration of point-defects, avoiding the consideration of events that do not produce any recognisable change in the material, such as thermal vibrations around static equilibrium positions. This is achieved by taking a higher level of abstraction compared to MD:

- The material is no longer described as a group of atoms that can be positioned at any place in the volume. Instead, the volume is partitioned into sub-volumes, for which the only available information is the chemical nature of the atom present there, if any. Partitioning into sub-volumes is a relatively obvious matter for single-crystals, such as the one depicted in Fig. 2.3. As atoms remain essentially near their equilibrium positions, the obvious choice is to partition the volume with Wigner-Seitz cells, centred around the perfect lattice positions, as depicted in Fig. 2.4.

- Migration events are defined at every step, and a frequency Γ is assigned to each of them. For the example depicted in Fig. 2.4, an obvious choice of events is the migration of the vacancy to any of its nearest-neighbour positions.

- One event is stochastically chosen according to the MC algorithm described in what follows, and applied.

The AKMC is therefore said to be a rigid lattice model, because information about static or dynamic relaxation is not available in the way the system is described. Off-lattice AKMC concepts also exist from a recent time (106), but are quite naturally more complex and require longer CPU times. We thus remain in a rigid-lattice framework in the thesis. Equivalence with MD, for long processes, depends on several conditions:

- One necessary condition is the existence of a unique relation between the same system, described in the AKMC rigid lattice world, and described in a the non-rigid lattice world of MD. This unicity depends entirely on the system being studied, and must be therefore evaluated case-by-case.

non rigid lattice world

rigid lattice world

Wigner-Seitz cells
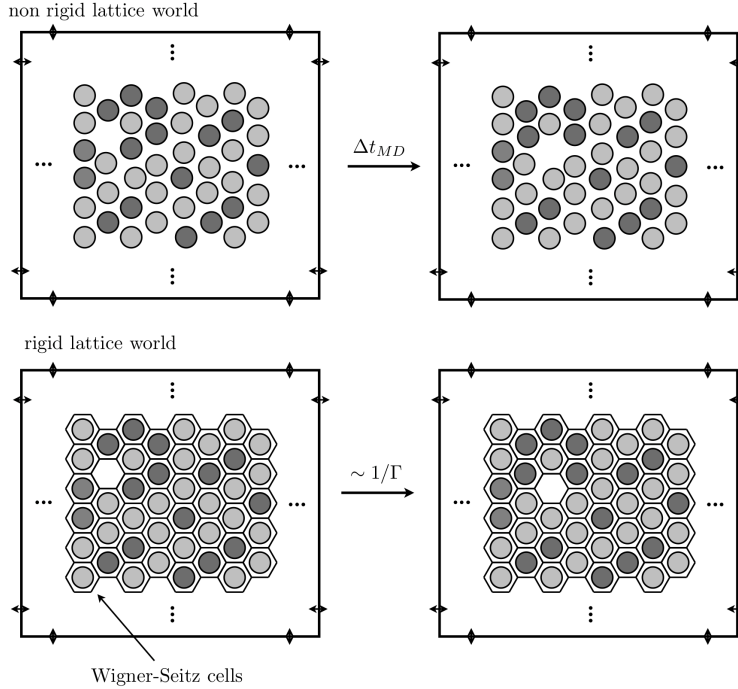
Figure 2.4: Vacancy migration event described in a non rigid lattice world (top part), such as for example in a MD simulation, or in a rigid lattice world (bottom part), such as for example the AKMC simulation. In the latter, the exact atomic positions are not given, only the chemical nature of the atom sitting in predefined volumes, often taken to be the Wigner-Seitz cells delimiting atomic volumes around the perfect lattice positions of the crystallographic structure.

- For some systems, the respect of this unicity can depend on the advancement of the simulated process. A good example is the simulation of thermal ageing of dilute Fe-Cu alloys, as will be discussed later in section 4.3, because crystallographic changes occur in the bulk of big clusters of Cu atoms, if they are formed.

- The last condition is that the migration events are properly defined. The example depicted in Fig. 2.4 is, again, obvious because the only relevant events are the migration of the vacancy to a 1nn position. In other cases the definition of the migration events is not equally straightforward.

Figure 2.5: Wigner-Seitz cells for the bcc crystallographic structure.

Regarding the discussions above, it is clear that AKMC is an obvious model to describe simple systems such as bulk materials with only one vacancy. For more complex systems, however, the definition of a proper algorithm is far less obvious, such as for example in presence of interfaces like grain boundaries, dislocations, systems were crystallographic changes occur,... In these cases, a proper partitioning of the space, and a relevant definition of migration events, are not straightforward at all. Similar problems occur for the simulation of the effects of neutron irradiation in steels, even if the attention is focussed on bulk materials, because of the presence of many vacancies and self-interstitial atoms, as will be discussed later. This explains why AKMC methods have been especially popular only for the simulation of processes where only one vacancy must be included in the box, such as thermal ageing, or a number of cases when attempting to simulate irradiation effects.

2.2.2   Algorithm

Fig. 2.5 shows the Wigner-Seitz cells that correspond to the case of bcc bulk materials. If one vacancy is introduced in the simulation box, the only relevant migration events to be defined are the migration of the vacancy to one of its eight 1nn positions, as depicted in Fig. 2.6.

After the list of N possible migration events is made, the next step is to assign a frequency of occurrence Γ_i for all of them ($i = 1, \ldots, N$). One event is then chosen with a simple Russian roulette algorithm[1]: a random number R is first chosen:

$$R = rand\,(0,1) \tag{2.1}$$

---

[1]The name "Monte Carlo" comes in fact exactly from the association of the method with the gambling games in Monte Carlo's casino.
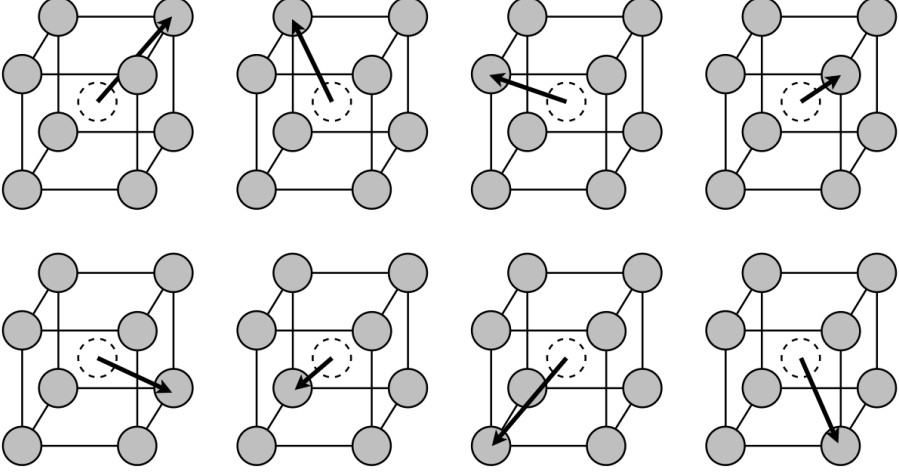
Figure 2.6: Eight migration events defined in the AKMC simulation for the single-vacancy case in bulk bcc materials: jump of the vacancy to one of its 1nn.

The event k to apply is the only one that fulfils the condition:

$$\sum_{i=1}^{k-1} \Gamma_i < R \sum_{i=1}^{N} \Gamma_i < \sum_{i=1}^{k} \Gamma_i \qquad (2.2)$$

Finally, time is incremented following the mean residence time principle:

$$r = rand(0,1) \qquad (2.3)$$

$$\Delta t = \frac{-ln(r)}{\sum_{i=1}^{N} \Gamma_i} \qquad (2.4)$$

Asides from the relevance of how the volume has been partitioned in sub-volumes, and how migration events were defined, the quality of the AKMC algorithm depends **entirely** on how the migration frequencies Γ_i are calculated.

2.2.3   Calculation of point-defects jump frequencies

As explained in the previous section, frequencies of occurrence must be assigned to all events that are encountered during the AKMC simulation. These events correspond to the migration of point-defects to a nearest-neighbouring position, that
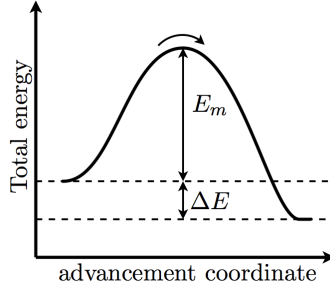
Figure 2.7: Schematic illustration of the minimum energy path followed by the system during a vacancy migration event. The migration energy is denoted as E_m, and the energy difference between the final and the initial state is denoted as ΔE.

are thermally activated. The frequency of occurrence can therefore be expressed as:

$$\Gamma = \Gamma_0 \cdot \exp\left(\frac{-E_m}{k_B \cdot T}\right) \tag{2.5}$$

Here, k_B is Boltzmann's constant and T is the absolute temperature in Kelvin. The prefactor Γ_0 is an attempt frequency, and E_m is the migration energy, assumed to be independent of temperature, as shown in Fig. 2.2. Such a description is derived from the theory of thermally activated processes, and is justified when the thermal oscillations of atoms around their equilibrium positions are smaller than the activation energies, i.e. if $k_B T \ll E_m$.

The migration energy E_m is the energy barrier that the migrating atom must overcome to escape from its present equilibrium position to another one found in the lattice site occupied by the vacancy. Because of the exponential function in Eq. 2.5, a proper calculation of E_m is unquestionably the most important requirement to ensure that the migration frequency is accurately calculated. The attempt frequency Γ_0 is potentially less influencing the migration frequency, because it is usually expected to vary from one case to another by one order of magnitude at the most. Many authors of AKMC algorithms therefore consider it as a constant, and all the efforts are devoted to a proper calculation of the migration energy.

Kang and Weinberg (87) proposed that the energy barrier for the migration of a point-defect can be decomposed as:

$$E_m = \epsilon + \Delta E \tag{2.6}$$

Here ΔE is the total energy change associated with the change of thermodynamic state between after and before the jump, and ϵ is the excess energy, to be added to obtain the complete barrier, whose value is a priori unknown and will also be, in

principle, a function of the local atomic environment (LAE) (97). Based on these premises, they proposed a recipe to 'properly' assess the energy barrier, given a harmonic potential for the atoms at equilibrium (87). In practice, however, such a decomposition has been later most often used assuming ϵ = constant (e.g. FISE method in (178), see also (13; 21; 23; 164; 176; 177)). Somewhat more sophisticated methods, based on broken-bond considerations (13; 14; 158; 164) extended also to the saddle point (97), have been used as well. In recent times, in addition, attempts at increasing the reliability of these heuristic methods have been made, by fitting the pair energy parameters of the model directly to DFT calculations (13; 163; 164; 176–178). These heuristic methods do take into account the influence of the chemical environment on the energy barrier, but they do so in an oversimplified manner and totally disregard the effect of faraway atoms.

At the other end of the spectrum of possible AKMC models, Henkelman and Jónsson (75) envisaged a way to eliminate not only the rigid lattice approximation, but also the predefined list of possible migration events. In their scheme, the choice of the migration events, together with the calculation of their corresponding migration barriers, is made on-the-fly by applying the dimer method (72). The latter is a numerical method that allows, given the initial state, all possible transition paths to another nearby local minimum in the potential energy surface to be found. The advantage of such an approach is clearly its flexibility with respect to systems where the rigid lattice approximation would no longer be valid, such as at free surfaces and grain boundaries, or in the bulk of materials containing dislocations or nanostructural features such as nano-voids and dislocation loops. In addition, the dimer-method-calculated migration energy will always take into account all effects, chemical and due to strain fields (at zero Kelvin). The main drawbacks of this method are its complexity and its high computational cost. Similar ideas were also explored by other authors, such as for example Trushin et al in Ref. (172) or Mellouhi et al. in Ref. (113).

Recently, methods using sophisticated mathematical techniques to regress energy barriers as functions of the LAE, such as cluster expansion (155; 173), or genetic programming (157), have been proposed for use in AKMC models. These approaches keep a rigid lattice description of the system, but calculate a number of examples of energy barriers between two given states on a non-rigid lattice, i.e. allowing for relaxation effects, using nudged elastic bands (see below) or other methods. These sets of examples are then employed to fit mathematical expressions to them. However, so far these approaches have been based on a limited number of examples (that can practically be calculated with DFT methods) and, either little importance has been given to verifying the capability of the obtained mathematical expression to predict never seen cases (local environments) (173), or the total amount of possible cases was any way relatively small (157).

2.3 New algorithm, based on artificial neural networks

In this section, a new approach is proposed for the calculation of the point-defect migration energy. As already discussed, AKMC algorithms are faced with a dual challenge: the point-defect migration frequencies should be calculated with as little approximations as possible in order to respect the physics of the process being studied; On the other hand, the complexity should be reduced as much as possible, in order to allow the simulation of long and complex processes in an affordable CPU time.

The new methodology proposed here happens to be a reasonable compromise between these antagonistic requirements:

- The migration energy (E_m in Eq. 2.5) is calculated with the nudged elastic band (NEB) method (74; 84). The latter takes into account the effects of long-range chemical interactions and static relaxation, and can therefore be regarded as very rigorous. The migration attempt frequency (Γ_0 in Eq. 2.5), that in fact also varies with the LAE like the migration energy, is, in first approximation, considered as a constant, similarly to the work performed by mist authors working with AKMC methods (see discussion in section 2.2.3). It could be calculated using the Vineyard method (179), that requires, to summarize, the CPU-time-consuming calculation of the vibrational modes of the system at the initial state and at the saddle point.

- The use of the NEB method on-the-fly during the simulation cannot practically be performed, because of the relatively high number of operations that are necessary to accomplish such calculation. An artificial neural network (ANN) is therefore trained to replace NEB, by providing an estimation of the migration energy, given as input a description of the local atomic environment (LAE) of the migrating point-defect.

- The ANN is therefore used as a powerful and general numerical method, that is trained on the basis of examples, calculated with the NEB method. A successful design of the ANN regression, however, inevitably requires a certain number of these examples, that cannot be easily obtained with ab initio methods. For this reason, the work performed in the thesis is entirely based on the use of IAP. The methodology proposed in this section can, however, be applied using ab initio calculations in the future.

Early attempts at developing this new approach were first published in 2007 by Djurabekova et al. in Ref. (46). At that time, the bottleneck was mostly the efficiency of the ANN regression, and the potential applicability of the method was, globally, limited to simple cases. Later, in 2008, improvements to the ANN techniques gave

more satisfactory results, and the feasibility was demonstrated in a more convincing way in Ref. (32) (Paper I in appendix B). A detailed description of the approach was finally published in Ref. (35) (Paper IV in appendix E).

In this section, the fundamentals of the proposed algorithm are described, to provide the reader with a concise overview of the different aspects necessary to implement the method. In section 2.3.2, the exact procedure followed to calculate migration energies with the NEB method is summarized. Next, section 2.3.2 explains how an ANN can be trained to replace NEB. Lastly, section 2.3.3 discusses how the AKMC algorithm can be parallelised.

2.3.1 Calculation of the point-defects migration energies with the nudged elastic band method

Several methods exist for the calculation of migration energies. The interested reader can find a general overview in a publication by Henkelman et al. in Ref. (73). Existing methods can, roughly, be divided in two categories:

1. In the first category, the initial and the final states of the system are given. In this case, the method searches for the minimum energy barrier that must be overcome, at zero Kelvins, for a transition from the initial metastable state to the final one. One example is the drag method: the migrating atom is gradually pushed in the direction of its final position, and the whole system is completely relaxed at every step. The migrating atom is of course constrained in the normal plane to its trajectory, to prevent it from going back to the initial position. This approach can be practically implemented in several different ways, that vary in the way the drag direction is defined, and the migrating atom is constrained. A drag method was used at the beginning of the thesis (see Ref. (32), Paper I in appendix B), for reasons of simplicity, but was later on abandoned for the NEB, because the latter finds usually lower energy barriers.

2. In the second category, the final state is unknown. The methods are therefore not only intended to calculate migration energies, but more generally to make a list of possible state transitions from a given metastable state. A popular example is Henkelman's dimer method (72) as already discussed before. An ameliorated version of the dimer, named monomer, was later proposed by Ramunni et al. (148). Another example of transition search methods is the activation-relaxation technique (ART), proposed by Barkema and Moussaud in Ref. (11; 12; 119).

In the framework of the AKMC algorithm that we use, where possible migration events are explicitly pre-defined, the first category of methods is a natural choice, and the NEB is, amongst the other possibilities, probably the most appropriate.

The NEB method, applied to the calculation of vacancy migration energies in AKMC simulations, proceeds in 5 steps:

1. The initial and final states, defined in a rigid lattice way in the AKMC simulation box, are copied in two other boxes for a future use with molecular statics methods. The atomic coordinates of the atoms, that are unknown in the AKMC, are chosen to correspond to the regular lattice positions. To limit the calculation complexity, only portions of the AKMC boxes are in reality copied: all positions that are situated within a given range R_{LAE} from the migrating vacancy and the migrating atom. The latter are henceforth called the "local atomic environment" (LAE). The size of the new box, with non-rigid lattice, is chosen to be much smaller than the AKMC box. The missing sites, that are not copied from the AKMC box, are filled with matrix atoms, i.e. Fe.

2. The initial and final states are both statically relaxed (see discussion in section 2.1). In this work, this operation is performed with a conjugate gradients method (141).

3. A chain of intermediate states is constructed. They are initially taken to be a linear interpolation, for all atomic coordinates, between the initial and final positions.

4. All intermediate states are step-wisely relaxed. During this operation, at every step, the component of atomic forces tangent to the trajectory are modified to keep the intermediate images at a constant distance (from the advancement coordinate point of view) from each other: virtual tangent spring forces are defined for this purpose.

5. Last, the migration energy is computed as the difference between the total energy at the saddle point, and the total energy of the initial state.

This method is in practice easily applied automatically without any human intervention. A detailed analysis, afterwards, of the calculated energy paths is most of the times unnecessary. Parameters such as the size of the box, and the appropriate value for spring forces constant are easily optimized. The only practical problem is, as already discussed, the CPU time necessary to perform the complete calculation.

A necessary discussion, however, is the determination of the number of atomic neighbours of the vacancy and of the migrating atoms that should be copied from the AKMC simulation. The safest choice is to take the range R_{LAE} defined above very large, for example 3 or 4 times the cut-off of the potential, to ensure that the effects of further away atoms on the migration event, via atomic relaxations, vanish. But such a high precision is in practice not necessary. A discussion in this respect is given in Ref. (35) (Paper IV in appendix E, section II). In the case of the single

Local atomic environment (LAE): | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 2 |
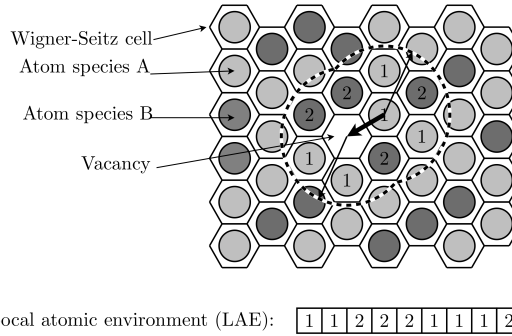
Figure 2.8: Coding of the local atomic environment describing a vacancy migration event as a string of integers.

vacancy in Fe-Cr or Fe-Cu alloys, the energy barrier is influenced by neighbours up to, roughly, a distance equal to 1.5 times the IAP cut-off. This range of sensitivity can be a priori much larger for more complex cases, such as the migration of SIA. For this reason, an appropriate study must be performed for every new application to determine R_{LAE}, following a similar procedure as described in Ref. (35) (Paper IV in appendix E, section II).

2.3.2 Prediction of the migration energies with artificial neural networks

"C'est ici que les romains s'empoignèrent !"

Captain Haddock, in "the adventures of Tintin", Hergé.

It was explained in the previous section that the vacancy migration energy can be automatically calculated with the NEB method. The simulation with AKMC algorithms of any physical process (such as segregation and precipitation) requires more than billions of events, and multiple migration energies must be calculated for everyone of them. The systematic use of NEB is, therefore, too prohibitive. In this section, a predicting tool is designed to replace NEB, for practical use at every step of the AKMC simulation.

The migration energy E_m depends entirely on the LAE. In the rigid lattice world of the AKMC simulation, the only available information about the LAE is the chemical species of the atoms sitting in the neighbouring positions of the migrating vacancy

1nn	2nn	3nn	E_m (eV)
1 2 1 1 ... 1 1 1 1	1 1 1 1 ... 1 2 1 1	1 1 1 1 ... 1 2 1 1	0.594
1 1 2 1 ... 1 2 1 1	1 1 2 1 ... 1 1 1 1	1 2 1 1 ... 1 1 1 1	0.671
2 1 1 1 ... 1 1 1 1	1 1 1 1 ... 1 1 2 1	1 1 1 1 ... 1 1 2 1	0.668
1 1 2 1 ... 1 2 1 1	1 2 1 1 ... 1 1 1 1	1 1 2 1 ... 1 1 1 1	0.525
1 1 2 2 ... 1 1 1 1	1 1 1 1 ... 2 1 1 1	1 2 1 1 ... 1 1 1 1	0.993
2 2 2 1 ... 2 1 2 1	1 1 2 1 ... 1 1 1 1	1 2 1 1 ... 1 1 2 1	0.479
1 1 1 1 ... 1 1 1 1	1 1 1 1 ... 2 1 1 1	1 1 2 1 ... 1 2 1 1	0.653
2 1 1 1 ... 1 2 1 1	1 1 1 1 ... 1 1 2 1	1 1 1 1 ... 1 1 2 1	0.098
1 2 1 1 ... 1 1 1 1	1 2 1 1 ... 1 1 1 1	1 1 1 1 ... 1 1 1 1	0.798
2 2 2 2 ... 1 2 1 2	1 1 2 1 ... 1 2 1 1	2 1 1 2 ... 1 2 1 1	0.865

Table 2.1: Table of examples of migration energy E_m versus LAE, for a binary alloy, that includes the 1nn, 2nn and 3nn of the migrating vacancy and migrating atom.

and of the migrating atom. An obvious and simple choice of input variables for the predicting tool is thus to represent the LAE as a string of integers, as depicted in Fig. 2.8: every integer of the string represent one atomic site of the LAE, and the value it takes describes the chemical species, e.g. 1 stands for Fe, 2 stands for Cu, 2 stands for Cr,... These integers are of a categorical type, meaning that the value of the integer is arbitrarily chosen. The generation of large tables of examples of E_m versus LAE, for varying LAE, is relatively straightforward to perform in practice, the only bottleneck being the limited CPU time. Table 2.1 shows an example of such a table. For convenience, the LAE is sorted by groups of atomic sites that correspond to the same shell of close neighbours of either the vacancy or the migrating atom. The groups are denoted 1nn, 2nn,..., Xnn. Fig. 2.9 shows the evolution with Xnn of the distance from the vacancy or migrating atom, and the corresponding number of atomic sites (cumulated with all previous shells). The order of the LAE entries, within the same Xnn group, is of course not important, as long as it remains the same for the whole table.

As an order of magnitude, the vacancy migration energy is, according to the IAP used in the thesis, sensitive to the LAE up to a distance as large as, roughly, $1.5 \cdot r_{co}$, where r_{co} is the cut-off of the IAP. This will be discussed later in chapter 4. For the potentials used on the thesis, r_{co} encompasses the 5nn. As a consequence, a number as large as 300 to 400 atomic sites should be taken into account by the predictive tool !

The objective to design the predicting tool can be regarded as the design of a mere numerical regression, though, perhaps, a bit ambitious considering the number of input variables involved, and the fact that these variables are categorical integers. The numerical regression should have the following qualities:
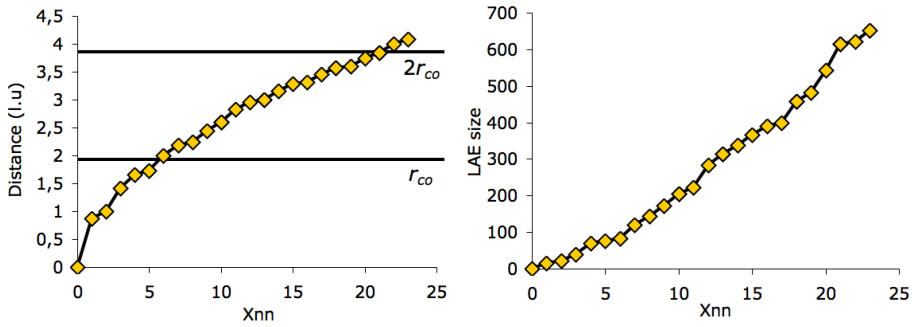
Figure 2.9: Evolution with the number Xnn of shells of close neighbours defining the LAE, for bcc single-crystals, of (left) the distance from the migrating vacancy of the migrating atom; (right) the total number of atomic sites concerned.

- Ideally, the regression should be expressed in a way that 'makes sense' from the physical point of view, i.e. it should be written in a form where all terms stand for a given contribution from the environment to the energy barrier. This quality would give a certain confidence that the regression, that could in that case be called a model, is capable to make relevant predictions of E_m for new cases that were never encountered before.

- The average error of prediction should the the smallest possible, ideally of the order of the NEB accuracy, in which case the regression could be considered as perfect.

- Lastly, the number of computing operations required to calculate E_m should be the smallest possible. This requirement is obvious, since the objective is the avoid the use of NEB for the only reason of CPU time limitations.

The development of a predicting tool with all these qualities is not at all an easy task. The first point above is, for instance, particularly idealistic, or at least very ambitious. Energy formalisms like the broken-bond can be regarded as an attempt to accomplish it: the migration energy is, there, expressed as the consequence of the breaking of atomic bonds while the migrating atom is attempting to escape from its equilibrium position towards another one. This kind of formalism can give satisfactory results, especially if the number of atomic sites to take into account is limited, but is in the end a simplified description of the real complexity of the interactions between the migrating atom and the LAE. Nothing therefore guarantees that the formalism is including all important contributions. Another problem that occurs with this kind of approaches is that the number of degrees of freedom explodes if more neighbouring sites are taken into account. This last point is

particularly limiting in practice, because any model or numerical regression must be fitted using at least a number of examples of I/O that is larger than the number of degrees of freedom.

The use of artificial neural networks as predicting tool follows a totally different philosophy. As will be discussed later in chapter 3, ANNs can be regarded as a general regression method, that is not at all intended to implement a description of the physical phenomenon being fitted: if written down, the expression of the output, giving the inputs, is a mixture of arbitrarily chosen non-linear functions, that are embedded in each others. Individual terms with a physical interpretation of the role they play on the final output cannot be identified in practice. This disadvantage is, however, compensated by a number of qualities:

- The application of ANNs to any problem, for instance finding a numerical logic between the LAE and the corresponding E_m, is completely automatised, and self-constructed. The user of ANN must therefore not make any assumption at all concerning the mathematical form of the relation.

- More importantly, the flexibility of the ANN is such that complex regression problems, such as the one we have here, can be efficiently solved with a rather simple expression that does not contain a huge number of degrees of freedom. In other words, ANNs can, in practice, be trained on the basis of a limited amount of examples, and this number does not explode if the complexity of the problem is increased.

The drawback of this generality is that the logic constructed during ANN training can, eventually, be out of control from the user's point of view. The choice of the examples constituting the training table can, consequently, make that logic unsuitable for a general use after training. Overspecialisation, for instance, is the extreme pathology where the network commits a very small prediction error on the particular list of examples used for training, but very large errors for other cases.

A more detailed description and discussion of the ANN technique is reserved for chapter 3. The questions addressed there are:

- How should the table of examples of E_m versus LAE be designed ? Should we take examples of LAE at random, or are other choices more suitable ?

- After training is completed, how confident can we be in the ANN ability to make relevant predictions for new cases ?

What is already clear at this point, is that the ANN qualities are entirely depending on the quality of the table used to train it. A necessary condition, or at least a wise precaution, is to design the table of example in such a way that it contains examples

that are reasonably representative of the different situations that the vacancy can encounter during the simulation in our specific case. This consideration will be discussed also in chapter 3, but in reality is more a problem of application than a problem of ANN technique. The definition of "reasonably representative" is to be debated before every application, and is therefore discussed in chapters 4 and 5.

2.3.3 Parallelization

Aside from designing a fast predictive tool for the vacancy migration energy, AKMC simulations can be speeded up with parallelisation techniques too. Parallelisation consists in defining, or re-defining, an algorithm is such a way that the whole calculation can be divided in parts that are independent of each other. If such a division is possible, it becomes possible to execute them separately, but simultaneously, using different processors. The total CPU time required to perform the calculation is, obviously, identical, but the real (wall) time to do so is divided by the number of processors used (assuming a perfect efficiency).

The AKMC algorithm can be parallelised in two different ways, that are not mutually excluding:

1. The simulation box can be partitioned in equal pieces, and each of them managed by one processor. This kind of approach has been successfully taken for MD simulations for example. Unfortunately, the AKMC algorithm is far less easy to parallelise in that way, for the simple reason that the simulated time is a consequence of the events that are encountered and chosen, differently from MD where, like in real life, events are a consequence of time. In other words, if the AKMC box is partitioned and if each pieces are dealt with a separate processor, the simulated time is going to be heterogeneous from one box to another. A possibility to annihilate this effect is to introduce null events in the simulation, that can be defined in such a way that all partitions advance at the same pace. This approach was proposed in Ref. (105).

2. The most time consuming operation in the AKMC algorithm is, clearly, the calculation of the point-defects migration frequencies. This is especially true for the ANN-based AKMC, because the calculation is performed with ANN that are fast in absolute basis but nonetheless represent a certain amount of elementary operations. In addition, the time required to read in the simulation box the necessary information to generate the ANN input variables (the LAE coded as a string of integers) is a somehow complex operation, too. These operations must be repeated for all possible events, 8 in the case of single-vacancy in bulk bcc materials, and are completely independent of each other. To share them between several processors is therefore relatively easy. The

disadvantage is that the maximum number of processors is limited to the maximum number of events encountered at every steps of the simulation, i.e. 8, which is not very large. Parallelisation can thus only speed-up the simulation by less than an order of magnitude.

Only the second approach has been used in the thesis, using the open-MPI library (202; 203). As an order of magnitude, a simulation for a single-vacancy problem could be speeded up by a factor 6 on a 8-processors personal computer.

2.4 Conclusion

In this chapter, the different tools of the multiscale modelling scheme linking the electronic level (ab initio) to the coarse-grained models have been briefly described. A new atomistic kinetic Monte Carlo (AKMC) algorithm has been proposed, where the vacancy migration energy is calculated with little approximations, taking all effects of long-ranged atomic relaxations and chemical interactions into account. The algorithm relies on the use of artificial neural networks (ANN) to allow for faster on-the-fly calculation, thereby making the simulation of long processes possible. The chapter focussed on the single-vacancy case, for reasons of simplicity, but also for reasons of traditions, because AKMC algorithms have been hitherto mostly used in that framework. Applications of the new algorithm to such problems are presented in chapter 4. Before, in chapter 3, the ANN technique is described and discussed in details, to provide all the necessary theoretical background. In particular, the exact procedure that is followed for the design of suitable ANN for the application of concern is described.

3 Artificial neural networks: a practical form of artificial intelligence

In chapter 2, a new AKMC algorithm was proposed, where migration energies associated with vacancy migration events are first rigorously calculated with a CPU-time consuming method, the NEB, and then, to speed-up the simulation and make it applicable to real problems within an affordable computing time, a predicting tool was designed to replace NEB for fast on-the-fly calculations. Artificial neural networks (ANN), that belong to the artificial intelligence family, were chosen to build this predicting tool, in consideration of their high flexibility and, more importantly, their ability to solve complex numerical problems.

ANNs are a concept of *weak artificial intelligence*, that more precisely belong to the branch of *machine learning* or *computational intelligence*. The objective that is followed is therefore not to create paradigms that exhibit real intelligence like human beings, but, in a pragmatic approach proper to engineers and also to many scientists, rather to propose systems, in our case here computer programs, that are more autonomous. ANNs, in this respect, are self-learning systems that extract hidden knowledge from their environment and manage to take relevant actions when entirely new situations are encountered. For the particular application to the design of numerical predicting tools, which is the major interest for the thesis, ANNs are *surrogate models* that present the advantage to spare to the user the need to formulate knowledge about the problem at hand explicitly. This is the reason why, despite the fact that the inexpert eye could qualify ANN as mere regression tools, just like polynomial functions for example, the words "artificial intelligence" do have both philosophical and practical meanings.

A complete description of the theoretical and mathematical frameworks of ANN's

Figure 3.1: Example of feed-forwards multilayer perceptron with 4 input variables (x_1, x_2, x_3 and x_4), one hidden layer with $H = 6$ nodes and one output (O). The right part of the figure shows the detail inside a node of the hidden layer. The signal is propagated from the left to the right of the network, in a layer-by-layer fashion.

is far beyond the scope of this document. For a complete and detailed reference textbook, the author recommends Bishop's book (16) to the interested reader, which is one of the most popular in this field. A quick and more general overview of the subject can also be found in Ref. (15) by the same author. Several other websites proposing answers to FAQ's such as the one in (156) also provide many useful information. The chapter starts in section 3.1 by a detailed description of the implementation of ANN that is used in the thesis, namely the *multilayer perceptron*, and summarizes its fundamental theoretical and practical aspects. Next, the application of ANNs to two sensibly different problems, both connected to nuclear materials science, is presented. In section 3.3, ANNs are used to extract knowledge from a limited amount of experimental data that are very expensive to generate, by predicting the radiation induced hardening of reactor pressure vessel steels. Finally, the application to the prediction of point-defects migration energies in AKMC simulations, which is the objective of the present thesis, is described in detail in section 3.4.

3.1 The multi-layer perceptron: a universal approximation machine

Several types of ANN exist, and each of them is best suited for a particular type of application. The so-called *multilayer perceptron* depicted in Fig. 3.1, for instance,

provides appropriate solutions for the design of general numerical regressions:

- The network is constituted of several layers of nodes. Each node is a simple processing unit that takes, alone, a simple behaviour. A unique output signal results from the passing through of input signals.

- On the left side of the figure, the input layer is the collection of the raw input signals of the whole network. On the right side, the last layer of nodes is called the output layer. In the middle, several intermediate layers of nodes can be introduced. They are called *hidden layers*, for the reason that they are, in practice, invisible to the user. The network depicted in the figure counts four input signals, one single hidden layer with H nodes, and, finally, one single node in the output layer.

- There is in theory no restriction about how nodes of different layers can be connected to each other. For simplicity, many networks do not allow for backwards connections: a given node never receives as input output signals coming from nodes of either the same layer, or from the next layers (see e.g. Ref. (54)). Such a network is qualified as *feedforwards*, and present the advantage that the calculation of the output is, as the name tells, straightforward, because no iteration due to backwards propagation of the signal is necessary.

- Last, the network shown in the figure is said to be *fully connected*, because all nodes in a given layer receive as input all output signals coming from the immediately preceding layer, and no *layer bypass* is allowed. The inputs for the first (and only) hidden layer are the raw input signals of the network.

The output y_j of a hidden node j reads:

$$y_j = \phi(v_j) \tag{3.1}$$

$$v_j = w_{j0} + \sum_{i=1}^{N} w_{ji} \cdot x_i \tag{3.2}$$

Here, v_j is the *internal activity* of the node. Function ϕ can, in principle, be any non-linear function. For convenience, to prevent large magnitudes of the signal, function ϕ is almost always taken to be bounded, most of the time a hyperbollic tangent or sigmoid function such as depicted in Fig. 3.1. Abnormally large internal activities are therefore blocked by *node saturation*, and the impact on the rest of the network is limited. The internal activity is calculated as a weighted sum of the input signals to the node, i.e. the input signals of the network. The output z_O of the only node in the

output layer is calculated in the same way:

$$z_O \quad = \quad \phi(v_O) \tag{3.3}$$

$$v_O \quad = \quad w_{O0} + \sum_{j=1}^{H} w_{Oj} \cdot y_j \tag{3.4}$$

In Eq. 3.2 and 3.4, the weights w_{0j}, w_{ji}, w_O and w_{Oj} monitor the strength of interaction between the nodes, often denoted as *neurons*, and are therefore called *synaptic weights*, or also *synapses*, in analogy to the role of real synapses in biological neural networks.

The interest of the feed-forwards multilayer perceptron, in its simplest form as depicted in Fig. 3.1, is that it fulfils the *universal approximation theorem* (78): For any continuous function $F(x_1, \ldots, x_n)$, there exist an ANN with a finite number of hidden nodes that fulfils the following condition, for all set of inputs x_i and all ϵ:

$$|F(x_1, \ldots, x_n) - O(x_1, \ldots, x_n)| < \epsilon \tag{3.5}$$

Here, $O(x_1, \ldots, x_n)$ is the output of the ANN. The multilayer perceptron can therefore be regarded as a universal approximation machine that can, in theory, perfectly solve any problem of numerical regression[1]. The problem of designing such a network is discussed in the next section.

3.2 Practical application of ANN to a given problem

In the thesis, ANNs are designed to implement numerical regression tools. This is therefore a *supervised training problem*, meaning that the ANN is trained to "understand" and reproduce at best the logic between input and output, on the basis of a finite number of examples provided by the user. Unfortunately, the universal approximation theorem only demonstrates the existence of a perfect ANN to solve the problem a hand, but does not give any indication about how it can be found in practice, and more importantly, if it can be obtained by training on the basis of a finite number of examples. For this reason, ANN training, that can be regarded as the problem of determining the ideal architecture of the network, and also the ideal numerical value for the synaptic weights, is, in the absence of an appropriate theoretical framework, solved as a mere optimization problem that consists in the minimization of the objective function:

$$f = \sum_{i=1}^{N} (o_i - d_i)^2 \tag{3.6}$$

---

[1]The interested reader can find a general discussion about that theoretical ability, for any problem other than numerical regressions, for example in the section "What can you do with an NN and what not?" in (156).

Here N is the number of available examples of I/O, d_i is the desired output for the example i, and finally o_i is the corresponding actual output of the network.

3.2.1 Training the ANN, given the architecture

Let's assume for the moment that the ANN architecture is fixed, i.e. that the number H of nodes in the hidden layer is chosen. The minimization of function f is thus the problem of determining the optimal numerical value of the synaptic weights, and can be undertaken by a number of different algorithms that cannot be exhaustively listed here. All these algorithms are iterative, meaning that synapses are first randomly initialized, and then iteratively updated following a given algorithm. These iterations are called *epochs* in ANN jargon. To the author's opinion, only algorithms based on the use of the analytical expression of the first order derivatives of f with respect to the synapses, that can be obtained using the so-called *back-propagation* algorithm (153), are really promising in general. Four algorithms were used during the thesis:

- The so-called *Delta-bar-delta* (DBD), described for example in Ref. (70), based on a mere "steepest descent" principle, with a fixed update factor of the synapses. For this reason, it is often wrongly denoted as back-propagation in old articles or books. This algorithm presents the advantage of being simple of use, because only the gradients must be calculated at every epoch, but gives in practice poor results, at least for the problems of interest in this thesis.

- A much more efficient algorithm named *resilient propagation* (RPROP), that is in fact almost equally simple of use, was proposed by Riedmiller and Braun in Ref. (151). In simple words, it is based on the idea that the magnitude of update of synapses should be decoupled from the magnitude of the gradients. Quite satisfactory results were obtained using this algorithm, but it still presents limitations of slow convergence, as probably unavoidable for first order optimization methods.

- Faster convergence is achieved with a more sophisticated algorithm, qualified in the field of optimization as a "1.5 order method", that is the conjugate gradients (CG). The application of CG to ANN training proceeds as for any non-linear minimization problem. The complexity is higher compared to the DBD and RPROP algorithm, mainly because a one-dimensional optimization must be performed at every epoch.

- In general, significantly better results were obtained during the thesis using an even more sophisticated approach, proposed by Levenberg (98) and Marquardt (103) (LM). In simple words, this is an exact approximation of the second order Newton method if the crossed second order derivatives are nil.

The major drawback of LM is the relatively high complexity, mainly because a system of linear equations, one per synapse in the network, must be solved. It is thus a priori best suited for small size problems. Despite this drawback, LM is by far the most successful training algorithm used in the thesis, and all results shown in this document were exclusively obtained using it, or more complex algorithms based on it.

- In addition, other algorithms were explored for the sake of completeness. Methods for improving back-propagation algorithms can for example be found in Ref. (191). Other algorithms such as Fahlman's *Quick-prop* in Ref. (53), a *training algorithm utilizing multiple sets of linear equations* proposed by Chen et al. in Ref. (42), and last a modification of the LM algorithm to reduce its complexity in Ref. (184), were tried.

All these algorithms are starting point dependent, meaning that there is a problem of decision of the initial values for the network synapses before training is started. This question is somehow long-debated in reference textbooks on ANN (16), for example in the book of Bishop, but also in the book of Haykin (70). Several strategies are proposed in the later, and were all investigated during the thesis. No major difference in the final quality of the network is, finally, observed if the LM training algorithm is used, contrary to the other ones listed above, which is another indication of the superiority of this algorithm. Random initialization between -1 and 1 is proved to be good enough in the framework of this thesis. It however does not mean that improvement, in the future, or for totally different applications, could not be achieved if a more appropriate initialization method is found. Other authors propose different methods, such as for example genetic algorithms in Ref. (2), or the method proposed by Nguyen and Widrow in Ref. (123) that has some popularity.

Finally, a major concern while training ANNs is to guarantee, or at least provide good indication, that the predictions for new sets of inputs are equally accurate compared to predictions on the available set of examples used for training. This concern actually translates the major drawback of the generality of ANN, and of the fact that no knowledge about the problem at hand is explicitly formulated by the user. In other words, the risk that the ANN, during training, does not develop a general logic but in reality, somehow, merely memorizes the complete set of available examples, is real. An example with one input variable is illustrated in the left part of Fig. 3.2. We see that reasonable interpolation is achieved by the ANN if the latter is not too complex, i.e. if not too many nodes are included in the hidden layer. Predictions are however not equally accurate for all training examples. If the number of hidden nodes is increased, the ANN manages more accurate predictions for all known points, but clearly looses generality, and is in fact merely reproducing the training data without being able to make any sensible interpolation. This pathology can rather easily be detected for one-dimensional problems, especially

such as depicted in Fig. 3.2 that is of a simple kind, but in general, however, over-specialization cannot be easily identified only on the basis of a limited amount of examples of I/O, especially if the dimensionality of the problem at hand is large. ANN training must therefore be *regularized*, to favour the development of extrapolation or interpolation skills. The most common regularization approach is called *early stopping*, and is based on the idea that memorization of the provided example, or more generally can we say *network overspecialisation*, develops only at a certain moment of training, i.e. after a certain number of training epochs. The most natural way to prevent it is therefore to divide the available table of examples of I/O into two different and non overlapping sets:

- The training set is used to minimize function f in Eq. 3.6. Only these examples are thus used to calculate the gradients with the backpropagation algorithm, and the synapses are updated during each epoch taking only them into account.

- The reference set is used to measure, after every epoch, the average error of prediction on new cases.

Fig. 3.2 shows the typical evolution of the average error of prediction on both sets during the training epochs. The error committed on the training set always decreases, or can in the worst case remain constant. The error on the reference set, however, ceases to decrease from a certain epoch, and then starts to increase, as a clear sign of the onset of overspecialisation. Training is therefore interrupted at that moment.

3.2.2 Determination of the optimal ANN architecture

In section 3.2.1, algorithms to find the optimal numerical value for the network synapses were listed and discussed. The next concern for the user of ANN is thus to determine the optimal architecture of the network. Assuming that networks with one hidden layer are enough, the problem can therefore be tackled by a mere mono-parametric study: networks with increasing number H of hidden nodes are trained separately, and the one committing the lowest error on the training set is finally retained. The typical evolution of this error with H is depicted in Fig. 3.2. Too small a H naturally leads to higher errors of prediction, for the reason that not enough degrees of freedom are available in the network. Oppositely, too high a H increases the risk of overspecialisation, and the error can increase as well.

This approach presents the major drawback that, in practice, many independent trainings must be performed: several H must be explored, but also trainings for a given H must be repeated several times because of variance of the trained network
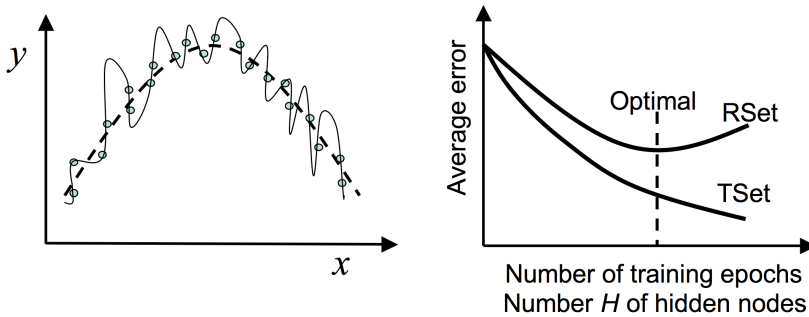
Figure 3.2: (left) Schematic illustration of ANN over-specialization problem for a simple problem with one variable x and one output y. Blue points represent data in the training set. The dashed line shows predictions by an ANN with a small number H of hidden nodes, whereas the plain line shows predictions of another ANN where too many hidden nodes were introduced. (right) Typical evolution of the average error committed on the training set (TSet) and reference set (RSet) with: (a) the number of training epochs, the number of hidden nodes being fixed; (b) the number of of hidden nodes, at at end of training.

with respect to synapse initialization. Many authors therefore proposed more involved approaches. The interested reader can find general overviews and critical reviews of many existing approaches in Ref. (39; 94; 135), that can, roughly, be classified in two categories:

- *Pruning* approaches (149) start from deliberately too big networks and iteratively remove nodes from the network[2]. The drawback is, a priori, that these approaches require as a first step the training of big networks, and is therefore probably not efficient from the computing time point of view.

- Oppositely, *constructive* approaches start from small networks and add new nodes iteratively. Two different popular algorithms were considered during the thesis:

  - The *dynamic node creation* of Ash (4), in which hidden nodes are added one by one in the sole hidden layer of the network.
  - The *cascade correlation* of Fahlman and Lebiere (55), in which hidden layers with one single node each, fully connected to all preceding nodes (therefore allowing for *layer bypass*), are iteratively added.

---

[2]Philosophically speaking, pruning approaches, to the author's experience, are a bit like writing a scientific paper or a thesis: in a first stage, too much details are written, and then the text is shortened to retain only the essential.

Both approaches gave satisfactory results, but the dynamic node creation was retained for reasons of simplicity, and was modified according to the topology of the problem at hand, as discussed in section 3.4.

These algorithms are aimed at automatizing the process of ANN training, avoiding manual parameter studies, but are in fact generally not necessarily expected to lead to more accurate networks. Nothing, indeed, guarantees that constructive or pruning algorithms are optimal training procedures because of, again, the absence of an appropriate theoretical framework.

3.3 Application to the prediction of radiation induced reactor pressure vessel steel hardening

As a spin-off application of the thesis, at least for the sake of curiosity and of making a broader use of the experience that was learned in the field, ANNs have been applied to a problem of sensibly different topology compared to the major problem of concern described in the next section. Namely, the ANN was used for the prediction of neutron irradiation induced hardening, a subject of interest in the field of nuclear material science. This work is completely described in a full length paper (37) (Paper VI in appendix G), but is here briefly summarized.

It is well known that reactor pressure vessel (RPV) steels used in light water nuclear reactors harden and embrittle under neutron irradiation (79). Hardening is customarily measured as the increase of the yield stress, with tensile tests performed on samples of the RPV steel. Nuclear regulations impose safety margins on this increase[3], according to rules that may somewhat change depending on the country, as safeguard against RPV failure in both service and accidental conditions. In the absence of a complete physical model, from the atomic to the macroscopic level, that can describe with accuracy the relevant processes taking place under irradiation, hardening and embrittlement are predicted by semi-empirical, or totally empirical models that are mostly based on numerical fittings using experimental data (20; 143). Although inadequate to cover all possible conditions, a large amount of data from surveillance capsules and from material test reactors does exist. One of the most important goals for utilities and other nuclear stakeholders is the development, based on "clever" interpolations and extrapolations of the available data, of reliable trend curves, providing estimates of steel embrittlement as a function of the most important influencing variables. ANN are therefore a potentially interesting candidate to achieve adequate predictions, because on the one

---

[3]All safety authorities actually impose a limitation in the shift of the ductile-to-brittle transition temperature, but in the case of RPV steels a strong correlation, generally linear, exists between this increase and the yield strength increase.
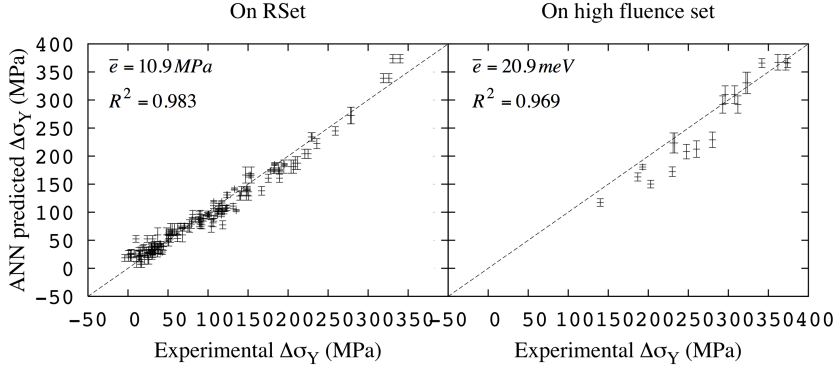
Figure 3.3: ANN quality of prediction of the neutron irradiation induced increase $\Delta\sigma_Y$ in the yield stress of RPV steels. The number of input variables is 4, and the number of hidden nodes is 4. Error bars show the standard deviation of predictions made by 25 different ANN, trained with the same training and reference sets. (left) Prediction on the reference set. (right) Prediction on a separate set that corresponds to higher neutron fluences, that was not used for training.

hand, as already discussed, of their abilities to extract hidden knowledge from data, and on the other hand thanks to the fact that knowledge about the physical process must not be explicitly formulated.

In Ref. (37), ANN was trained to predict the increase $\Delta\sigma_Y$ of the yield stress taking four inputs into account: the Cu content of the steel, the Ni content, the neutron fluence and finally the irradiation temperature. The RADAMO database (40) was used as set of 346 examples of I/O. For this application, the definition of a training and a reference set from the complete database is a delicate exercise, mainly because of the limited number of available examples, but also because of their sparse distribution in the space of the input variables and output. Two different algorithms to define the sets were proposed and compared. The achieved ANN quality of predictions on the reference set is shown in Fig. 3.3. Predictions on a separate set that corresponds to higher neutron fluences are also shown.

The conclusions in Ref. (37) were that ANNs can accurately predict hardening, and that the algorithm followed to define training and reference sets have an impact on the skills of extrapolation for higher neutron fluences. In a future work, the extrapolation skills for other steels chemical compositions will be evaluated, by using other databases of I/O in addition to RADAMO.

3.4 Application to the prediction of the vacancy migration energy in AKMC simulations

The problem of designing an ANN to predict the migration energy E_m associated with a vacancy migration event, using as examples of I/O tables similar to the one shown in Tab. 2.1, has a sensibly different topology compared to the application discussed in section 3.3. The number of input variables is rather large, as discussed in section 2.3.2, and is of the order of several hundreds. The number of available examples of I/O, however, is virtually infinite, meaning that tables as large as required can be produced[4], with the only limitation of computing time. Fortunately, larger amounts than 100 times the number of input variables can be produced.

In this section, the application of ANNs to this problem is discussed for the prediction of the single-vacancy migration energy in binary or ternary alloys. The trained networks are used later in chapter 4 for the simulation of thermal annealing experiments.

3.4.1 Design of training and reference sets to ensure high extrapolation and interpolation skills

The notion of extrapolation or interpolation for new cases has here no real meaning from the numerical point of view, because the inputs are categorical integers. New cases therefore correspond to a never seen combination of these integers. The only required precaution, when designing a training and reference set, is therefore to ensure that every category is, ideally, equally represented for every input variables.

Oppositely, extrapolation and interpolation do have a meaning from a physical point of view, because new cases can correspond to never seen categories of atomic configurations that exhibit new features of, for example, stress field or chemical interactions. The design of proper table of I/O is therefore mainly a problem of gathering a set of LAE's that are representative of all situations that can be encountered during the AKMC simulation. Discussions in Ref. (35) (Paper IV in appendix E) lead to the conclusion that, for the simulation of thermal annealing experiments, two kinds of random examples should be generated in equal proportion, to represent extreme situations that can be encountered by the vacancy:

1. Solute atoms are randomly distributed around the migrating vacancy, with a homogeneous spatial distribution. These cases correspond to the beginning of the experiment, when the alloy is still a solid solution, while at equilibrium

---

[4]In reality the number of possible LAE's is countable, but is so large that it can be considered as infinite from any practical point of view.

they are representative of alloy compositions below the solubility limit. The local concentration of solute atoms should not, however, be constrained to respect the solubility limit value, because the local concentration around the migrating pair in the course of the AKMC simulation changes and can in fact take any value in principle. So, different local concentrations should be explored.

2. Clusters of solute atoms are created and the vacancy is randomly located around them, or inside them. ANN trained with such tables of I/O is, therefore, a priori able to make relevant predictions in any encountered configurations. Training can thus be performed on the basis of randomly generated examples, and the obtained ANN can be used during the complete AKMC simulation without restriction.

Last, the definition of training and reference sets from the table of I/O, is for this application relatively easy. Regarding the fact that the number of examples is large, and that they can be considered as independent of each others, a simple shuffle-and-cut procedure is totally appropriate.

3.4.2 Specialized training algorithm

ANN trainings on fixed architectures, using the LM algorithm, require in practice rather long computing times, because of the large number of input variables, and the large number of examples of I/O used. The number of training epochs before the onset of overspecialisation is not properly speaking large (typically around 150 or 200 at the most), but the time required to compute each epochs is. Constructive approaches, in particular the dynamic node creation of Ash, are thus a priori the only viable approach.

As discussed in section 2.3.2, energy barriers are calculated taking within long-range LAE's, that is beyond two times the cut-off of the used IAP. Including all these neighbouring atomic sites corresponds to a number of several hundreds of input variables. It is therefore reasonable to anticipate the fact that ANN can probably not efficiently take into account all these inputs, especially considering the fact that they do not all have the same degree of influence on the migration energy. ANN training is thus, a priori, equally efficient if the LAE is redefined to a shorter range, i.e. if the information about the furthest away atomic sites is discarded.

In order to propose a totally automatized training procedure, that not only adds nodes in the hidden layer similarly to the dynamic node creation, but also self-determines the maximum range of the LAE that can efficiently be taken into account, an original training algorithm, named *gradually improving accuracy constructive algorithm* (GIACA) was elaborated and completely described in Ref. (35) (Paper

IV in appendix E). Training proceeds in several stages, were entries of the LAE that correspond to increasing shells of close neighbours of the migrating pair are gradually connected to the network. The computing time spent for training is thus mainly focussed on the most influencing input variables, and training is interrupted when no further progress is achieved.

Finally, the LM algorithm, on which the GIACA is based, could be efficiently parallelized using the open-MPI library (202; 203), achieving a speed-up factor of 6 on a 8-processors personal computer.

3.4.3 Application to binary problems

Fig. 3.4 shows the quality of prediction, after training with the GIACA algorithm, for the vacancy migration energy in binary alloys. For the Fe-Cu system, 60000 examples of I/O were calculated with the NEB method, using Pasianot's IAP as cohesive model (140). The same amount of examples was calculated for the Fe-Cr system using Olsson's IAP (128). In both cases, a separate ANN was trained depending on the chemical nature of the migrating atom. The average error of prediction was calculated as:

$$\bar{e} = \frac{1}{R} \sum_{i=1}^{R} |d_i - o_i| \tag{3.7}$$

Here, R is the number of examples in the reference set, d_i is the desired ANN output for the example i, i.e. the NEB calculated energy barrier, and o_i is the actual ANN output. We see that the ANN predictions are, in general, very accurate, especially for the Fe-Cu system. The error is somehow larger for the Fe-Cr system, which is explainable by the fact that Olsson's IAP is more complex compared to Pasianot's, because of the presence of an extra embedding function accounting for the s band of electrons. Predictions are in any case not only highly correlated with the NEB values ($R^2 > 0.98$), but are in fact accurate for all individual cases. Thus, ANNs can be considered, in fact, as near-perfect substitutes to NEB, in these cases.

3.4.4 Application to ternary or more complex problems

Fig. 3.5 shows the quality of prediction, after training with the GIACA algorithm, for the vacancy migration energy in the ternary Fe-Ni-Cu alloy. Migration energies were calculated using the potential developed by Bonny et al. in Ref. (25). In this case, less atomic sites were efficiently connected to the network, and the final error of prediction committed on the reference is larger compared to the achievements for binary alloys. This is explained by the increased complexity of the regression problem from a numerical point of view, because each entries of the LAE can take

Figure 3.4: ANN quality of prediction of the vacancy migration energy in binary Fe-Cu and Fe-Cr alloys. In both cases, the GIACA algorithm converged after connection to the network of the LAE entries encompassing the 11nn of the vacancy and migrating atom (223 atomic sites), and the addition of 20 to 25 nodes in the hidden layer.

three different integers: "1" for Fe, "2" for Ni and "3" for Cu. One can imagine that the development of a relevant behaviour in nodes of the network is less easy, because a clear distinction between the different categories of atoms is not straightforward to make. For this reason, it is recommended (see for example (16)) to reformulate the input variables in a different way, using bits instead of integers. This formulation is named 1-of-c coding, and requires a number of bits that is equal to the number of categories minus one. For our example here, each entry of the LAE is thus coded with two bits: Fe is coded with "00", Ni with "10" and finally Cu with "01". We see, in the right part of Fig. 3.5, that the global quality of predictions is improved, despite the fact that the number of input variables is doubled.

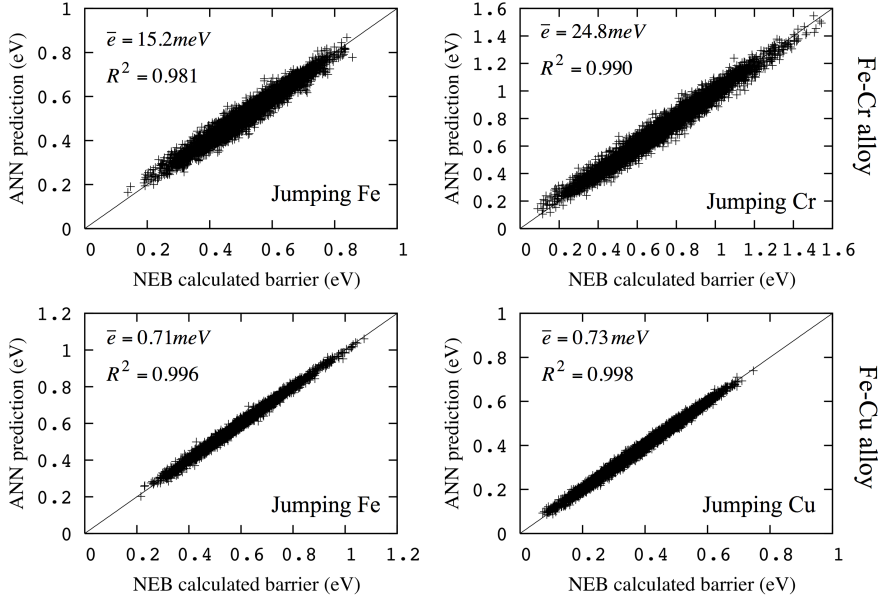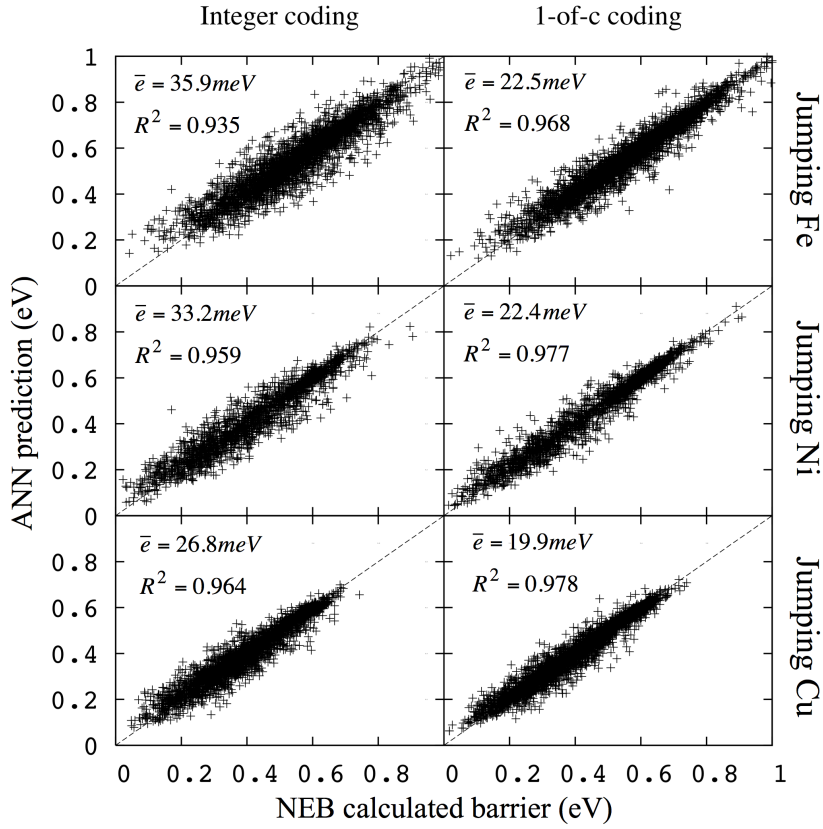**Figure 3.5:** ANN quality of prediction of the vacancy migration energy in ternary Fe-Ni-Cu alloys. The GIACA algorithm converged after connection to the network of the LAE entries encompassing the 6nn of the vacancy and migrating atom (83 atomic sites), and the addition of 30 nodes in the hidden layer.

4 Simulation of thermal ageing experiments

Thermal annealing experiments (TAE) are typically conducted on binary or ternary alloys to study their thermodynamic properties: to identify the phases that appear at equilibrium, and also to study the kinetics of the relevant thermodynamic transformations. Thermal ageing experiments are often also conducted on complex alloys and steels used as structural materials, to study their thermal stability under conditions that either mimic or accelerate operational conditions. A thermodynamic process that is of special concern for applications is the precipitation of new phases during prolonged exposure to high temperature, because these precipitates will generally make the material harder and less ductile. Here we are specifically interested in steels used or to be used as structural materials in nuclear installations, which will be exposed during operation to both high temperature and neutron irradiation fields. The behaviour of these steels will be studied by choosing suitable model alloys. Therefore, two model alloys are considered:

- Fe-Ni-Cu, as model alloy for bainitic ferritic reactor pressure vessel steels. In these steels, the formation of Cu-rich precipitates is one of the most important reasons for the deterioration of their mechanical properties (5; 77; 126), causing hardening and embrittlement. Copper, therefore, is unquestionably the most influencing chemical element in this process, and its precipitation in α-Fe has been intensively studied during the last decades, both experimentally (7; 30; 66; 67; 86; 101; 109; 114; 115; 142; 178; 189), and using theoretical models (43; 44; 64; 65; 109; 131; 132) or computer simulations (13; 14; 18; 34; 48; 91; 97; 158; 160; 162–164; 176–178). Nickel is considered as the second most influencing elements by many specialists (5; 126).

- Fe-Cr, as model alloy for high Cr (9-12 at.% Cr) ferritic-martensitic stainless steels, that are candidate structural materials for many concepts of new

generation nuclear reactors. The presence of Cr has many positive effects, amongst which the reduction of radiation-induced swelling (63; 90; 99; 100; 146) and possibly even reduction of irradiation embrittlement (88; 89). Steels of this type containing more than 14at%Cr undergo phase separation after thermal annealing at about 475°C (19; 51; 59; 68; 71; 82; 95; 111; 122; 168; 185; 186), with the formation of finely dispersed nano-metric nearly pure Cr precipitates coherent with the matrix. The latter cause hardening and provoke embrittlement of these steels. Irradiation accelerates this process and makes precipitates appear also at concentrations well below 14%Cr.

The TAE proceeds as follows: (1) Samples of the material are annealed at high temperature, in the region of complete solubility, in such a way that the insoluble alloying elements are distributed as a random solid solution; (2) the material is cooled down to room temperature, where diffusion is too slow to produce any redistribution of atoms towards the thermodynamic equilibrium and therefore remains a random solid solution; (3) the actual thermal ageing experiment consists in exposing the material at a higher temperature, that is kept constant for a long time (hours, days, even months or years). Samples are taken out of the oven at different times, then examined with an appropriate technique to look at the advancement of precipitation, typically in terms of average cluster radius, clusters density, and also content of insoluble species in the matrix. Several measurement techniques exist, amongst which, atom probe tomography, small angle neutron scattering, small angle X-rays scattering, electron transmission microscopy, resistivity measurements,...

The simulation of TAE is an interesting exercise for the multi-scale modelling community. At least three reasons could be evoked, that are equally important and in fact intimately linked with each others:

- The simulation of these processes can provide us with a better understanding of the clustering and precipitation mechanisms. The essential difference is that, in computer simulations, a very precise description of the system being studied is naturally available. A quantity of detailed information, that is normally inaccessible experimentally, can be collected, being obviously limited, however, by the realism of the simulation method.

- TAE are a relatively simple problem to model, because, differently for example from irradiation experiments, only one point-defect is introduced in the materials to be studied. This is therefore a first test case for tools developed with the purpose of simulating the effects of irradiation. In particular, this test allows the energy models used as physical input to be tested: for methods based on ab initio data, the formalism used to calculate the vacancy migration energy, and in our ANN-based AKMC, the quality of the interatomic potential used.

- Last, but not least, TAE are considerably less expensive to conduct than irradiation experiments: more data are therefore available in the literature. Direct comparison of results of the simulation with larger amounts of experimental data becomes then possible, concerning the kinetics of precipitation of insoluble species.

The ANN-based AKMC described in section 2.3 has been successfully applied to simulate TAE experiments for Fe-Cr alloys and Fe-Cu alloys. This work has been published in full length papers: Ref. (35; 36) (Papers IV and V in Appendix E and F) and Ref. (38) (Paper VII in Appendix H), respectively. In this chapter, the methodology that was used and the obtained results are summarized. The interested reader is referred to the corresponding papers for a more detailed discussion. Simulations conducted in Fe-Cr alloys are the subject of section 4.2, while section 4.3 concerns Fe-Cu alloys. The chapter starts with a first section, 4.1, describing in detail the general methodology used for the simulation of TAEs.

4.1 Simulation details

The process of clustering and precipitation of insoluble species can occur in different places of the material: in the bulk, at grain boundaries, at interfaces between different materials, in the vicinity of dislocations,... Depending on the chemical composition of the system being studied, some of these places can be more or less interesting to study. In this chapter, only bulk materials are considered, without extended defects such as dislocations or grain boundaries. The simulation by AKMC of systems containing extended defects is more delicate, because of the loss of periodicity of the atomic structure, which complicates the definition of the algorithm in many ways: mainly the definition of atomic volumes and of migration events, as discussed in section 2.2.

Bulk materials are modelled as infinite grids of atoms, arranged in space according to an appropriate crystallographic structure (bcc in this work): the simulation is conducted on a cubic sample of the infinite material, and periodic boundary conditions are applied to emulate infinity of the volume. As the material being studied is unirradiated, the only mobile point-defects introduced in the simulation box are vacancies. The equilibrium concentration of vacancies in real materials, however, is not accurately know, neither experimentally, nor by first principles calculations. In theory, this concentration is given by:

$$C_v^{real} = A \cdot \exp\left(\frac{-H_v^f}{k_B \cdot T}\right)$$ (4.1)

Here, H_v^f is the vacancy formation enthalpy, k_B is Boltzmann's constant, T is the absolute temperature and A is a constant. Both A and H_v^f depend on the alloy composition. The exact value of the vacancy formation enthalpy is delicate to estimate, either experimentally or theoretically. For pure Fe, first principle calculations situate it between 1.9 eV and 2.2 eV, while experimental measurements suggest values around 1.6-1.7 eV. Assuming for the sake of simplicity $A = 1$ and $T = 773$ K, we find that $C_v^{real} \simeq 10^{-15} \cdots 10^{-11}$. In other words, the AKMC simulation box should be large enough to contain from 10^{11} to 10^{15} atoms if only one vacancy was introduced. Simulations, however, cannot practically be performed in such a huge box, mainly for reasons of CPU time limitations. So, the box is always much smaller, and one vacancy is introduced anyway. Thus, the effective concentration of vacancies in the simulation is orders of magnitude higher than the real one and the atomic redistribution processes will proportionally occur much faster. As a compensation, the time of the simulation must be rescaled, as will be discussed in the end of this section. Optimal choice of the box size is a trade-off between affordable CPU time and statistics that can be eventually achieved concerning precipitate size distributions.

The alloy is initially a random solid solution: all positions in the box are occupied by matrix atoms (Fe in this chapter), and solute atoms are placed at random, with the appropriate number to respect the alloy composition, as well as one vacancy. The only event is the migration of the vacancy to one of its eight 1nn, as depicted in Fig. 2.6. Other migration events could be imagined, for example the migration to a further away distance than the 1nn position. This and other possibilities have been considered, but they have been found not relevant, because of the significantly higher migration energies associated with these jumps. The jump frequencies are calculated using Eq. 2.5. The jump attempt frequency is taken to be a constant, in first approximation, and an ANN is trained to predict the migration energy that would be calculated with the NEB band method, using the relevant IAP as cohesive model.

The simulation time t_{MC} must be rescaled into the real time t_{real}, to take into account the difference between the vacancy concentration in the simulation box $C_v^{(MC)}$ and the concentration in the real material given in Eq. 4.1. The simplest way to do so is to consider that the vacancy concentration in the box is simply given by the ratio between the number of vacancies in the box (one) and the number N_{at} of atomic sites in the box:

$$C_v^{(MC)} = \frac{1}{N_{at}} \tag{4.2}$$

This simple relation, however, does not take into account the fact that the vacancy formation energy, in principle, varies during the evolution of the system, as discussed by Le Bouar and Soisson in Ref. (97). These authors calculated $C_v^{(MC)}$ in a different way for the simulation of TAE in Fe-Cu alloys: they suggested that the

vacancy concentration in the box, $C_v^{(MC)}$, should be calculated taking into account the evolution of the formation energy during the experiment:

$$C_v^{(MC)} = \frac{f_v}{N_{at} \cdot X_{Fe}} \tag{4.3}$$

Here, f_v is the fraction of time spent by the vacancy in a pure Fe environment (up to the 2nn distance), and X_{Fe} is the proportion of Fe atoms in the box: $X_{Fe} \simeq 1$. f_v is directly measured during the simulation: in the beginning, when the alloy is still a random solid solution, $f_v \simeq 0.5 \ldots 1$, but rapidly decreases by several orders of magnitude as soon as clusters of Cu atoms are formed.

Other authors didn't obtain satisfactory results using this relation. The most probable reason is the relative uncertainty on several parameters involved at different levels of the simulation. One example is, as mentioned, the exact numerical value of the vacancy formation enthalpy H_v^f, but also the exponential prefactor A in Eq 4.1. Another uncertainty concerns the exact value of the vacancy jump attempt frequency Γ_0 in Eq. 2.5, that acts as a global time scaling factor if assumed to be constant. For these reasons, some authors such as Vincent et al. in Ref. (178) or Bonny et al. in Ref. (23), choose the value of H_v^f in such a way that the results of their simulation fit at best data obtained experimentally. This is equivalent to choosing a priori one value of H_v^f, for example the one predicted by ab initio or by the IAP used, and apply a global correction factor f_g on the simulation time. Time is thus, finally, rescaled as:

$$t_{real} = t_{MC} \cdot \frac{C_v^{(MC)}}{C_v^{(real)}} \cdot f_g \tag{4.4}$$

The AKMC model is considered as globally correct and consistent if, for given experimental conditions, a unique value of f_g can be found, in such a way that the AKMC predictions are superposed with all the different experimental data, and the value of f_g should, ideally, be as close as possible to 1.

4.2 Results for Fe-Cr alloys

The simulation of TAE in the Fe-Cr system is especially suitable to be performed with the ANN-based AKMC, for two reasons:

- Clusters of Cr atoms are coherent with the bcc matrix, so the description of the system as a bcc rigid lattice is always consistent with the real crystallographic structure of the material.
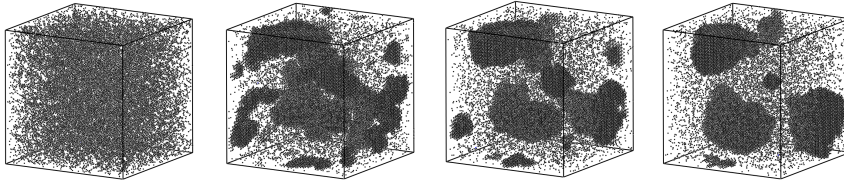
Figure 4.1: Snapshots of the AKMC simulation box for the simulation of a TAE with an Fe-20%Cr alloy at 500°C. Black dots, in these images, represent the Cr atoms. The left image is the initial state, that corresponds to a random solid solution, and the other images represent later stages of the simulation in a chronological order.

- The vacancy is not strongly attracted by Cr, or Cr clusters once they are formed. The simulation is thus never slowed down by the occurrence of vacancy migration events that do not contribute to the re-distribution of atoms in the system, typical of situations in which the vacancy is trapped by clusters of solute atoms..

A TAE in a Fe-20%Cr alloy at 500°C was simulated in a cubic box composed of 40x40x40 unit cells (128,000 atoms). The vacancy migration energies were calculated using the ANN shown in Fig. 3.4, trained on the basis of NEB calculated barriers using Olsson's IAP (128). A number of $1.3 \cdot 10^{10}$ AKMC events were computed in, roughly, 2 CPU months with a standard workstation. Fig. 4.1 shows a couple of snapshots of the simulation box. For this concentrated alloy, where the formed clusters are not compact and not pure (29; 50; 59; 71; 82; 93; 95; 110–112; 185; 186), and finally are surrounded by a depleted matrix that keeps a relatively high concentration of Cr (3; 21; 92; 107; 122) ,the identification of clusters of Cr atoms was performed with the algorithm proposed by Bonny et al. in Ref. (21): an atomic site is considered to belong to a cluster if the local concentration of Cr is larger than 95%. At the peak of clusters density, the box contained 60 Cr-rich clusters, whereas only 3 clusters remained when the simulation was stopped. Enough statistics were thus collected for the nucleation and growth stages of the TAE, as well as for the first phase of coarsening, whereas the simulation should be continued in a larger box to properly account for the final phase of coarsening.

Results obtained with the AKMC simulation are compared with experimental data in Fig. 4.2. The ANN-based approach is also compared with another AKMC algorithm, where the vacancy migration energy is calculated with a Kang-Weinberg (KW) approach (see Eq. 2.6). This simulation was also performed by Bonny et al. in Ref. (23; 26), though in slightly different conditions, but results are almost identical. The comparison between the results obtained with the two different AKMC algorithms allows an evaluation of the effects of static atomic relations on the global process of
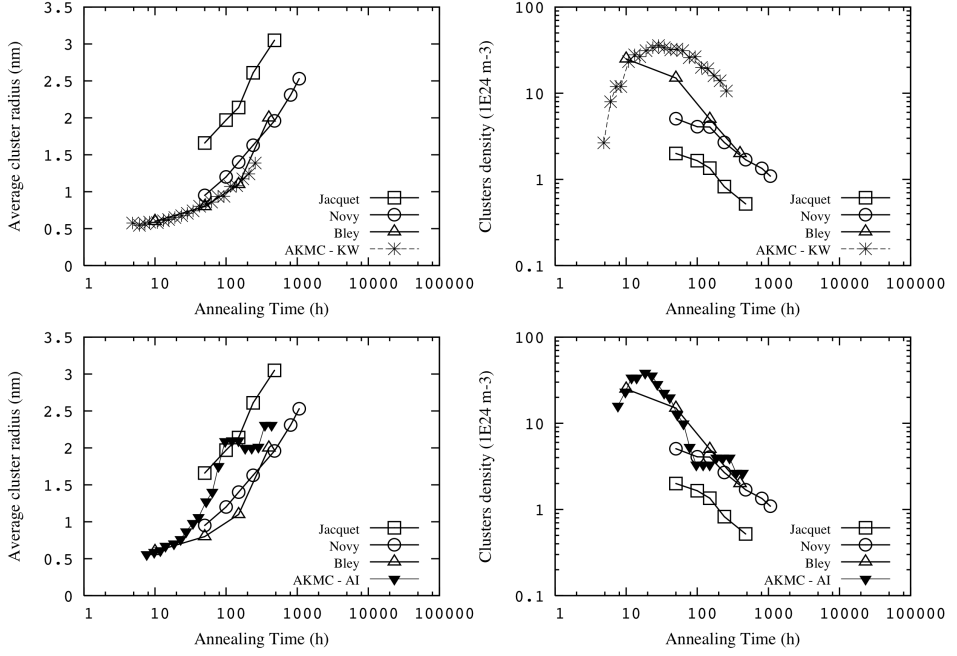
Figure 4.2: Comparison of experimental average cluster radius (left) and cluster density (right) with the results from Fe-Cr thermal annealing simulations with AKMC, using two different migration energy calculation methods: Kang-Weinberg decomposition (above) and ANN (below). The time on the top figures has been rescaled using H_v^f =1.68 eV, whereas H_v^f =1.80 eV has been used for the bottom ones. The experimental values are taken from Refs. (19; 81; 122)

precipitation, that are completely neglected in the KW approach (see section 2.2.3). In Ref. (36), these results are also compared with an AKMC model where the vacancy migration energy is calculated with a cluster expansion model (96), fitted using ab initio data. The results shown in Fig. 4.2 can be so summarized:

- When using the KW driven AKMC, the choice of $H_v^f \simeq 1.5$ eV fits Jacquet's experiments (81), but $H_v^f \simeq 1.68$ eV is required to fit Bley's (19) and Novy's (122) (this is the case shown in the upper panel of Fig. 4.2). However, for either choice, while the predicted average cluster radius is in very good agreement with the experiments, the final cluster density is overestimated by almost an order of magnitude.

- When using the ANN-based AKMC, $H_v^f \simeq 1.8$ eV fits the first point of

all experimental works. It is very close to the numerical value predicted by the potential: $H_v^f \simeq 1.81$ eV for Fe-20%Cr alloys, using the method described in Ref. (170). The rate of increase of the average cluster radius is somewhat larger than in the experiments, but the results can be considered in very good agreement with the experiments, especially considering the discrepancy existing between experimental data-sets. In addition, the cluster density is in much better agreement with the experiments than using the KW decomposition. Note that the AKMC curves exhibit a jerky shape, because the number of clusters is very small, varying between 3, 4 and 5, at the end of the simulation. To have better statistics and a smoother curve, a larger simulation box should be used, but the computational cost associated with an increased size becomes prohibitive.

Nevertheless, not all results of the simulation compare one-to-one with the experiments. Fig. 4.3 shows for example the evolution versus time of the Cr concentration in the matrix and in the clusters from experiments and according to the ANN-based AKMC. We see that Cr depletion in the matrix is faster in the simulation than in the experiment (81; 122). Consistently, the build up of the equilibrium concentration of Cr in the precipitates in the simulation is totally different from the experimentally observed one and much faster. Essentially, in the simulation Cr clusters are created in thermodynamic equilibrium since the beginning, this implying that the matrix is rapidly depleted and the only re-adjustement is the emission of some Cr atoms. On the contrary, the experiments suggest that precipitates are initially diffuse and only at the end approach the equilibrium concentration. However, this discrepancy may also partly be due to different definitions of clusters in the experimental analysis as compared to the simulation, as well as to limitations in the precision of the experimental technique (atom probe tomography in this case).

Finally, Fig. 4.4 shows data points for the Cr solubility limit in Fe, as predicted by various MC schemes, all based on the use of Olsson's interatomic potential as Hamiltonian. The experimental phase boundary, as reviewed by Bonny et al. in Ref. (22), is also shown. Metropolis MC methods (17; 192) are unable to provide the kinetics of the diffusion process leading to precipitation. However, they can correctly account for all contributions to the free energy of the system, as stemming from the used Hamiltonian, including the vibrational contribution. They are therefore more suitable and reliable to trace the phase diagram embodied by a given Hamiltonian than AKMC models. In addition, if the possibility of displacing atoms off lattice is switched off in the Metropolis MC, the vibrational and strain-field contributions to the free energy are switched off too, and the "rigid lattice phase diagram" can be thereby obtained. Both Metropolis MC options have been used to compute the points in Fig. 4.4. It can be seen that the solubility limit data points that were predicted by the ANN-based AKMC simulations are consistent with the Metropolis MC results. This shows that the good thermodynamic description of the Fe–Cr
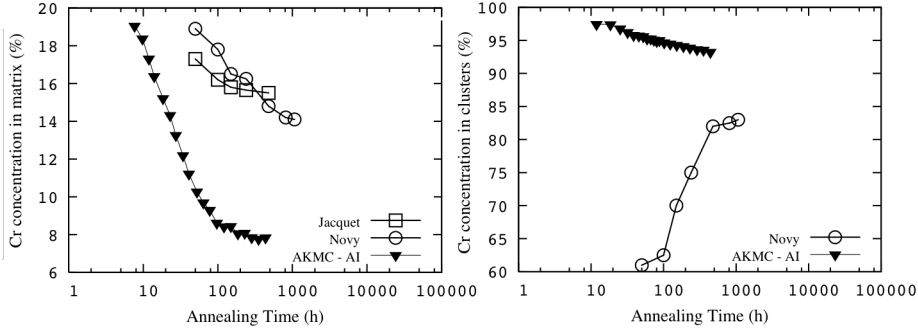
Figure 4.3: Comparison of atom probe tomography (122) and small angle neutron scattering (81) study of the experimental Cr concentration in the matrix and in the clusters with the Fe-Cr thermal annealing simulations performed with the ANN-based AKMC.

system provided by Olsson's potential is transferred almost untouched, via ANN regression of migration energy barriers, to the AKMC model. The fact that the ANN-based AKMC data points lie between those obtained by Metropolis MC in a rigid lattice and in an off-lattice approximation can be explained because the ANN-based AKMC allows for relaxation effects, but cannot allow for vibrational entropy effects, visible especially at high temperature. We therefore conclude that our AKMC model reproduces well the thermodynamic properties of the Fe–Cr system.

4.3 Fe-Cu alloy

The simulation of TAE in Fe-Cu alloys have represented and still represent a challenge that many authors have tried to face, with different degrees of success. One of the origins of the difficulty of simulating Cu precipitation in iron, especially when using AKMC methods, is that pure Cu has an fcc crystallographic structure, whereas ferritic Fe has a bcc structure. It is experimentally found (130) (145) that Cu clusters are, at the beginning of the experiment, coherent with the Fe matrix, up to a diameter between 4nm and 5nm. Assuming spherical clusters, this corresponds to, roughly, 6000 atoms at the most. The lattice parameter of these clusters is also almost identical to the one of Fe: about 2.86-2.87 Å. Growing further, they first take intermediate structures (9R, then 3R), to finally become fcc from about 12nm (142). This finding was also confirmed using molecular dynamics simulations, which moreover suggested that the stability of Cu precipitates is enhanced by the presence of vacancies inside (18). As a consequence, in the framework of the rigid lattice description of the system proper to AKMC, only the coherent stage of precipitation can be simulated, i.e. as long as Cu clusters are smaller than 5nm. The next stages
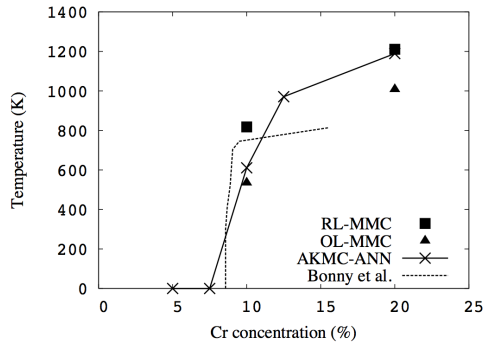
Figure 4.4: Fe-Cr phase diagram, in the low Cr concentrations range, as predicted with different atomistic Monte Carlo methods: rigid lattice Metropolis Monte Carlo (RL-MMC), off-lattice Metropolis Monte Carlo (OL-MMC), and ANN-based AKMC. The method followed to generate this diagram, as well as the MMC predicted phase diagrams, are described in Ref. (192). The dashed line is a recent revision of the experimental diagram (22).

can only be simulated if the AKMC algorithm is modified to take crystallographic changes into account, but this is not straightfoward, as discussed in section 2.2.

The interested reader can find a critical review of most of previous attempts to simulate TAE with AKMC methods in the paper of Vincent et al. (178). In these approaches, the vacancy migration energy was always calculated on a rigid lattice, using a KW or broken-bonds formalism, generally fitted on the basis of energy data calculated with ab initio. The conclusion by Vincent et al. is, however, that none of the models reviewed can consistently predict Cu precipitation in Fe: depending on the parametrization, either the average precipitate radius, or the precipitate density is correctly predicted, but hardly ever both. In particular, most models tend to overestimate the density.

Soisson and Fu (163) achieved a more satisfactory and consistent prediction. In their AKMC model, the vacancy migration energy was also calculated using a broken-bonds formalism, fitted to energy data calculated by density functional theory methods. However, they were able to predict accurately the evolution with annealing time of both the average precipitate radius and the precipitate density. The main reason for their success is probably that their model, by incorporating a very strong interaction between vacancies and Cu clusters, predicted that Cu clusters of all sizes, including large ones (i.e. full coherent precipitates), were mobile, as a consequence of complex series of vacancy hops at their surface or near the surface. They also observed direct coalescence of clusters, as the result of migration. Their results therefore suggested that the mobility of Cu clusters, as a consequence of strong interactions with vacancies, can play a non-negligible role in the kinetics of Cu precipitation in Fe. Mobility of clusters was also observed in previous AKMC

simulations, though limited to the smallest clusters (97; 160). The diffusion of Cu clusters, on the other hand, had never been considered as mechanism in classical diffusion theory models for Cu precipitation (43; 64; 65; 109). The main limitation of Soisson and Fu's model was a "collateral effect" of the strong binding between Cu clusters and vacancy: the latter remained trapped in the bulk of these clusters for a very large fraction of the simulation time, thereby increasing enormously the CPU cost and therefore limiting drastically the reach of the simulation, which had to be stopped at a very early stage of the coherent precipitation.

In this thesis, TAE were simulated using the ANN-based AKMC algorithm described in section 2.3. The essential differences with previous works is therefore that an IAP is used as cohesive model, but the vacancy migration energy is not calculated on a rigid lattice. The IAP used was developed by Pasianot et al. in Ref. (140). This potential was designed to fulfil two major objectives: (1) to be consistent with thermodynamics, by providing a correct prediction of the experimental Fe-Cu phase diagram; (2) to provide an accurate description of the interaction between point-defects and solute atoms, which is a necessary quality for the simulation of neutron irradiation effects in steels. The potential has been proven to be able to predict the correct final equilibrium for a thermal annealing experiment (34; 140), although its ability to fully predict the kinetic path from a random solid solution has not been evaluated yet. This was done in the work described in this section.

A first attempt to simulate TAE experiments on Fe-Cu alloys with the ANN-based AKMC was published in Ref. (34) (Paper III in Appendix D). Due to the negligible solubility of Fe in Cu, which is properly reflected by the interatomic potential used, the Cu clusters formed during the simulation are from the beginning completely pure and compact, similarly to the findings by other authors in previous AKMC simulations (97; 163; 178). Note that the validity of this finding is still a debate in the community, because Cu clusters are experimentally found to contain non-negligible amounts of Fe, also without neutron irradiation (see e.g. Ref. (58; 117)). Also consistently with previous work, in a full AKMC simulation with this method, the vacancy is strongly attracted by Cu clusters, and remains trapped inside them for a very large fraction of the AKMC events. The speed of the simulation is, therefore, significantly slowed down (6; 34; 163; 178). For this reason, the ANN-based AKMC algorithm has been modified to allow explicitly for the mobility of Cu clusters of all sizes, while reducing drastically the CPU cost of the simulation, as compared to a full AKMC simulation. This was achieved by combining the AKMC approach with a coarse-grained approach, of object kinetic Monte Carlo (OKMC) type (48). As a matter of fact, the AKMC is a priori the ideal approach, to describe atomic diffusion-driven processes, because this algorithm does not make any assumption on the coarser-scale mechanisms that determine the phenomenon that is being simulated: these stem spontaneously from the physical approximation used to calculate the migration energies and the underlying thermodynamic description of the system. As mentioned, however, in a fully atomistic model the simulation can be tremendously

slowed down, in terms of CPU time, in case of trapping of vacancies. We therefore treated Cu clusters above a certain size as objects, defined by size and centre-of-mass position, neglecting the details of their atomic configuration and therefore of the hops performed by the single vacancy on their surface. Instead, these objects are allowed to diffuse, or dissociate by emission of a single vacancy or a Cu-vacancy pair, based on specific, size-dependent and thermally-activated frequencies, in much the same way as in OKMC models (48). Also similarly to OKMC, Cu clusters are allowed to coalesce if two of them meet during migration. Seamless matching between the fully-atomistic model used to describe small Cu clusters and the coarser-grain model used to describe larger Cu clusters is guaranteed by calculating the diffusion coefficients and emission probabilities for the object-like clusters based on specific, full AKMC simulations, on which another ANN has been trained. In this way, the parameters of the model for large clusters are calculated in a way that is fully consistent with the atomistic description. We show in what follows that this "hybrid" AKMC approach proves both computationally efficient and physically very accurate, thanks also to the high quality of the interatomic potential, which is here exploited in the most complete way possible.

The remaining part of the section is organized in the following way. In section 4.3.1, the fundamentals of the hybrid AKMC are briefly described. In section 4.3.2, this original algorithm is applied to the simulation of several TAE in Fe-Cu, for different alloy compositions and temperatures. Results are compared with sets of experimental data. Finally, in section 4.3.3, the mechanism of Cu precipitation that stems out of the model is analysed, and compared with mechanisms observed by other authors.

4.3.1 Hybrid atomistic kinetic Monte Carlo algorithm

A complete description of the hybrid AKMC algorithm has been published in a full length paper: Ref. (38) (Paper VII in Appendix H). Here, only the fundamentals of this new algorithm are briefly summarized:

- At the beginning of the simulation, the alloy is a random solid solution, and no Cu objects are defined. The regular ANN-based AKMC algorithm is applied without modification. The ANN quality of prediction is shown in Fig. 3.4.

- At regular steps, the simulation box is analysed, and the list of clusters of Cu atoms is made. Clusters are defined as groups of atoms that are linked by 1nn or 2nn bonds. Clusters bigger than a predefined minimum size are re-shaped into a perfect sphere, and considered as objects for the rest of the simulation.

- The minimum size to define cluster objects was chosen to be a number of 15 atoms, because it corresponds, in bcc, to a central atom surrounded by other

atoms filling completely its first and second shells of close neighbours (so it is very stable), thereby being the smallest size describable as a sphere, and it is larger than the critical size for nucleation of Cu clusters in our AKMC, at least for temperatures lower than 700°C.

- Once the vacancy enters in contact with one atom of a cluster object, new events are defined in replacement of the vacancy migration events. These events are briefly described below. The speed of the algorithm is tremendously enhanced, because the new defined events are very simple to apply, and are defined in such a way that they approximate at best the complex series of events that would, in regular AKMC, take place.

- This algorithm is therefore named hybrid AKMC, because the rigorousness of the AKMC algorithm is globally kept, except for events concerning big Cu clusters.

The new events defined for Cu clusters, are the following:

Dissociation: The cluster dissociates most of the time by the emission of the sole vacancy. In accordance to existing OKMC algorithms, a second dissociation mechanism, via the emission of a VCu pair, was also defined, as a simplification of the Ostwald ripening mechanism for coarsening: It is the only mechanism that can result to a complete dissolution of the cluster.

Migration: The whole cluster migrates to a 1nn position. This process is observed in the ANN-based AKMC, as the consequence of complex series of vacancy jumps at the surface of the clusters, similarly to Soisson and Fu findings with their own model (163).

Frequencies were calculated for the two events, for all coherent clusters sizes (up to a diameter of 5nm), with series of independent AKMC simulations. Following the procedure described in Ref. (138) and Ref. (139) (Paper VIII in Appendix I), diffusion coefficients and lifetimes of VCu_N clusters where calculated for $N = 15, \ldots, 6000$. These calculations had to be performed, most of the times, at high temperature, to avoid too strong trapping of the vacancy in the bulk of the cluster that would make the calculation prohibitive in term of CPU time. Extrapolation with Arrhenius plots was therefore necessary to obtain the required values at the desired temperature. Extrapolation was improved, in practice, with the help of artificial neural networks.

4.3.2 Simulation of thermal annealing experiments

Several TAE's could be simulated, up to the end of the coherent stage, with the hybrid AKMC algorithm described in section 4.3.1. Fig. 4.5 shows snapshots of the simulation box for one experiment. All results obtained are shown in Fig. 4.6.
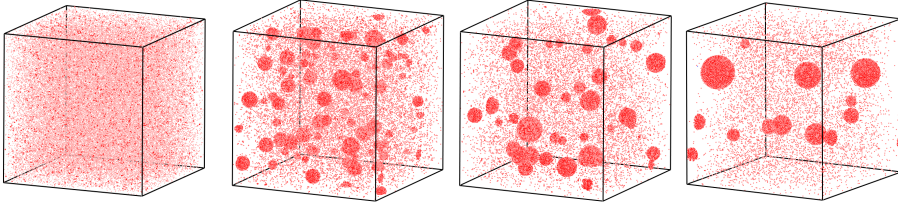
Figure 4.5: Snapshots of the hybrid AKMC simulation box for the simulation of a TAE with an Fe-1.34%Cu alloy at 500°C. Red dots, in these images, represent the Cu atoms. The left image is the initial state, that corresponds to a random solid solution, and the other images represent later stages of the simulation in a chronological order.

Cu content (at%)	T (○C)	Ref.	box size (l.u.)	Nbr. AKMC events	f_v Eq. 4.3	f_g Eq. 4.4
1.34	500	(66; 67; 86; 109; 142)	64	1.4E9	1.5E-5	0.60
1.34	500	(66; 67; 86; 109; 142)	128	4.2E9	1.5E-5	0.60
1.34	600	(142)	64	3.5E9	1E-4	0.60
1.34	700	(142)	64	6.3E9	8E-4	0.60
1.1	550	(30)	128	15.3E9	4E-5	1.30
0.6	500	(178)	128	30.2	2.7E-5	0.10

Table 4.1: Summary of TAE for Fe-Cu that were simulated with the hybrid AKMC, and of the sets of experimental data used as comparison reference.

Fig. 4.6 shows at a glance that the predictions of the hybrid AKMC model are in very good agreement with experimental data. For example, the increase of the average precipitate radius versus time in Fe-1.34at%Cu is very closely reproduced at all three temperatures investigated (500, 600 and 700°C). In the case of the experiment at 500°C the measured evolution of the density of precipitates is also provided and the model very nicely predicts nucleation (density increases), growth (density reaches a peak and remains temporarily constant while the radius keeps increasing) and coarsening (density decreases while the radius keeps increasing, because large precipitates grow at the expense of smaller ones). The curves obtained in the simulation box with sides of 64 lattice parameters are jerky because, especially when the coarsening stage is reached, only a few precipitates remain in the box and the disappearance of a small one to make a big one produces significant oscillations in the overall density. In particular, step-like increases/decreases are observed: this is a clear indication that the predicted mechanism leading to the density decrease and radius increase is the coalescence of mobile precipitates. Simulations conducted in the larger box (side of 128 lattice parameters) allow the jerks to be damped, thanks to better statistics (larger number of precipitates in the box). In all cases, the simulations finished with one single cluster in the box, of varying size, consistently with the
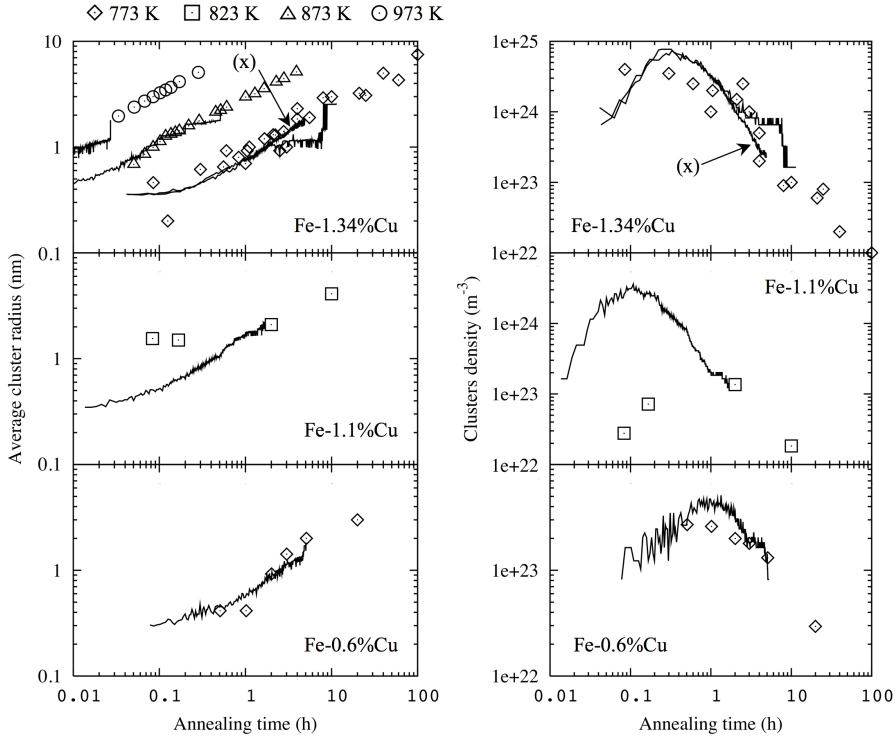
Figure 4.6: Comparison of the experimental average cluster radius (left) and cluster density (right) with the results from Fe-Cu thermal annealing simulations with the hybrid AKMC algorithm. Dots are experimental values taken from Ref. (30; 66; 67; 86; 109; 142; 178). For the case Fe-1.34%Cu at 500°C, the (X) mark indicates the simulation conducted in a 128x128x128 box.

increasing solubility limit with temperature.

The annealing of Fe-0.6at%Cu at 500°C leads to a precipitate density significantly smaller than in the above case, at all stages, consistently with the halved solute concentration: the simulation could only be meaningfully performed in a 128x128x128 unit cells box with 4 million atoms. These simulations were particularly demanding in terms of CPU time, also because the Cu concentration is lower and the acting thermodynamic force correspondingly weaker. Nonetheless, good agreement with experimental data is achieved.

Finally, in the simulation of the annealing of a Fe-1.1at%Cu alloy at 550°C we observe that the first two experimental points are not predicted by our model. The good agreement with experiments obtained in all other cases, however, gives us

63

sufficient confidence to believe that those two experimental points are probably affected by large uncertainty, possibly as a consequence of the limited resolution of the experimental technique used, i.e. small angle neutron scattering, which is not sensitive to precipitates below 1nm in diameter and therefore can be supposed to have overestimated the average size at the early stage of the precipitation, especially because in the experiment there was no support from any other complementary technique. This is confirmed by the fact that the third experimental point, still within the limit of coherent precipitation, is correctly reached by the model, and that a visual extrapolation of the curve will lead to reach the fourth point as well, even though this lies well into the regime where crystallographic transformation must have started, i.e. strictly speaking outside the range of validity of the model. Our model, as most experimental data, is roughly consistent with a dependence of the radius on a 1/2 power of time during growth that decreases to a dependence on a 1/3 power of time during coarsening, as should be expected (142; 178) (in logarithmic scale this is a roughly linear dependence, though with gradual change of slope), while no regression interpolating the four experimental points from (30) will respect such a law.

4.3.3 Analysis of the mechanism of Cu precipitation in Fe

An analysis of the results presented in section 4.3.2 strongly supports the idea that the mobility of Cu clusters and even precipitates plays a significant role in the process of precipitation in Fe, similarly to the conclusion of Soisson and Fu in Ref. (163). To highlight this conclusion, two extra simulations were conducted for a Fe-1.34at%Cu alloy at 500°C, as shown in Fig. 4.7, in which some events where deliberately prohibited, namely: in one, the emission of VCu pairs from Cu clusters was suppressed; in the other, the migration frequency was artificially modified in order to progressively inhibit the migration of the biggest clusters (more than 100 atoms). Compared to Fig. 4.6, in the former case the average precipitate radius is unsurprisingly somewhat larger and the average precipitate density almost unaltered, the results remaining in good agreement with experiments. In the latter, however, the results deviate significantly from the experimental data: from a certain annealing time on, the average precipitate radius ceases to increase, and the clusters density ceases to decrease, thus remaining higher than the experimental one. Therefore, it is only by allowing large clusters to be mobile that experimental results can be matched by the model. The one-by-one emission of VCu pairs, on the contrary, is not a sufficiently efficient mechanism to enable coarsening as observed in experiments.
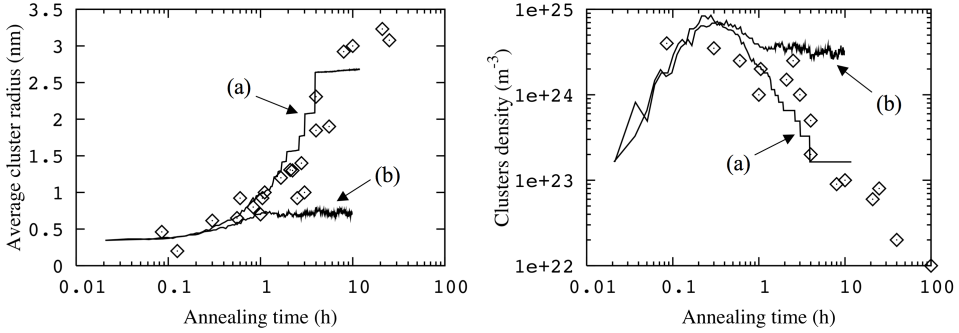
Figure 4.7: Thermal annealing experiment on a Fe-1.34at%Cu alloy at 500°C. Parameters in the model were changed compared to Fig. 4.6: (a) the probability for clusters objects to dissolve by the emission of a CuV pair is set to 0; (b) The Cu clusters migration frequency is modified to inhibit the migration of clusters bigger than 100 atoms.

4.4 Conclusions

In this chapter, the ANN-based AKMC algorithm was successfully applied to the simulation of TAEs for two representative alloys of importance for the field of nuclear material science. In both cases, good comparison with different sets of experimental data was achieved, which shows that this physical model is globally satisfactory, and that the most important mechanisms are properly taken into account. This can be attributed to two reasons: (i) the IAP used, that are the only physical input in the model, are correct from the thermodynamic viewpoint and, clearly, also adequate to describe the kinetics of precipitation processes (ii) the "filter" introduced by NEB and ANN between the predictions of the potential and those of the model is not such that the good qualities of the former are lost.

For the case of Fe-Cr alloys, a comparison with results obtained using a Kang-Weinberg approach, where relaxation effects are disregarded, highlighted for instance the importance of calculating the vacancy migration energy rigorously, that is achieved thanks to the use of artificial neural networks trained on the basis of relevant energy data.

The special case of Fe-Cu alloys was successfully tackled thanks to a "hybridization" of the AKMC algorithm with a coarse-grained model, which allowed thermal ageing experiments to be simulated in an appreciatively better way than the existing approaches used by other authors. The obtained results allowed, also, to shed some light on the mechanism of Cu precipitation in Fe, by suggesting that the mobility of Cu clusters containing one vacancy plays a central role in the precipitation mechanism, confirming the speculations put forward by other authors. This feature

65

of the process was so far never included in theoretical models, thereby possibly explaining why no fully satisfactory prediction of the kinetics of Cu precipitation in Fe could be achieved in the past.

The results obtained and presented in the chapter therefore open the way for the modelling, in the future, of either more realistic systems, or taking the effects of neutron irradiation into account. In the next chapter, the ANN-based AKMC algorithm is generalized to allow several point-defects to be introduced in the simulation, which is the necessary condition for the modelling of the effects of neutron irradiation in metals.

5

Simulations with interacting point-defects

The previous chapters focussed mainly on AKMC algorithms where only one vacancy is introduced in the system, an approach that is only applicable to the simulation of thermal annealing experiments, as in chapter 4. The simulation of irradiation damage in metals, however, requires simultaneous treatment of several interacting vacancies, as well as their counterpart, i.e. self-interstitial atoms (SIA). These point-defects and their clusters are created for example as debris after atomic collision cascades, initiated by impinging high energy neutrons or ions, in much larger quantity than the equilibrium concentration dictated by thermodynamics. In these conditions, the ageing of the material does not correspond to equilibrium states; instead, it corresponds to a non-equilibrium steady state, driven by external forces (104). SIA move very quickly, compared to vacancies, and recombine with the latter when they meet at a close enough distance, thereby decreasing the amount of mobile defects. In addition, these defects affect the diffusion of chemical species. For Fe-based metals, ab initio results indicate that P (49), Mn (175) or Cr (129) are very likely to diffuse via both SIA and vacancy migration mechanisms. On the contrary, elements such as Cu, Ni and Si, that are not strongly bound to SIA's, are likely to diffuse via vacancy migration mechanisms only, as under thermal ageing (174; 175). It is vital, if irradiation processes are to be described, that a correct treatment of both point-defects is introduced in AKMC simulations, including an acceptable description of their mutual interactions and of their individual or clusters migration properties. Since AKMC models are inherently rigid lattice approaches, however, this poses a number of methodological and practical problems that will be discussed and partially tackled in this chapter.

Section 5.1 considers the problem of the introduction of multiple vacancies, more precisely of the study of their clusters, with the purpose of calculating their diffusion properties that are a required input for OKMC simulations. In that section, the

general methodological and practical problems are discussed, and the AKMC algorithm presented in sections 2.2 and 2.3 is successfully generalized, allowing for the introduction of any number of vacancies.

The problem of SIA is discussed in section 5.2. The complexity of full generalization of the AKMC in this case is first discussed and evaluated. Then, a simplified approach to simulate isochronal annealing experiments, with a reasonably accurate description of vacancy-SIA recombination, is finally proposed.

5.1 Study of vacancy clusters using AKMC simulations

Parameters such as diffusion coefficients and lifetimes of point-defect clusters are a necessary input for OKMC simulations of irradiation processes in metals. As already discussed in section 4.3, AKMC methods are a priori the ideal approach to determine these parameters, especially in the case of vacancy clusters, the diffusivity of which cannot be studied using molecular dynamics tools. In general, they are ideal to describe atomic diffusion-driven processes. As a metter of fact, this algorithm does not make any assumption on the coarser-scale mechanisms that determine the phenomenon that is being simulated: these stem spontaneously from the physical approximation used to calculate the migration energies and the underlying thermodynamic description of the system.

The AKMC algorithm presented in chapter 2 must, however, be generalized, to allow the introduction of several vacancies in the simulated volume. Early attempts in this direction were already achieved by Pascuet et al. in Ref. (138) and Ref. (139) (Paper VIII in Appendix I), though limited to a small number of vacancies, as a first "careful" step not too far from the single-vacancy case. The methodological and practical problems encountered with the introduction of more vacancies, ideally up to any number, are discussed where the ANN-based AKMC algorithm is generalized. Next, the competitiveness of envisaged migration mechanisms is analysed in section 5.1.2, in order to determine the most important events that should be defined in the generalized AKMC algorithm. Finally, diffusion coefficients for vacancy or vacancy-copper clusters are calculated in sections 5.1.3 and 5.1.4, while section 5.1.5 shows how nickel, i.e. a third chemical element, in addition to vacancies, can also be introduced in the system.

5.1.1 AKMC Methodology

The simplest, and also most natural, way to generalize the AKMC algorithm, described in section 2.2 for the single-vacancy case, is to consider vacancies as an additional chemical species. As depicted in Fig. 5.1, the LAE associated to a
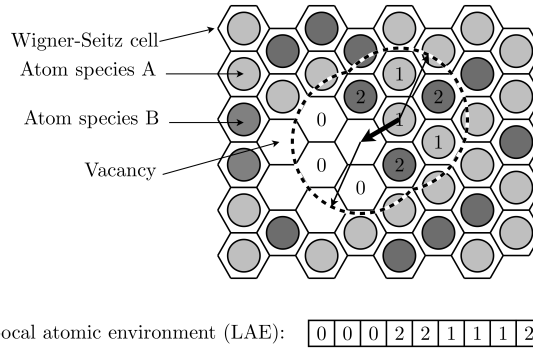
Local atomic environment (LAE): | 0 | 0 | 0 | 2 | 2 | 1 | 1 | 1 | 2 |

Figure 5.1: Coding of the local atomic environment (LAE) describing a vacancy migration event as a string of integers. Other vacancies in the LAE are described by an integer, and are therefore considered as another alloying element.

vacancy migration event is then calculated describing the other vacant sites with an integer that is not already attributed to a chemical element of the alloy ("0" in the figure). As discussed in section 2.2.2, the AKMC algorithm inherently makes the assumption that all states of the system that can be encountered are stable: any state of the system, after relaxation with a CG method, remains identical in the rigid lattice world. This condition was in practice always fulfilled for single-vacancy cases (according to the IAP used in the thesis). In principle, the introduction of several vacancies in the AKMC simulation box is straightforward as long as this condition remains fulfilled:

- Possible migration events are pre-defined, similarly to the single-vacancy case. The envisaged examples are depicted in Fig. 5.2. The migration to a first nearest neighbour (1nn) position of one individual vacancy, the others remaining immobile, is the natural extension of the single-vacancy case. Because here many vacancies are close to each others, the migration to a further away distance (2nn distance on the figure) is also possible a priori, as well as the joint migration of several vacancies at the same time: these types of event should be evaluated to decide if they must be taken into account or not.

- The NEB method can be used without modification to calculate the migration energies associated with any of the jumps depicted in Fig. 5.2. The presence of other vacancies within a close distance makes, however, the calculation more delicate in some cases, because of additional deformations of the lattice.

- An ANN can be trained to predict the migration energy, given as input the description of the LAE as a string of integers. The additional difficulty here,
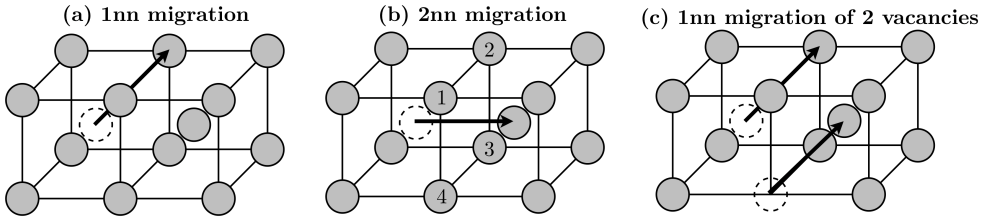
69

Figure 5.2: Examples of migration events defined in the AKMC simulation, for a bcc crystallographic structure. (a) Migration of a vacancy to a 1nn position. (b) Migration of a vacancy to a 2nn position. The condition to consider this event is that at least two atomic sites within those denoted at 1,2,3 and 4 on the figure are occupied by a vacancy. (c) Joint migration of two close vacancies to respective 1nn positions.

> compared to the single-vacancy case, is that the relationship between the LAE and the migration energy is a priori harder to fit numerically, because of the stronger deformations of the lattice provoked by the vacancies. In addition, the LAE contains an additional integer, "0" in Fig. 5.1, i.e. an additional possible value for the configurational variables, which makes the problem even harder from the numerical point of view.

The possible occurrence of unstable states can be taken into account in the algorithm, at the cost of a redefinition of some migration events, as depicted in Fig. 5.3. In this example, starting from a stable state, the final state associated with a migration event is in reality unstable: with CG relaxation, one atom, that can be a priori anywhere in the volume, spontaneously migrates to a vacant site. Ideally, these non-physical transitions should be removed from the list of events, and redefined as depicted in Fig. 5.3. One practical condition to achieve this redefinition is to detect, one way or another, during the AKMC simulation, spontaneous transitions. Preliminary studies have shown that ANNs can be efficiently trained for this purpose: these ANNs, given the LAE associated to a migration event, returns as output either the integer "1" or "0", telling if the migration is spontaneous or not, respectively. The second practical requirement is to train another ANN to predict the migration energies associated with the redefined events.

5.1.2 Competition between migration events

The following study was performed in order to evaluate the probability of complex migration events such as 2nn jumps (b) or simultaneous jumps (c), in addition to simple 1nn jumps (a). In the simulations of Cu-vacancy clusters (see section 5.1.4)
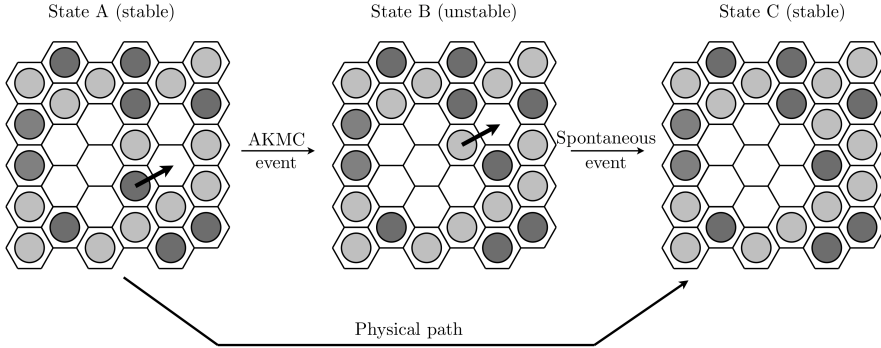
Figure 5.3: Example of redefinition of migration event in the AKMC simulation, because of the instability of states that can be encountered. State A is the initial one, assumed to be stable. Migration from state A to state B is one of the events defined during the simulation, following the pre-defined procedure. If state B is actually unstable, because for example of the spontaneous migration of one vacancy to a close neighbouring position, this migration from state A to state B does not correspond to a physical path followed by the system. It should therefore be removed from the list of events, and replaced by the migration from state A to state C.

100 configurations were randomly extracted, for 30 different cluster compositions (3000 configurations were thus studied). For each configuration, the complete list of possible events of type (a), (b) and (c) was generated and the corresponding energy barriers were calculated with the NEB method. Type (a) and type (b) events are easily defined, and the maximum numbers of these events, removing irrelevant events were vacancies exchange their position with another vacancy, is equal to 8 times or 6 times the number of vacancies in the cluster at the most. Type (c) events, however, are tremendously more numerous, because the number of possible joint migrations of vacancies explodes if the number of vacancies increases. For this reason, the study has been limited to a few examples of possible events of this type:

- Only jumps of two vacancies to 1nn positions were considered;

- The distance between the migrating vacancies was limited to the 2nn distance;

- The distance between the migrating atoms was, also, limited to the 2nn distance.

In total, roughly 50000 energy barriers were calculated and the corresponding relative probability for events type (a), type (b) or type (c) could thus be evaluated, at different temperatures, in terms of jump frequencies (Eq. 2.5). The conclusion

was that type (a) events are clearly dominating the other events. Type (b) events have a probability to take place that varies between 2.5% and 15%, depending on temperature and cluster composition, whereas type (c) events are extremely rare, with a probability between 0.001% and 0.01%. It is therefore clear that the migration of Cu-vacancy clusters is mainly driven by the migration of individual vacancies to a 1nn position and therefore only these types of events where retained in the calculations reported in what follows. Migration jumps to a 2nn position can in some cases play a minor role, but are very unlikely to change the global migration properties in a significant way. Note also that these probabilities were computed assuming a constant attempt frequency in Eq. 2.5. Jumps to a 2nn position likely have a lower attempt frequency than jumps to a 1nn position. As a consequence, the real probability for these jumps is likely to be even smaller than 15%.

For the sake of completeness, this study has been complemented by one with a different approach. A few configurations were also used as input for the monomer method proposed by Ramunni et al. in Ref. (148), that can be regarded as an ameliorated version of Henkelman's dimer (72). As already discussed in section 2.2.3, these methods allow the energy landscape to be explored, searching for possible transitions from a given initial state. No other migration events than type (a) were found, consistently with our conclusions.

5.1.3 Vacancy clusters

In this section, we present the results of using the algorithm described in section 5.1.1 to calculate diffusion properties and lifetime of vacancy clusters in pure Fe. The first step is to generate a table of examples of vacancy migration energies versus the corresponding LAE's, that will be used to train the ANN. In this case, the only species other than Fe defining the LAE of the migrating vacancies are other vacancies. The interatomic potential for Fe used for the NEB calculations was the one proposed by Ackland et al. in Ref. (1). Two categories of LAEs were generated at random, in equal proportion, as suggested in Ref. (35) for the single-vacancy case: (i) LAEs corresponding to vacancies dispersed in the volume as a random solution; and (ii) LAEs corresponding to vacancies in clusters. This ensures that the ANN is not specialized for a specific category of configurations, and therefore develops an ability to make accurate predictions for never-seen LAE's. Two different ANNs were trained: one on LAEs containing from 3 to 21 vacancies; another on LAEs containing from 15 to 60 vacancies.

The quality of the prediction, after training using the GIACA algorithm described in section 3.4.2, is shown in Fig. 5.4. Predictions are shown for cases where vacancies are either dispersed in the volume, or in cluster. We see that ANN predictions are in general very accurate, showing that the methodology described in section 5.1.1 is applicable, even when large vacancy clusters are introduced in
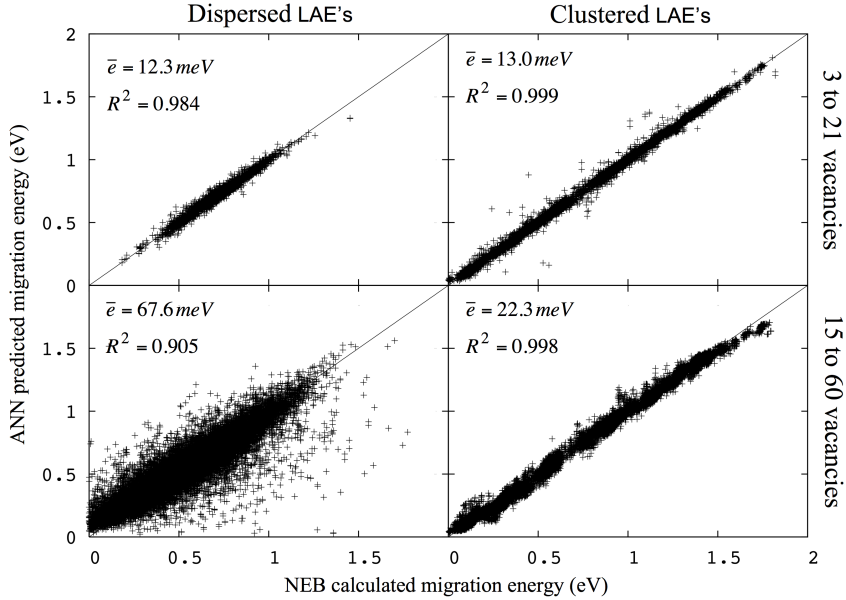
Figure 5.4: ANN quality of prediction for the vacancy migration energy in pure Fe. The ANN was trained with a database of 50,000 examples containing an equal number of configuration where vacancies are dispersed in the volume and where vacancies are in cluster. Predictions are shown for the separate cases, for clarity. The GIACA algorithm converged after connection to the ANN of the 9th nearest neighbours (173 atomic sites), and the addition of 16 hidden nodes.

the AKMC simulation. Predictions for the case of dispersed vacancies, however, are less accurate when the number of vacancies is large. This is not attributed to a failure of the methodology, but rather to two facts: (i) configurations with vacancies in clusters tend to be more easily classified into a limited number of similar families, as compared to the combinatorial number of possibilities given by random vacancy distributions; (ii) randomly generated configurations with vacancies randomly distributed are likely to correspond to very unrealistic cases, in which odd lattice distortions are produced. One example is the artificial creation of branches of atoms floating in a vacant environment. Such configurations in reality are extremely unlikely to be encountered during an AKMC simulation, because thermodynamic driving forces, correctly embedded in the model, will make nearby vacancies cluster and will give these clusters a compact shape.

In the case of clustered configurations, even for large numbers of vacancies, unstable configurations were very occasionally observed. As an order of magnitude, the proportion of unstable states, for clusters of 60 vacancies in pure Fe, was observed to
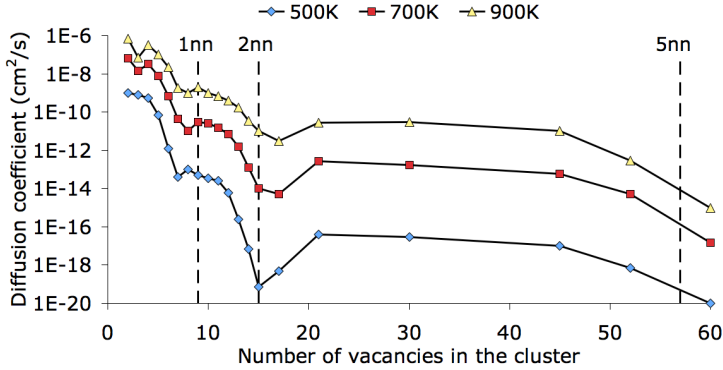
Figure 5.5: Evolution with the number of vacancies in the clusters of the diffusion coefficient calculated with AKMC simulations, at three different temperatures. Dashed lines show cluster radii that correspond to shells of neighbours (first nearest neighbours 1nn, 2nn and 5nn) that are completely filled by vacancies if the shape of the cluster is spherical. Values for clusters of 2, 3, 4, 5 or 6 vacancies are taken from Ref. (138; 139)

be around 5% at the most. These configurations, if encountered, were discarded from the ANN training table, and the occurrence of spontaneous migration events, that in principle requires the redefinition of the pre-defined migration events as depicted in Fig. 5.3 can thereby be ignored. Ideally, a post-analysis should be performed after every AKMC simulation to verify the validity of this hypothesis: states of the simulation box are randomly extracted and relaxed with CG, to compute statistics about the occurrence of unstable states. Regarding the fact that unstable states are not often observed for randomly generated clusters, and that these configurations are the most representative of AKMC simulations aimed at calculating diffusion coefficients of clusters, no such post analysis was performed.

As an application of the AKMC model described above, diffusion coefficients of clusters of 3 to 60 vacancies were calculated with series of 100 independent AKMC simulations, at different temperatures, following the procedure described in Ref. (138) and Ref. (139) (Paper VIII in Appendix I). The complete series of data obtained can be found in Appendix A. Fig. 5.5 shows the evolution of the diffusion coefficient with the number of vacancies in the cluster, at three different temperatures. We see that the diffusion coefficient is not continuously decreasing for increasing size. Two local minima, for instance, are visible on the figure. They translate the fact that some clusters are more stable than others, and are consequently less mobile. The increased stability of these clusters is explained by the fact that they correspond to numbers of vacancies that form a compact structure, filling completely some shells of close neighbors around the geometrical centre of the cluster. We see on the figure that the most stable clusters are close to the perfect sphere on the 1nn (first nearest
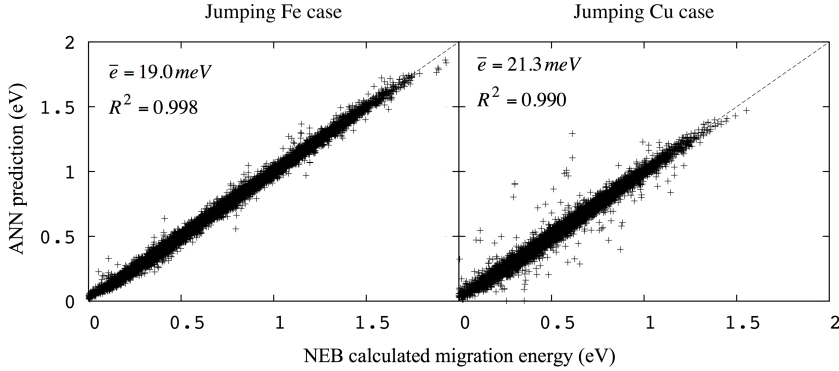
Figure 5.6: ANN quality of prediction for the vacancy migration energy in the Fe-Cu system. The ANN was trained with a database of 50,000 examples containing an equal number of configuration were vacancies and Cu atoms are dispersed in the volume and were vacancies are in clusters. Predictions for both types of configurations are mixed on the figure. The GIACA algorithm converged after connection to the ANN of the 9th nearest neighbours (173 atomic sites), and the addition of 12 hidden nodes.

neighbours), 2nn and 5nn sense, at all temperatures. As a consequence, clusters of 60 vacancies are found to be almost equally mobile as clusters of 15 vacancies, and clusters of intermediate sizes are more mobile, by up to two orders of magnitude (depending on temperature). During the AKMC simulation, all clusters larger than 6 vacancies were found to be very long-lived, and did not dissolve until after long series of AKMC events. A proper calculation of the lifetime therefore requires longer simulations to be conducted, and at high temperatures, in order to have enough statistics. For the moment, therefore, there are no results concerning the stability of vacancy clusters larger than size 6.

5.1.4 Cu-vacancy clusters

The same procedure as in section 5.1.3 was followed to calculate diffusion coefficients of Cu-vacancy clusters in an otherwise pure Fe matrix environment. The potential used for the NEB calculations was taken from (140). The ANN was trained, similarly to section 5.1.3, on the basis of examples where vacancies and Cu atoms are either dispersed in the volume or in cluster. The number of vacancies in this case was limited to 12, and the composition of the clusters was restricted to $N_{Cu}/N_{Vac} = 1/3$ to 3, where N_{Cu} is the number of Cu atoms in the cluster and N_{Vac} is the number of vacancies.

The quality of the prediction, after training, is shown in Fig. 5.6. As discussed
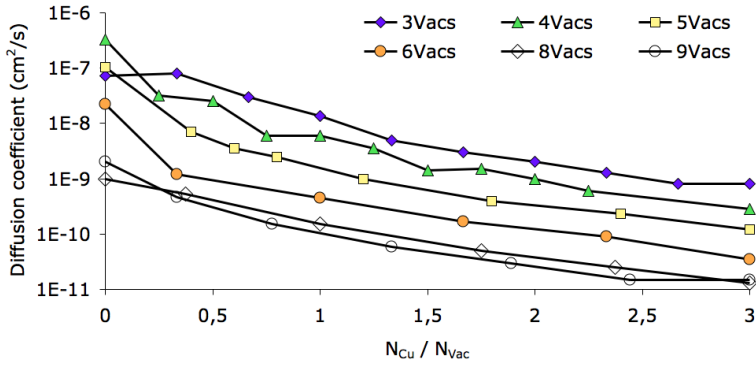
75

Figure 5.7: Diffusion coefficients calculated with 100 independent AKMC simulations, at 900K, for Cu-Vac clusters in a pure Fe matrix.

in section 3.4.4, the "1-of-c coding" convention is more appropriate to formulate the LAE as ANN input variables, because the configurational variables can take 3 different values. Here, 2 bits are necessary, and the number of input variables is thus two times as large as the number of atomic sites included in the LAE. We see that the predictions are very accurate in all cases, independently of the spatial arrangement of the vacancies. Diffusion coefficients and lifetimes were also calculated with series of 100 independent AKMC simulations for copper-vacancy clusters. The complete series of data obtained can be found in Appendix A, and are summarized in Fig. 5.7 for one temperature. We see that, the number of vacancies being constant, the diffusion coefficient is generally decreasing if more Cu atoms are added to the cluster. Also, the number of Cu atoms being constant, the diffusion coefficient generally decreases if more vacancies are added. These trends, however, are not always respected. This can be explained by two facts: (i) similarly to vacancies in pure Fe, "magic" numbers of vacancies and Cu atoms correspond to more stable configurations (these are, however, more difficult to identify than in the case of only vacancies); (ii) the global migration mechanism of the cluster is the consequence of a sequence of elementary migration events, that can either be added or removed if a couple of elements of the cluster are either added or removed. We therefore expect the dependence of migration properties to deviate from general trends for small sizes, but to become more self-consistent if the size of the cluster is increased.

The complete tabulation of all possible clusters compositions is a computationally extremely demanding work, although of little conceptual difficulty, and interpolation/extrapolation from the data for other compositions can be delicate, because of the non-respect of general trends in particular cases. For this reason, a more convenient approach is to design a general numerical regression of the diffusion
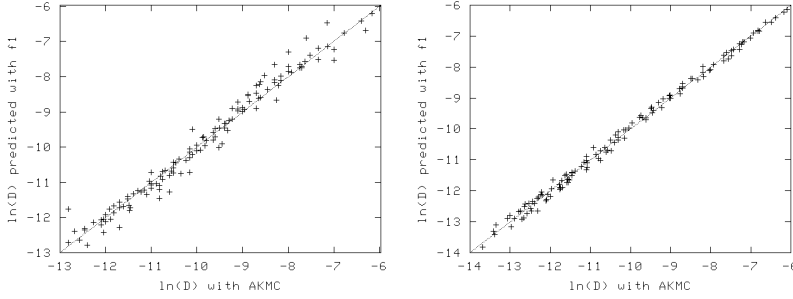
Figure 5.8: prediction of the diffusion coefficient of Cu-Vac clusters using function f_1 in Eq. 5.2, implemented by an artificial neural network. (left) For small clusters with 3 to 9 vacancies. The average error of prediction, in the natural logarithm space, is 0.197, and $R^2 = 0.987$. (right) For big Cu clusters with one vacancy. The average error of prediction, in the natural logarithm space, is 0.106, and $R^2 = 0.998$.

coefficient D and lifetime ν based on the following choice of variables:

$$\ln(D) = f_1(N_{Vac}, N_{Cu}, 1/k_B T) \tag{5.1}$$
$$\ln(\nu) = f_2(N_{Vac}, N_{Cu}, 1/k_B T) \tag{5.2}$$

Such functions are fitted on large sets of data and therefore potentially more accurate than extrapolations/interpolations with simple functions based on subset of the available data. For example, using the Arrhenius functions ($ln(D)$ versus $1/k_B T$) fitted on data concerning only one cluster at a time, and only varying temperature, the error bar associated with each regression is likely to be larger than for a function such as those in eqs. 5.2 and 5.2, fitted to all data. The functions f_1 and f_2 need, however, to be determined, using a sufficiently powerful and general regression technique. ANNs are again amongst the best candidate. An example of f_1 fitting using an ANN is shown in Fig. 5.8, obtained by training on the data of Tab. A.1 and Tab. A.2 and Tab. A.3 for $N_{Vac} < 10$. We see that the predictions are accurate on the average, although not satisfactory for all individual cases. Better accuracy can be probably achieved with a larger training dataset. More accurate fitting was already obtained in Ref. (35) (Paper IV in Appendix E), as shown in Fig. 5.8, for the simulation of thermal annealing experiments in the Fe-Cu system: there, function f_1 was designed to predict the diffusion coefficients of VCu clusters, with only one vacancy.

5.1.5 Fe-Ni-Cu alloy

A similar procedure as followed in sections 5.1.3 and 5.1.4 can in principle be followed for this ternary alloy, using the IAP developed by Bonny et al. (24; 25),
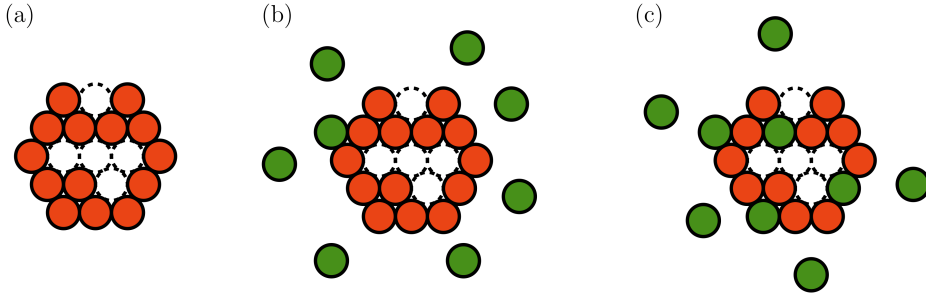
Figure 5.9: Generation of LAEs for the training of ANN in the Fe-Cu-Ni system. Red circles are Cu atoms, and green circles are Ni atoms. (a) configurations without Ni are taken from the table used for the Fe-Cu system; (b) Ni atoms are randomly included in the same LAE's, but never in positions already occupied by either vacancies or Cu atoms; (c) Ni atoms are randomly included in the same LAE's, but never in positions already occupied by vacancies.

which extends to the Fe-Ni-Cu system the Fe-Cu potential developed in (140): the presence of an additional chemical species can straightforwardly be taken into account in the model. The design of an accurate ANN to predict the vacancy migration energy, however, is more delicate because the numerical complexity of the problem increases. The ultimate objective when considering this alloy is to evaluate the effect of Ni on Cu precipitation in Fe, enhanced by neutron irradiation. It is therefore important to ensure that the influence of Ni on the migration energies is properly understood by the ANN, but at the same time that the ANN remains accurate if, as a limiting case, no Ni is found in the LAE. To provide direct ways to evaluate this ability, a different procedure was followed to generate the training table. The latter was composed of three different parts, in equal proportion:

- The first part is a collection of examples taken at random from the table generated for the Fe-Cu system. These LAE's, therefore, do not contain Ni atoms.

- For the second part, the same examples chosen in the first part were used, but a small number of Ni atoms were added at random positions of the LAE, as depicted in Fig. 5.9b. Ni atoms were however not placed at positions already occupied by either a vacancy or a Cu atom. These new configurations therefore correspond to Cu-vacancy clusters migrating in a dilute Fe-Ni environment.

- For the last part, the same procedure as for the second part is followed, but Ni atoms are not forbidden to replace Cu atoms. The formation of Cu-Ni-Vac clusters is therefore allowed, as shown in Fig. 5.9c.
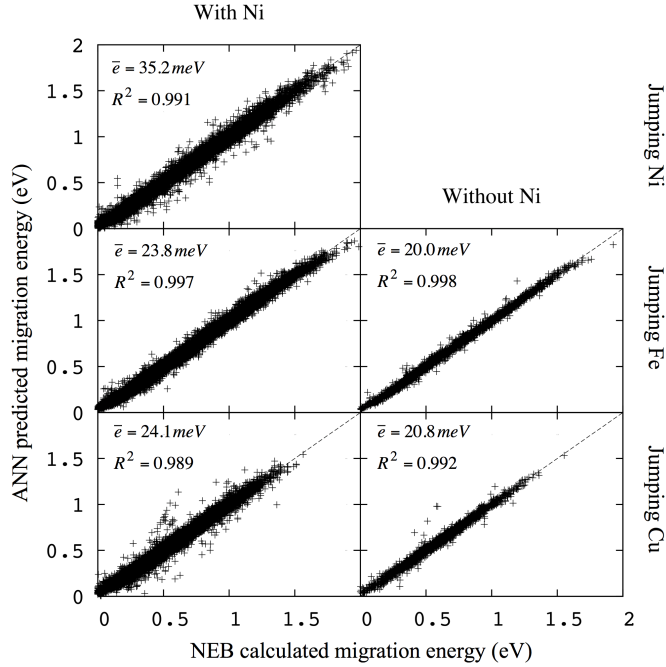
Figure 5.10: ANN quality of prediction for the vacancy migration energy in the Fe-Cu-Ni system. Predictions are also shown for Fe-Cu configurations, i.e. without Ni, for comparison. Using as training set a database of 75,000 examples, the GIACA algorithm converged after connection to the ANN of the 5th nearest neighbours (77 atomic sites), and the addition of 14 hidden nodes.

The quality of prediction, after training, is shown in Fig. 5.10. Again, the "1-of-c coding" convention is more appropriate to formulate the LAE as ANN input variables: here, 3 bits are necessary, and the number of input variables is thus three times as large as the number of atomic sites included in the LAE. The mean error of prediction is somewhat larger compared to the Fe-Cu case, but remains reasonably acceptable. Predictions for Fe-Cu configurations are also shown. The mean error of prediction, for these cases, is comparable to the specialized ANN shown before in Fig. 5.6. It is therefore clear that predictions are less accurate for the Fe-Cu-Ni system because of the higher complexity from the numerical point of view, due to an additional chemical species that must be taken into account. As an additional test, Fig. 5.11 shows the comparison of the difference in the migration energy due to the presence of Ni atoms in the LAE, replacing Fe atoms. This comparison can be performed by computing the difference between the energy barriers from the second and first parts of the table of examples. This difference is predicted with the same

Figure 5.11: ANN quality of predictions of the change in energy barrier caused by the presence of Ni atoms.

accuracy, globally, as migration energies themselves, which shows that the effect of Ni on the energy barrier is globally understood by the ANN. From a physical, rather than numerical, point of view, it is difficult to draw any conclusion from just these numbers on the effect of Ni on the diffusion coefficients of Cu-vacancy clusters and on Cu precipitation in general, because positive and negative differences seem to compensate. So, whether or not Ni influence Cu precipitation in the model and how will depend on the detail of which jumps are affected in one sense or another. Only a full simulation can tell whether such effects exist and in which sense it goes.

Diffusion coefficients and lifetimes of Vac-Cu-Ni clusters can thus be calculated following the same method as described in the previous sections. For convenience, numerical regressions similar to f_1 in Eq. 5.2 and f_2 in Eq. 5.2 can be designed: only a new variable, the number N_{Ni} must be added.

5.2 Introduction of self-interstitial atoms and their clusters in the AKMC simulation

The last necessary ingredient to simulate the effects of irradiation in metals is to allow for the presence of self-interstitial atoms (SIA). This point-defect is the counterpart of the vacancy, because a self-interstitial atom corresponds to an additional atom as compared to the number of available lattice positions. In metals, this extra atom is accommodated by sharing lattice site with another atom, i.e. two atoms are found in the atomic volume associated with only one atom in the perfect lattice. Very few attempts have been made in the past to include SIAs in AKMC

models, for example in Ref. (13; 91; 162; 176; 177). This can be easily explained by the fact that SIAs, and especially their clusters, induce significant perturbations in the lattice, and a rigid-lattice-like description of their migration and interactions with vacancies is therefore not realistic.

Soisson, in Ref. (161; 162), introduced large amounts of SIAs and vacancies in generic A-B alloys. The migration energies were calculated on a rigid lattice with a simplified cohesive model based on pair interactions (up to the 1nn distance only), fitted to ab initio energy data. Opposite point-defects, in his approach, are blind to each other, i.e. the migration energy of vacancies is not influenced by the presence of nearby SIAs, but only by other vacancies, and conversely for SIAs. Recombination of point-defects, also, was treated in a simplified manner, occurring as soon as they are closer to each other than a predefined spherical range. Radiation induced segregation and precipitation was, in spite of these simplifications, satisfactorily predicted. Later, Vincent et al. (176; 177) conducted similar simulations in a Fe-Cu-Ni-Mn-Si alloy, which is in fact only possible with ab initio energy data up to now, because of the absence of a corresponding IAP.

In the framework of the general methodology developed in this thesis, attempts have been made to introduce SIAs in the most rigorous possible way, using the ANN-based AKMC algorithm proposed in section 2.3. Ideally, the goal would be to avoid the assumptions made by previous authors, by calculating the point-defects migration energies in a non-rigid lattice (using the NEB method), thereby taking into account the complete effects of atomic relaxation up to the convergence range for the LAE. In addition, point-defect recombinations should be described without approximation, i.e. not assuming a constant spherical range of recombination, but instead taking the decision of whether to apply recombination or not in a more realistic way. Artificial neural networks, for instance, can eventually succeed in providing the relevant information about the migration of SIA, in a similar way as was achieved in the case of vacancies. The methodological and practical limitations of the ANN-based AKMC algorithm proposed in section 2.3, and later on generalized in section 5.1.1, are in fact identical if SIAs are included, except that they are more intense in magnitude and pose more difficult problems:

- The number of migration events that should be considered, i.e. that correspond to migration energies that are in balanced competition with each other, is sensibly larger compared to the case of vacancies. As a matter of fact, SIAs can take several different orientations after the migration, and migration can occur a priori to various distances: 0nn migration events correspond to a change of orientation without migration to another lattice site, next the same reorientation can be accompanied by a migration to a 1nn or 2nn distance. Migrations to farther away distances are also not to be neglected, particularly in the case of crowdions (additional atoms on a close-packed atomic direction) and their clusters. Finally, the simultaneous migration of two or more dumbells

is, contrary to the case of vacancies, often favourable in terms of energy barriers (see for example *ab initio* barriers in (61)).

- The relation of unicity between a given state of the system in the rigid lattice world of the AKMC and in the off-lattice world is not at all guaranteed. SIAs can often, depending on the LAE, take several different equilibrium positions, all of them corresponding to the exact same state in the rigid lattice world, because of two degrees of liberty: the distance between the atoms of the dumbell and also the orientation of the vector joining them, that defines the type of dumbell.

- The occurrence of spontaneous migration events is not at all negligible in general, especially when SIA's are close to each others.

- Last, but certainly not least, the computation of the migration energy between two given (stable) states using NEB is in general not a routine operation in the case of SIAs. The calculation is first of all delicate because of the large deformations of the lattice, and the obtained migration path with NEB may sometimes not converge to a physically acceptable transition because of numerical problems. Secondly, several transition paths can, a priori, exist between two given metastable states. In practice, it has been observed that the migration path found by NEB can differ if the initial guess, usually calculated as a linear interpolation (see discussions in section 2.3.2), is calculated with another method, such as a drag for example.

In short, it is clear that a completely rigorous treatment of SIAs in AKMC simulations is probably unrealistic in practice, because no approach could on the one hand account for all states of the systems and all transitions between them, and on the other, be fast to compute at the same time. The key point for the development of a physically acceptable and numerically affordable model is, therefore, to search for the most important mechanisms to be allowed for, while limiting simplifications as much as possible.

Two different approaches were used in this thesis, and are described in the following two sections. In section 5.2.1, an attempt is made to include several SIAs in a pure Fe matrix, without restrictions on the elementary migration mechanisms that could be defined in the AKMC simulation. Next, in section 5.2.2, a possible simplification of the problem is proposed, going along the same line as Soisson and Vincent et al. in Ref. (162; 177). Vacancies and self-interstitials are introduced together in the system, but their interactions is treated in a simplified way, as well as the formation and migration of SIA clusters.
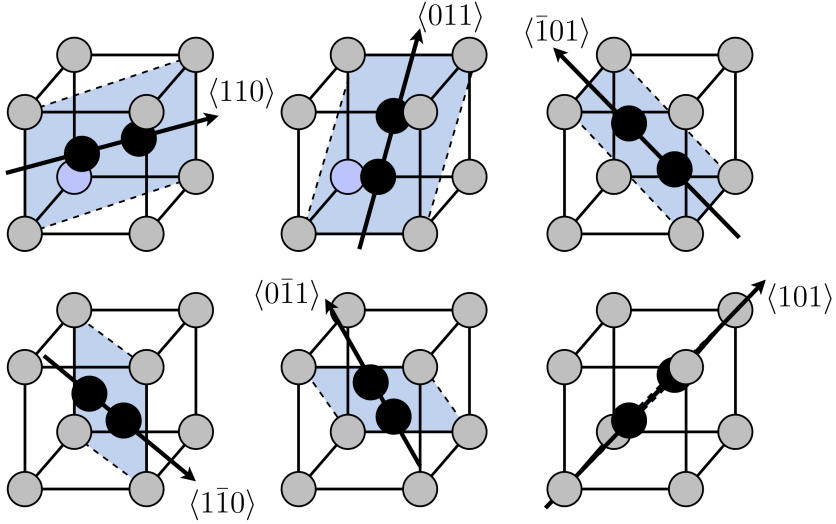
Figure 5.12: Six possible Dumbell orientations in bcc crystallographic structures.

5.2.1 Clusters of dumbells in pure Fe

As a preliminary study, to evaluate the complexity of the problem and fix the ideas, attempts have been made to consider in a general way the migration of several SIAs in pure Fe. For simplicity, the study was limited to dumbells orientated in <110> directions, which correspond to the stable SIA configuration in pure Fe, even in the presence of nearby solute atoms (175). Taking a given crystal, with a fixed origin and xyz axis, a total of six different <110> dumbell orientations can be found, as depicted in Fig. 5.12. To improve the stability of these orientations, independently of the spatial arrangement of SIAs in the volume, the total number of dumbells was limited to 5, because bigger clusters take stable configurations in <111> *crowdion* orientations. If successfully generalized, the AKMC would then be able to model the formation and migration of small clusters of interstitials, without assumptions on the migration mechanism. This is however known to be a delicate task in a rigid lattice framework, because for example the configuration of even small clusters of only three dumbells was observed in Ref. (167) to be difficult to describe if they are not aligned. The approach that was attempted can be so summarized:

- Many possible migration events were defined. In a first step, only the migration of one <110> SIA to a 1nn position was considered, as depicted in Fig. 5.13, limited to the case in which the final orientation after migration is <011>, as this could be easily extended to the other two possibilities.
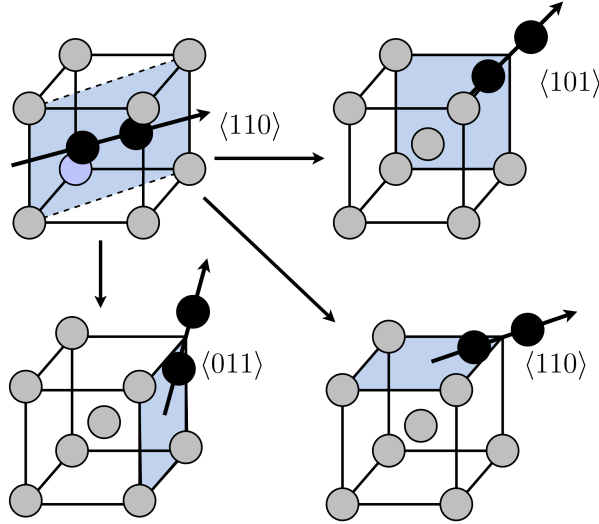
Figure 5.13: Three migration events to a 1nn position considered for a dumbell. The position of the destination is the same, but the dumbell can take, after migration, three different orientations.

- Similarly to the introduction of several vacancies in the AKMC, as discussed in section 5.1.1, the most natural way to define the LAE of the migrating SIA is to consider neighbouring SIA's as additional chemical species of the alloy, and different integers are assigned to describe every possible orientations they can take. The problem is thus unquestionably sensibly harder from the numerical point of view, because of the large number of categories of entries that the LAE contains, but also because of the stronger deformations of the rigid lattice that make the relationship between the LAE and the migration energy harder to describe with a simple ANN.

- As a first step, no migration energies were calculated with the NEB method, but only the initial and final states were relaxed with CG. The ANN was then trained to predict the difference between the relaxed energy of the final state of the migration event and the total energy of the initial state. This energy difference is the input for a Kang-Weinberg formalism, as already discussed in section 2.2.3.

- The existence of several metastable states corresponding to the same rigid lattice description was confirmed for a non-negligible proportion of the explored configurations. To get round this problem, every state was relaxed several times with CG, each of them starting from the rigid lattice configuration

Figure 5.14: ANN quality of prediction for the migration energy associated with the migration to a 1nn position of the <110> dumbell in pure Fe, taking a final <110> orientation, allowing for the presence of up to 4 other dumbells in its vicinity. (left) The 76 input variables are a description of the LAE up to the 5nn distance, coded as integers that can take 7 different values. (righ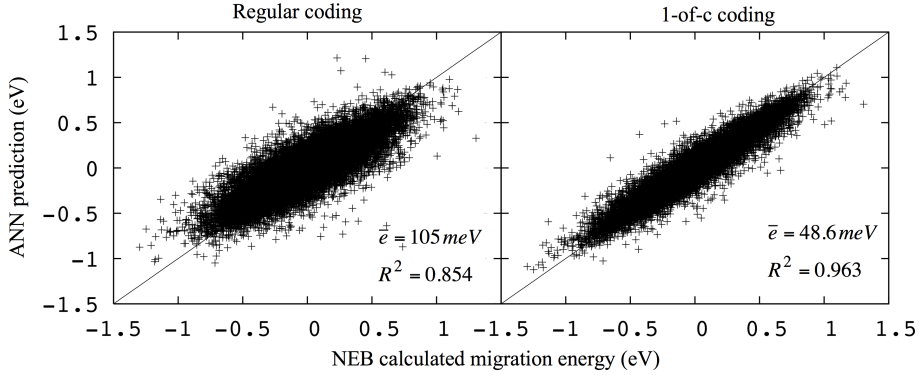t) The 456 input variables are a description of the LAE up to the 5nn distance, coded with the 1-of-c convention with 6 bits. The network counts 9 hidden nodes.

slightly perturbed by random atomic displacements of small magnitude. In some cases, it was observed that the variation of the energy of the system, after relaxation, can be as large as several tens of meV. Only the state corresponding to the lowest energy was retained, the other ones were discarded.

- A number of 90000 examples of LAEs versus relaxed energy difference were calculated. Up to 4 dumbells of random orientations were placed in random positions within the 5nn distance from the migrating one. If encountered, unstable states were discarded from the table.

- Some positions of the LAE are situated in regions of large strain caused by the migrating dumbell. In these areas, some dumbell orientations are almost never found to be stable, especially of course if their strain fields are pointing in the direction of the migrating dumbell. These entries of the LAE have therefore been restricted in such a way that the concerned dumbell orientations are forbidden. In total, the number of restrictions was 31.

The quality of ANN predictions, after training, is shown in Fig. 5.14. For the sake of comparison, two different networks were trained: The first one receives as inputs the LAE coded as 76 integers that can take 7 different values, whereas for the second one the LAE is coded using the 1-of-c convention (see section 3.4.4). In the latter case, the input variables are thus a number as large as 456 bits. We see that predictions of the ANN with 76 input variables is poor compared to the accuracy that was achieved, in

general, in the case of vacancies. This, shows the relative difficulty of treating SIAs compared to vacancies. The quality of predictions, however, is sensibly improved if the LAE is coded using the 1-of-c convention, despite the fact that the number of input variables is multiplied by 6. This interesting observation can be ascribed to two different reasons:

- Training the ANN that receives as input variables integers that can take 7 different values is, unquestionably, a challenge. From a purely numerical point of view, the difference between the two networks shown in Fig. 5.14 can thus be explained by the fact that bitwise variables more easily express, in the nodes of the ANN, a clear and distinguishable difference between the different dumbell orientations. Nothing, therefore, tells that an equal quality of prediction for both ANN's could not be achieved if the number of hidden nodes was increased, and if the initial synapses in the ANN were chosen, before training is started, with a more appropriate procedure. This kind of uncertainty is proper to the ANN technique, and not specific of this particular application.

- From the physical point of view, the 1-of-c coding convention is, in this case, equivalent to a re-partitioning of the simulated volume. Each Wigner-Seitz cell is effectively divided in 6 sub-cells, each encompassing the volume than can be occupied by the atoms of the SIA if they take one of the 6 possible orientations. This point of view is interesting, because it suggest that, after all, one should not be surprised that the ANN manages a better understanding of the physics of the problem at hand if the latter is described in a more "explicit" way.

As a conclusion, the results discussed in this section suggest that the definition of a fully atomistic description of SIAs in an AKMC is apparently feasible. The problem is, however, numerically very demanding and delicate, with the consequence that a successful extension to more realistic and complex cases would probably not give sufficiently satisfactory results. Simplified approaches, such as the one proposed in the next section, must therefore be found.

5.2.2 Simplified model for the simulation of isochronal annealing experiments with Fe-Cr alloys

As discussed in the previous section, the complete and rigorous modelling of SIA migration in AKMC, describing properly the formation and migration of clusters and their interactions or recombination with vacancies, is probably an utopian objective. One alternative is thus to elaborate less complete models, embedding different levels of approximations, somehow similarly to the hybrid

AKMC algorithm that was proposed and applied for the simulation of TAE in Fe-Cu alloys in Ref. (38) (Paper VII in Appendix H) and section 4.3.1. In this respect, for instance, the migration of clusters of SIAs can be hybridized in a similar way, to avoid the considerations of many complex elementary transitions where problems of temporary configurational instability would be encountered. The biggest challenge, in reality, is probably to find an adequate, if simplified, way to describe the recombination of SIAs with vacancies, while possibly accounting for chemical effects due to different atomic species present in the LAE of the migrating SIA. As a first step in this direction, the case of only single SIAs was considered. Such a model can be used to simulate isochronal annealing experiments following low temperature electron irradiation (generally monitoring electrical resistivity recovery), for the sake of validation by allowing comparison with experimental data, as introduced in Ref. (169) (Paper IX in Appendix J).

The case of the migration of a single SIA in Fe-Cr alloys was considered, because Cr atoms interact fairly strongly with SIAs and therefore their effect is worth being taken into account. In pure bcc Fe, the ground state for SIA is a <110> dumbbell (124), and the most probable migration mechanism suggested by Johnson is a single hop to a 1nn position (83), as depicted in Fig. 5.13. On the contrary, in pure bcc Cr, the ground state of the SIA is the <111> crowdion: consequently, the migration mechanism in Fe-Cr alloys with high Cr content changes from Johnson's to the crowdion's glide. As a consequence, Cr influences also the configuration of the SIA and, for concentrated alloys containing more than 25% of Cr, it was observed that <110> dumbells are often unstable, and tend to glide for long distances (sometimes several lattice units !) to reach a stable position as a <111> crowdion with Cr atoms. The frequency of occurrence of such spontaneous migrations, however, becomes negligible for lower Cr concentrations. Thus, to start with, the case of alloys of not too high concentration was considered and, an ANN was trained to predict the migration energies associated only with the three jumps depicted in Fig. 5.13:

- A number of 27500 examples of dumbbell migration jumps in changing chemical environments were calculated with the NEB method, for the three types of events.

- The LAE was extended up to the 9nn (173 atomic sites), and the Cr concentration was randomly chosen between 5% and 25%. Cr atoms were placed in dispersed configurations only.

- The problem of the existence of multiple metastable states corresponding to the same rigid lattice description was confirmed for a non-negligible proportion of the explored configurations. To get round this problem, every state was relaxed several times with CG, each of them starting from the rigid lattice configuration, slightly perturbed by random atomic displacements of small magnitude. In some cases, the difference in energy of the system, after

Figure 5.15: ANN quality of prediction for the migration to a 1nn position of the <110> dumbell in Fe-Cr alloys. The input variables are a description of the LAE up to the 9th shell of close neighbours (173 atomic sites), and the network counts 10 hidden nodes.

relaxation, can be as large as several tens of meV. Only the state corresponding to the lowest energy was retained, and the other ones were discarded.

The ANN quality of prediction, after training with the GIACA algorithm for each of the three types of jump individually, is shown in Fig. 5.15. We see that predictions are less accurate than achieved in general for the migration of vacancies, but are in any case highly correlated with the values calculated with NEB ($R^2 > 0.95$). The average error committed is here probably higher than the accuracy of the NEB, especially if the final dumbell orientation is <110>. This is attributed to the numerical difficulty of constructing an accurate regression of the relation between the LAE and and the migration energy.

AKMC simulations were performed to evaluate de diffusion properties of dumbells in Fe-Cr alloys, as predicted by the model. The migration energies were, however, estimated with a Kang-Weinberg formula, using the ANN shown in Fig. 5.16 to predict the relaxed energy difference associated with the jumps. This work will be repeated in a near future using the ANN shown in Fig. 5.15 instead. Two different studies were performed:

- In the first one, the SIA diffusion coefficient was calculated in Fe-Cr alloys, for temperatures in the range between 500 K and 1200 K, in a box composed of 20x20x20 lattice units (i.e. 16 000 atoms). As a benchmark, the same calculations were also performed with MD (this is possible, at least in the range of high temperatures, because the simulation box was deliberately chosen to be small). Results are compared in Fig. 5.17. We can see that migration energies, fitted in Arrhenius plots, are not rigorously identical, but the curves shown in the figure are reasonably similar: these curves are not strictly parallel, as would be the case if the migration energies were the same, but the degree of non-parallelism visible on the figure does not significantly change the evolution with temperature of the diffusion coefficient, at least in the range of diffusion coefficients covered by the figure. Last, the (nearly constant) distance between the curves in the figure is due to the choice of the jump attempt frequency in the AKMC simulation (exponential prefactor in Equ. 2.5), that was taken to be $\nu_0 = 10^{13} \ s^{-1}$. A higher value, by roughly one order of magnitude, would have been more appropriate, apparently.

- In the second one, AKMC simulations were performed in much bigger boxes composed of 100x100x100 lattice units (i.e. 2 million atoms), to obtain better statistics, in Fe-Cr alloys with varying Cr content. The obtained values of the SIA diffusion coefficient are shown in Fig. 5.18, as well as the correlation factor. The latter is the ratio between the frequency of jump (to a given distance, for example 1nn) calculated on the basis of the diffusion coefficient, and the one directly measured from the results of the simulation. In other words, if the correlation factor equals to 1, the SIA migration is "uncorrelated", i.e. completely random in the space, whereas a smaller value reveals that series of migration jumps are correlated to each other and do not contribute to the value of the diffusion coefficient. We see in Fig. 5.18 that such a correlation is observed for all alloy compositions at low temperature, implying that SIAs can be locally trapped in low energy state configurations. This result is in good agreement with experimental evidences, for example the work of Dimitrov et al. in Ref. (45) and the work of Nikolaev et al. in Ref. (121).

The next step to be taken in the future is to attempt the design of ANNs that could predict not only the SIA migration energy (as in Fig. 5.15), taking the LAE of Cr atoms into account, but also the possible presence of a vacancy. In addition, another ANN can be trained to predict, given the respective positions of the point-defects as input, whether immediate recombination should take place or not. This work requires the generation of databases of examples that are, each, rather CPU-time demanding, mainly because the presence of close antagonist point-defects increases the complexity of states relaxations with CG. This work is currently in course.

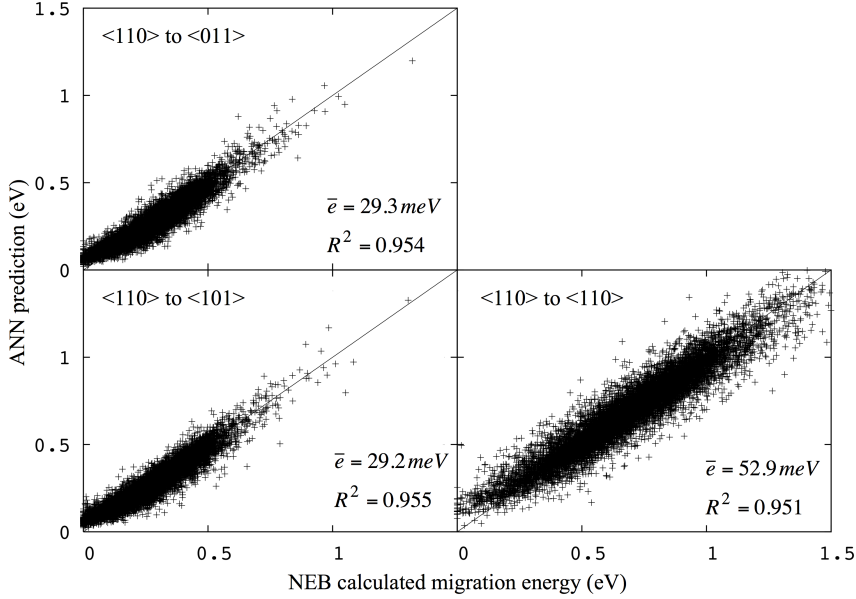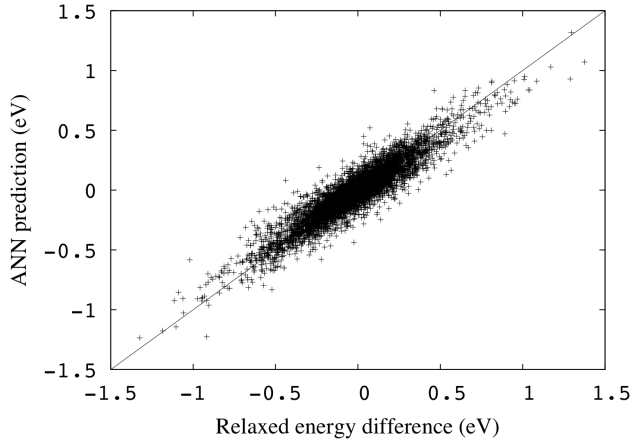**Figure 5.16:** ANN quality of prediction of the relaxed energy difference, calculated with CG, for the migration to a 1nn position of the <110> dumbell in Fe-Cr alloys. The input variables are a description of the LAE up to the 9th shell of close neighbours.

5.3 Conclusions

In this chapter, the new atomistic kinetic Monte Carlo (AKMC) scheme proposed in this thesis, based on the use of artificial neural networks (ANN), and originally developed in chapter 2 for the single-vacancy case, has been successfully generalized to allow for the presence of several point-defects in the simulated volume.

The case of vacancies, without self-interstitial atoms (SIA), was relatively straight-forward thanks to the fact that spontaneous migration events can be neglected, and that the migration mechanism of their clusters was observed to be dominated by simple elementary events. As an application, Diffusion coefficients and lifetimes of Cu-vacancy clusters, that are necessary input values for the simulation with coarse-grained models of the effects of irradiation in steels, have been calculated and tabulated in Appendix. The calculation of more data, for larger numbers of vacancies and Cu atoms, is however necessary to complete the database and allow for simulations in more general conditions. One possible way to reduce the amount of AKMC simulations is to design general regression methods, for example implemented with ANN, fitted on the basis of a limited amount of examples, to predict these parameters for any case. The bases to do so have been set.

The case of SIAs, however, turned out to be sensibly more delicate. It has been shown that they can be introduced in the ANN-based AKMC model, at the price of simpli-fications of the possible states of the systems, and of the possible transition paths between them. As a next step to investigate the feasibility of simulating neutron
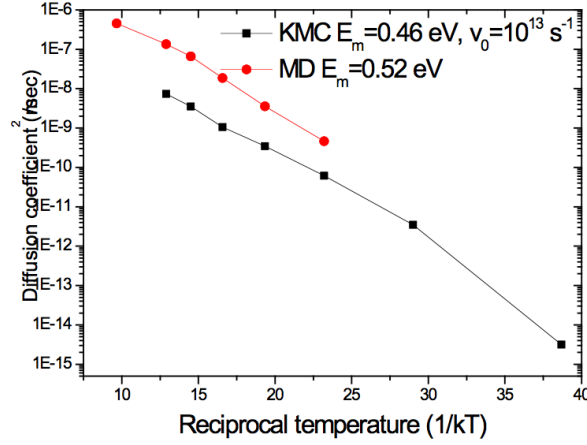
Figure 5.17: Diffusion coefficients for a single SIA in Fe-10%Cr alloys, versus the reciprocal temperature $(1/k_B T)$, calculated using AKMC and MD simulations. The corresponding migration energies E_m, fitted with Arrhenius plots are also reported.

irradiation damage, attempts to describe rigorously the immediate recombination between SIAs and vacancies will soon be made, relying again on ANN. In a longer term perspective, hybrid approaches such as the one proposed for thermal ageing simulations in FeCu (see chapter 4) can be devised to allow the model to treat SIA clusters in a physically acceptable way, without the need of retaining the full atomistic description.
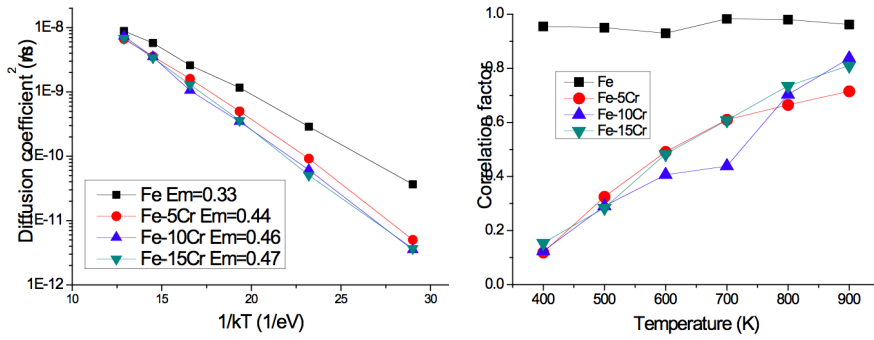
Figure 5.18: (left) SIA diffusion coefficients calculated with the AKMC method at T = 400-900 K in different Fe-Cr alloys. (right) Correlation factors between the SIA diffusion coefficient and the jump frequency, in the same conditions.

6 Conclusions and future perspectives

6.1 Summary and conclusions

In this thesis, a novel and innovative atomistic kinetic Monte Carlo (AKMC) model, based on artificial neural networks (ANN) has been proposed in chapter 2 to perform a link, in the general multiscale modelling scheme, between the atomic level and coarse-grained models. The simulated process is naturally decomposed in elementary thermally activated migration events involving point-defects, i.e. vacancies or self-interstitial atoms (SIA), that are in competition with each other according to their respective frequencies of occurrence. The latter are calculated on the basis of the migration energies, that are calculated with minimal approximations using a method that takes into account all the effects of long-range static atomic relaxation and chemical interactions. The proposed AKMC algorithm is therefore a physical model, entirely based on a given interatomic potential which is exploited in the most complete possible way, without the need to define arbitrary parameters that should be, for example, fitted to experimental data.

The model was first applied, in chapter 4, to the simulation of thermal annealing experiments, for the main purpose of validation by comparison of the obtained results with experimental data. Very satisfying comparison was achieved for two model alloys of interest for nuclear materials science: (i) Fe-Cu, which is a model alloy for bainitic reactor pressure vessel steels of western light-water nuclear reactors: (ii) Fe-Cr, which is a model alloy for high Cr ferritic-martensitic stainless steels, that are candidates for many concepts of new generation nuclear reactors. In both cases, the evolution with annealing time of the average clusters radius and the clusters density were in excellent agreement with experimental data, more closely, or up to longer simulated times, than previously obtained by other authors. This showed that the model is globally correct, and that all important mechanisms are

included with the relevant accuracy, a result which can be ascribed to two reasons: (i) the interatomic potentials used are correct from the thermodynamic viewpoint and, clearly, also adequate to describe the kinetics of precipitation processes (ii) the "filter" introduced by the artificial neural network between the predictions of the potential and those of the model is not such that the good qualities of the former are lost. In addition, this showed that the modelling technique developed here is particularly efficient (affordable computing time while retaining high accuracy) and versatile (the method is suitable to be generalised to more complex problems).

Indeed, in chapter 5, the ANN-based AKMC algorithm was successfully generalized to treat any number of vacancies, which is one of the two necessary ingredients for the simulation of neutron irradiation damage in metals. The methodology could be generalized without fundamentally changing the way the system is described, i.e. inherently as a rigid lattice as in most of the existing AKMC models, thereby limiting the complexity of the simulation by keeping all advantages of simplicity of the definition of migration events, and the simplicity of calculating their corresponding frequencies of occurrence. This allows for the simulation of long and complex processes, for example the calculation and tabulation of diffusion coefficients and lifetimes of Cu-vacancy clusters, that are necessary input values for coarse-grained models. The calculation of these parameters by molecular dynamics would be simply unaffordable; the technique developed here is almost equivalent in terms of physical information to using MD, but makes these calculations possible in a routine way.

Last, convincing evidence has been given that the ANN-based AKMC algorithm can be generalized to the introduction of SIAs as well. Simplifications must however be made, as it has been shown that all possible states of the system and all possible migration mechanisms could not be practically included in the model. A possible approach, for the simulation of radiation processes under simplifying assumptions, i.e. involving simultaneously SIAs and vacancies and including a rigorous description of recombination processes, almost as detailed as in an MD simulation, has been proposed, and its feasibility has been demonstrated.

6.2 Future perspectives

In this thesis, the newly developed ANN-based AKMC algorithm was mainly applied to binary systems. In chapter 4, thermal annealing experiments were simulated for the Fe-Cu and Fe-Cr systems, that are model alloys for steels. More realistic simulations require the introduction of additional chemical species. It has been demonstrated, in sections 3.4.4 and 5.1.5, that this is possible in the case of Fe-Cu-Ni systems (The potential was developed by Bonny et al. in Ref. (25)), despite the somewhat less satisfactory numerical precision that is achieved by ANNs for

the prediction of point-defects migration energies. Our future objective will be to simulate thermal ageing experiments in Fe-Ni-Cu alloys, taking a hybrid approach similar to the work described in section 4.3, as a first step before diffusion coefficients and lifetimes of Cu-vacancy clusters, including also Ni atoms, are calculated, as a natural continuation of the work presented in section 5.1.5. The main bottleneck for other systems is in fact the development of the corresponding interatomic potentials, that is a more delicate task for ternary or more complex alloys compared to binary systems. Yet, new ternary potentials are under development and will soon be available, such as for example:

- A potential for the Fe-Cr-Ni system, model alloy for austenitic steels. In this case, the phenomenon to be studied is radiation-induced depletion of Cr at surfaces or grain boundaries, accompanied by Ni enrichment, that is known to increase susceptibility to stress corrosion cracking. This process can be simulated using the algorithm proposed in section 5.2.2.

- A potential for the Fe-Ni-Mn system, which is a model alloys for low-Cu RPV steels, of relevance for western Europe reactors. In this case, the first objective will be the study of the formation of so-called "late blooming phases", that are Ni-Mn precipitates, containing little or no Cu, that are believed to be an important origin of hardening and embrittlement. Their formation is, nowadays, presumed to be linked to small interstitials loops. This is therefore likely to be a radiation-induced phenomena, that cannot be studied as a thermal annealing process in the absence of irradiation. Its simulation represents a challenge as it implies being able to treat both types of point-defects and their clusters, as well as three chemical species.

Even remaining in the framework of binary alloys such as Fe-Cr or Fe-Cu, the next challenge to adequately complete the achievements of this thesis will be to develop methods providing an accurate description of the recombination between SIA and vacancies and, next, of SIA clusters (including of course recombination processes). Concerning recombination, the purpose is to overcome simplified descriptions of this process, for example by simply defining a constant range from which recombination is assumed to take place immediately, disregarding the influence of the local chemical environment. This can probably be achieved using, again, several ANNs: One ANN that predicts, given as input the relative positions of the SIA and the vacancy, as well as the distribution of solute atoms, whether recombination takes place or not; if not, a second ANN is used to predict the SIA migration energy, taking into account the presence of the near-by vacancy. This work requires the generation of databases of examples that are, each, rather CPU-time demanding, mainly because the presence of close antagonist point-defects increases the complexity of relaxation of metastable states with CG. Concerning SIA clusters, the exploratory work performed in the framework of this thesis showed that, while

a fully atomistic KMC description is probably feasible within certain boundaries (a few SIAs, limited amount of migration mechanisms, ...), its generalization beyond those boundaries is likely to be hopeless. Thus, in this case appropriate hybrid methods inspired to the one developed for Cu precipitation will have to be identified and applied.

Bibliography

[1] G.J. Ackland, M.I. Mendelev, D.J. Srolovitz, S. Han and A.V. Barashev, J. Phys. Condens. Matter 16 (2004) p.S2629.

[2] E. Alba and J.F. Chicano, "Training Neural Networks with GA Hybrid Algorithms", Lecture Notes in Computer Science, 2004, Volume 3102/2004, 852-863, DOI: 10.1007/978-3-540-24854-5_87

[3] J.-O. Andersson and B. Sundman, CALPHAD 11 (1987) 83.

[4] T. Ash, Connect. Sci. 1, 365 (1989).

[5] ASTM International, Predicting Radiation-Induced Transition Temperature Shift in Reactor Vessel Materials, E706 (IIF), E 900-02, ASTM, Materials Park, OH, 2007.

[6] M. Athenes, P. Bellon, and G. Martin, Defect Diffus. Forum 143–147, 297 (1997).

[7] B. P. Auger, P. Pareige, M. Akamatsu, and J. C. Van Duysen, J. Nucl. Mater. 211 (1994) 194.

[8] R.S. Averback, T.D. de la Rubia, Solid State Phys. 51 (1998) 281.

[9] D.J. Bacon, F. Gao, Yu.N. Osetsky, J. Nucl. Mater. 276 (2000) 1.

[10] A.V. Barashev, Philos. Mag. 85 (2005) 1539.

[11] G. T. Barkema and N. Mousseau, Phys. Rev. Lett. 77, 4358 (1996).

[12] G.T. Barkema and N. Mousseau, Comput. Mater. Sci. 20 (2001) 285.

[13] C.S. Becquart, C. Domain, Phys. Status Solidi b (2009), in press.

[14] P. Bellon, Thermodynamics, Microstructures and Plasticity 108 (2003) 395-409, and references therein.

[15] C.M. Bishop, Rev. Sci. Instrum. 65 (1994) 1803

[16] C.M. Bishop, "Neural Networks for pattern recognition", Clarendon press, Oxford, 1995.

[17] K. Binder, in "Monte Carlo Methods in Statistical Physics", edited by K. Binder Springer-Verlag, Berlin, 1979, Chap. 1.

[18] J.J Blackstock, G.J. Ackland, Philos. Mag. A 81 (2001) 2127.

[19] F. Bley, Acta Metall. Mater. 40, 1505 (1992).

[20] A.-S. Bogaerts, R. Gerard, R. Chaouadi, Comparison of embrittlement trend curves to high fluence RPV surveillance results, in: International Symposium on Contribution of Materials Investigations to Improve the Safety and Performance of LWRs (Fontevraud 7), Avignon, France, 2010 (26–30 September).

[21] G. Bonny, D. Terentyev, L. Malerba, Comput. Mater. Sci. 42 (2008) 107.

[22] G. Bonny, D. Terentyev, and L. Malerba, Scr. Mater. 59 (2008) 1193.

[23] G. Bonny, D. Terentyev, L. Malerba and D. Van Neck, Phys. Rev. B 79 (2009) 104207.

[24] G. Bonny, R. C. Pasianot, and L. Malerba, Modell. Simul. Mater. Sci. Eng. 17 (2009) 025010.

[25] G. Bonny, R.C. Pasianot, N. Castin, L. Malerba, Philos. Mag. 89 (2009) 3531.

[26] G. Bonny, D. Terentyev, and L. Malerba, J. Nucl. Mater. 385 (2009) 278.

[27] G. Bonny, "Many-body interatomic potentials for the study of radiation effects in nuclear structural materials", phD dissertation, Ghent University, Belgium, June 2009.

[28] B. Bortz, M.H. Kalos, J.L. Lebowitz, J. Comp. Phys. 17(1) (1975) 10.

[29] S.S. Brenner, M.K. Miller and W.A. Soffa, Scripta Metall. 16 (1982) 831.

[30] J.T. Buswell et al., Effects of Radiation on Materials: 14th International Symposium (Volume II), ASTM STP 1046, N.H. Packan, R.E. Stoller, and A.S. Kumar, Eds., American Society for Testing and Materials, Philadelphia, 1990, pp. 127-153.

[31] J.T. Buswell, W.J. Phythian, R.J. McElroy, S. Dumbill, P.H.N. Ray, J. Mace and R.N. Sinclair, J. Nucl. Mater., 225 (1995) 196-214.

[32] N. Castin, R. P. Domingos, International Journal of Computational Intelligence Systems 1, 4 (2008) 340.

[33] N. Castin and L. Malerba, in press: doi:10.1016/j.nimb.2009.06.041.

[34] N. Castin, L. Malerba, G. Bonny, M.I. Pascuet, M. Hou, in press: doi:10.1016/j.nimb.2009.06.092

[35] N. Castin and L. Malerba, J. Chem. Phys. 132, 074507 (2010).

[36] N. Castin, G. Bonny, D. Terentyen, M.Yu. Lavrentiev, D. Nguyen-Manh, in press: doi:10.1016/j.jnucmat.2010.12.193

[37] N. Castin, L. Malerba, R. Chaouadi, in press: doi:10.1016/j.jnucmat.2010.10.039

[38] N. Castin, M. I. Pascuet and L. Malerba, "Modeling the first stages of Cu precipitation in a-Fe using an hybrid atomistic kinetic Monte Carlo approach", submitted to the journal of chemical physics.

[39] C. Campbell, Neural Network Systems Technologies and Applications, Constructive Learning Techniques for Designing Neural Network Systems, edited by C. T. Leondes Academic, San Diego, 1997.

[40] R. Chaouadi, R. Gerard, J. Nucl. Mat. 345 (2005) 65.

[41] A. Chatterjee, D.G. Vlachos, J. Computer-Aided Mater. Des. 14 (2007) 253-308.

[42] H-H. Chen and M. T. Manry and H. Chandrasekaran, Neurocomputing 25 (1999) 55.

[43] F. Christien and A. Barbu, J. Nucl. Mater. 324 (2004) 90.

[44] A. Deschamps, C. Genevois, M. Nicolas, F. Perrard, and F. Bley, Philos. Mag. 85 (2005) 3091.

[45] C. Dimitrov, A. Benkaddour, C. Corbel, P. Moser, Annales De Chimie-Science Des Materiaux 16 (1991) 319.

[46] F. Djurabekova, R. Domingos, G. Cerchiara, N. Castin, E. Vincent, L. Malerba, Nucl. Instr. and Meth. Phys. Res. B 255 (2007) 8.

[47] C. Domain and C. S. Becquart, Phys. Rev. B 65 (2001) 024103

[48] C. Domain, C.S. Becquart, L. Malerba, J. Nucl. Mater. 335 (2004) 121.

[49] C. Domain and C. S. Becquart, Phys. Rev. B 71, 214109 (2005).

[50] S.M. Dubiel and G. Inden, Z. Metallkde. 78 (1987) 544.

[51] P. Dubuisson, D. Gilbon, J.L. Séran, J. Nucl. Mater. 205 (1993) 178.

[52] C. English, W. Phythian, R. McElroy, Mater. Res. Soc. Symp. Proc., vol. 439, MRS, Pittsburg, Pennsylvania, 1997, p. 471.

[53] S.E. Fahlman, "An Empirical Study of Learning Speed in Back-Propagation Networks", Technical Report, CMU-CS-88-162, 1988.

[54] S.E. Fahlman, "The recurrent Cascade-Correlation Architecture", in Advances in Neural Information Processing Systems, Richard P. Lippmann and John E. Moody and David S. Touretzky, 1991, pp 190-196.

[55] S.E. Fahlman and C. Lebiere, "The Cascade-Correlation Learning Architecture", in Advances in Neural Information Processing Systems, volume 2, pages 524-532, 1990.

[56] C. Fazio, A. Alamo, A. Almazouzi, S. De Grandis, D. Gomez-Briceno, J. Henry, L. Malerba and M. Rieth, European cross-cutting research on structural materials for Generation IV and transmutation systems, J. Nucl. Mater. (2009), in press, doi:10.1016/j.jnucmat.2009.03.020.

[57] K. A. Fichthorn, W. H. Weinberg, J. Chem. Phys., Vol. 95, No. 2, 15 July 1991.

[58] M.E. Fine, J.Z. Liu and M.D. Asta, Mater. Sci. Eng. A (2007) 271.

[59] R.M. Fisher, E.J. Dulis, K.G. Carroll, Trans. AIME 197 (1953) 690.

[60] C-C. Fu, F. Willaime, P. Ordejón, Phys. Rev. Lett. 92, 175503 (2004).

[61] C-C. Fu, J. Dalla Torre, F. Willaime, J-L. Bocquet and A. Barbu, Nature Materials 4 (2005) 68.

[62] K. Fukuya, K. Ohno, H. Nakata, S. Dumbill and J.M. Hyde, J. Nucl. Mater. 312 (2003) 163-173.

[63] F.A. Garner, M.B. Toloczko and B.H. Sencer, J. Nucl. Mater. 276 (2000) 123.

[64] S.I. Golubov, Y.N. Osetsky, A. Serra, and A. V. Barashev, J. Nucl. Mater. 226 (1995) 252.

[65] S.I. Golubov, A. Serra, Y. N. Osetsky, and A. V. Barashev, J. Nucl. Mater. 277 (2000) 113.

[66] S.R. Goodman, S.S. Brenner, J.R. Low, Metall. Trans. 4 (1973) 2363.

[67] S.R. Goodman, S.S. Brenner, J.R. Low, Metall. Trans. 4 (1973) 2371.

[68] P.J. Grobner, Metall. Trans. 4 (1973) 251.

[69] M.L. Grossbeck, "Effects of Radiation on Materials", ASTM international, 1982, DOI: 10.1520/STP782-EB

[70] S. Haykin, "Neural Networks: A Comprehensive Foundation", New York: MacMillian (1994).

[71] J.J. Heger, Metal Prog. (1951) 55.

[72] G. Henkelman, H. Jónsson, J. Chem. Phys. 111 (1999) 7010.

[73] G. Henkelman, G. Jóhannesson, and H. Jónsson, "Methods for Finding Saddle Points and Minimum Energy Paths", in Progress on Theoretical Chemistry and Physics, 269-300, Ed. S. D. Schwartz (Kluwer Academic Publishers, 2000).

[74] G. Henkelman and H. Jónsson, J. Chem. Phys. 113 (2000) 9901.

[75] G. Henkelman, H. Jonsson, J. Chem. Phys. 115 (2001) 9657.

[76] P. Hohenberg, W. Kohn, Phys. Rev. 136, B864–B871 (1964)

[77] E. Hornbogen and R. C. Glenn, Trans. Metall. Soc. AIME 218, 1064 (1960).

[78] K. Hornik, M. Stinchcombe and H. White, Neural Networks 2 (1989) 359.

[79] IAEA Nuclear Energy Series, "Integrity of reactor pressure vessels in nuclear power plants: assessment of irradiation embrittlement effects in reactor pressure vessel steels", IAEA Nuclear Energy Series, No. NP-T-3.11, 2009.

[80] D. Isheim, M. S. Gagliano, M. E. Fine, and D. N. Seidman, Acta Mater. 54, 841 (2006).

[81] V. Jacquet, "Effet de l'irradiation sur la demixtion des alliages modeles Fe–Cr autour de 15% de Cr," Ph.D. thesis, Ecole Polytechnique, 2000.

[82] P. Jacobsson, Y. Bergström and B. Aronsson, Metall. Trans. 6A (1977) 1577.

[83] R. Johnson, Physical Review 134 (1964) A1329.

[84] H. Jónsson, G. Mills, K. W. Jacobsen, in: Classical and quantum dynamics in condensed phase simulations, B. J. Berne, G. Ciccotti and D. F. Coker Eds.,World Scientific, Singapore, 1998.

[85] J. Kameda and A.J. Bevolo, Acta Metall 37 (1989) 3283.

[86] R. Kampmann, R. Wagner, in: C. Janot, W. Petry, D. Richter, T. Springer (Eds.), Atomic Transport and Defects in Metals by Neutron Scatterings, Springer, New York, 1986. p. 73.

[87] H.C. Kang and W.H. Weinberg, J. Chem. Phys. 90 (1989) 2824.

[88] H. Kayano, A. Kimura, M. Narui, Y. Sasaki, Y. Suzuki and S. Ohta, J. Nucl. Mater. 155-157 (1988) 978.

[89] A. Kohyama, A. Hishinuma, D.S. Gelles, R.L. Klueh, W. Dietz and K. Ehrlich, J. Nucl. Mater. 233-237 (1996) 138.

[90] Yu.V. Konobeev, A.M. Dvoriashin, S.I. Porollo and F.A. Garner, J. Nucl. Mater. 355 (2006) 124.

[91] P. Krasnochtchekov, R. S. Averback, and P. Bellon, Phys. Rev. B 75, 144107 (2007).

[92] O. Kubaschewski, Iron Binary Phase Diagrams, Springer, Berlin, 1982.

[93] H. Kuwano, Trans. Japan Inst. Met. 26 (1985) 473.

[94] T.-Y. Kwok and D.-Y. Yeung, IEEE Trans. Neural Netw. 8 (1997) 630.

[95] R. Lagneborg, Trans. ASM 60 (1967) 67.

[96] M.Yu. Lavrentiev, R. Drautz, D. Nguyen-Manh, T.P.C. Klaver, S.L. Dudarev, Phys. Rev. B 75 (2007) 014208.

[97] Y. Le Bouar, F. Soisson, Phys. Rev. B 65 (2002) 094103.

[98] K. Levenberg, Q. Appl. Math. 2, 164 (1944).

[99] E.A. Little and D.A. Stow, J. Nucl. Mater. 87 (1979) 25.

[100] E.A. Little and D.A. Stow, J. Metal Sci. 14 (1980) 89.

[101] S. Lozano-Perez, M. L. Jenkins, and J. M. Titchmarsh, Philos. Mag. Lett. 86, 367 (2006).

[102] J. Marian, B.D. Wirth, A. Caro, B. Sadigh, G.R. Odette, J.M. Perlado, and T. Diaz de la Rubia, Phys. Rev. B 65 (2002) 144102.

[103] D. Marquardt, SIAM J. Appl. Math. 11, 431 (1963).

[104] G. Martin, P. Bellon, and F. Soisson, Defect Diffus. Forum 143–147, 385 (1997).

[105] E. Martinez, J. Marian, M.H. Kalos, J.M. Perlado, Journ. Comput. Phys. 227 (2008) 3804.

[106] D.R. Mason, R.E. Rudd and A.P. Sutton, J. Phys.: Condes. Matter 16 (2004) S2679.

[107] T. B. Massalsky, H. Okamoto, P. R. Subramanian, and L. Kacprzac, Binary Alloy Phase Diagrams ASM International, Materials Park, 1990, p. 1273.

[108] J.P. Massoud, S. Bugat, J.L. Boutard, D. Lidbury, S. Van Dyck, F. Sevini, Prediction of irradiation damage effects on reactor components (PERFECT), in FISA 2006 EU Research and Training in Reactor Systems, G. Van Goethem, P. Manolatos, M. Hugon, V. Bhatnagar, S. Casalta and M. Deffrennes, eds., Directorate-General for Research, Euratom, Luxembourg, 2006.

[109] M.H. Mathon, A. Barbu, F. Dunstetter, F. Maury, N. Lorenzelli, C.H. de Novion, J. Nucl. Mater. 245 (1997) 224.

[110] M.H. Mathon, Y. de Carlan, G. Geoffroy, X. Averty, C.H. de Novion and A. Alamo, Effects of Radiation on Materials: 20th International Symposium, Eds. S.T. Rosinski et al., ASTM STP 1405, American Society for Testing and Materials, West Conshohocken, PA, 2001, p. 674.

[111] M.H. Mathon, Y. de Carlan, G. Geoffroy, X. Averty, A. Alamo, C.H. de Novion, J. Nucl. Mater. 312 (2003) 236.

[112] M.H. Mathon, Y. de Carlan, X. Averty, A. Alamo and C.H. de Novion, J. ASTM Int. 2 (2005) JAI12381.

[113] F. El-Mellouhi, N. Mousseau, and L. J. Lewis, Phys. Rev. 78, 153202 (2008).

[114] M. K. Miller, K. F. Russel, P. Pareige, M. J. Starink, and R. C. Thomson, Mater. Sci. Eng. A 250, 49 (1998).

[115] M.K. Miller and P. Pareige, Mater. Res. Soc. Symp. Proc. 650, R6.1.1 (2001).

[116] K. Morita, S. Ishino, T. Tobita, Y. Chimi, N. Ishikawa, A. Iwase, J. Nucl. Mater. 304 (2002) 153.

[117] A. Morley, G. Sha, S. Hirosawa et al, Ultramicroscopy 109 (2009) 535

[118] J. Moteff, "Effects of radiation on substructure and mechanical properties of metals and alloys", ASTM international, 1973 ,DOI: 10.1520/STP529-EB

[119] N. Mousseau and G. T. Barkema, Phys. Rev. E 57, 2419 (1998).

[120] Y. Nagai, Z. Tang, M. Hasegawa, T. Kanai and M. Saneyasu, Phys. Rev. B 63 (2001) 134110.

[121] A.L. Nikolaev, V.L. Arbuzov, A.E. Davletshin, Journal of Physics: Condensed Matter 9 (1997) 4385.

[122] S. Novy, P. Pareige, and C. Pareige, J. Nucl. Mater. 384, 96 (2009).

[123] D. Nguyen and B. Widrow, "Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights", Proceedings of the International Joint Conference on Neural Networks, 3:21–26, 1990.

[124] D. Nguyen-Manh, S. Dudarev, A. Horsfield, J. Nucl. Mater. 257 (2007) 367.

[125] G. Odette, Scr. Metall. 17 (1983) 1183.

[126] G.R. Odette, G.E. Lucas, JOM 53 (7) (2001) 18–22.

[127] P.R. Okamoto, L.E. Rehn, J. Nucl. Mater. 83 (1979) 2.

[128] P. Olsson, J. Wallenius, C. Domain, K. Nordlund, and L. Malerba, Phys. Rev. B 72, 214119 (2005).

[129] P. Olsson, C. Domain, and J. Wallenius, Phys. Rev. B 75, 014110 (2007).

[130] P.J. Orthens, M.L. Jenkins, G.D.W. Smith, Phil. Mag. A 70 (1994) 1.

[131] C. Ortiz, M.-J. Caturla, C.-C. Fu and F. Willaime, Phys. Rev. B 75, 100102(R) (2007).

[132] C. Ortiz and M.-J. Caturla, Phys. Rev. B 75, 184101 (2007).

[133] K. Osamura, H. Okuda, M. Takashima, K. Asano, M. Furusaka, Mater. Trans., JIM 34 (1993) 305.

[134] Y.N. Osetsky, D. J. Bacon, and A. Serra, Philos. Mag. Lett. 79 (1999) 273.

[135] R. Parekh, J. Yang, V. Honavar, IEEE Trans. Neural Netw. 11 (2)(2000)

[136] P. Pareige, J.C. Van Duysen and P. Auger, Appl. Surf. Sci. 67 (1993) 342.

[137] P. Pareige and M.K. Miller, Appl. Surf. Sci. 94/95 (1996) 370.

[138] M. I. Pascuet, L. Malerba, N. Castin, "Stability and mobility of small and large Cu-vacancy clusters in FeCu alloys using an artificial intelligence driven atomistic kinetic Monte Carlo method", external report of the Belgian nuclear research centre, SCK-CEN ER-93, May 2009.

[139] M.I. Pascuet, N. Castin, C.S. Becquart, L. Malerba, "Stability and mobility of Cu-vacancy clusters in Fe–Cu alloys: A computational study based on the use of artificial neural networks for energy barrier calculations", in press: doi:10.1016/j.jnucmat.2011.02.038

[140] R.C. Pasianot and L. Malerba, J. Nucl. Mater. 360 (2007) 118.

[141] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. 64, 1045 (1992).

[142] M. Perez, F. Perrard, V. Massardier, et al., Philos. Mag. 85 (2005) 2197.

[143] P. Petrequin, "A Review of Formulas for Predicting Irradiation Embrittlement of Reactors Vessel Materials", AMES Report No. 6, November 1996.

[144] W. Phythian, C. English, J. Nucl. Mater. 205 (1993) 162.

[145] S. Pizzini, K.J. Roberts, W.J. Phythian, C.A. English, G.N. Greaves, Philos. Mag. Lett. 61 (1990) 223.

[146] S.I. Porollo, A.M. Dvoriashin, A.N. Vorobyev and Yu.V. Konobeev, J. Nucl. Mater. 256 (1998) 247.

[147] U. Potapovs and J.R. Hawthorne, Nucl. Appl. 1 (1969) 27.

[148] V.P. Ramunni, M.A. Alurralde, and R.C. Pasianot, Phys. Rev. B 74, 054113 (2006).

[149] R. Reed, IEEE Trans. Neural Netw. 4 (1993) 740.

[150] L.E. Rehn, P.R. Okamoto, in: F.V. Nolfi (Ed.), Phase Transformations During Irradiation , Applied Science Publishers, London and New York, 1983, p. 247, Chapter 1.

[151] M. Riedmiller and H. Braun, "A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm", In Proceedings of the IEEE International Conference on Neural Networks (ICNN) (Ruspini, H. ed), p. 586-591, 1993, San Francisco.

[152] M.T. Robinson, I.M. Torrens, Phys. Rev. B 9 (1974) 5008.

[153] D.E. Rumelhart, G.E. Hinton, R.J. Williams, "Learning internal representations by error propagation", In D.E. Rumelhart, J.L. McClelland, and the PDP Research Group (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1, pp. 318-362, Cambridge, MA: MIT Press, 1986.

[154] K.C. Russel, Prog. Mater. Sci. 28 (1984) 229.

[155] J.M. Sanchez, F. Ducastelle and D. Gratias, Pysica A 128 (1984) 334.

[156] Sarle, W.S., ed. (1997), Neural Network FAQ, part 1 of 7: Introduction, periodic posting to the Usenet newsgroup comp.ai.neural-nets, URL: ftp://ftp.sas.com/pub/neural/FAQ.html

[157] K. Sastry, D. D. Johnson, D. E. Goldberg, P. Bellon, Phys. Rev. B 72 (2005) 085438.

[158] S. Schmauder, P. Binkele, Comput. Mater. Sci. 25 (2002) 174.

[159] R. Smith, "Atomic and ion collisions in solids and at surfaces: theory, simulation and applications", Cambridge University Press, 1997.

[160] F. Soisson, A. Barbu, and G. Martin, Acta Mater. 44, 3789 (1996).

[161] F. Soisson, Philos. Mag. 85, 489 (2005).

[162] F. Soisson, J. Nucl. Mater. 349 (2006) 235.

[163] F. Soisson and C-C Fu, Phys. Rev. B 76, 214102 (2007).

[164] F. Soisson, C.S. Becquart, N. Castin, C. Domain, L. Malerba, E. Vincent, J. Nucl. Mater. 406 (2010) 55.

[165] N. Soneda and T. Diaz de la Rubia, Philos. Mag. A 78 (1998) 995.

[166] A.P. Sutton, "Electronic structure of materials", Oxford University Press, 1993, pages 204-209

[167] D. A. Terentyev, T. P. C. Klaver, P. Olsson, M. C. Marinica, F. Willaime, C. Domain, and L. Malerba, Phys. Rev. Lett. 100, 145503 (2008).

[168] D. A. Terentyev, G. Bonny, and L. Malerba, Acta Mater. 56 (2008) 3229.

[169] D. Terentyev and N. Castin, in press: doi:10.1016/j.commatsci.2009.06.004

[170] D. Terentyev, N. Juslin, K. Nordlund, and N. Sandberg, J. Appl. Phys. 105 (2009) 103509.

[171] D. Terentyev et al., J. Nucl. Mater. 409 (2011) 167.

[172] O. Trushin, A. Karim, A. Kara, and T. S. Rahman, Phys. Rev. B 72, 115401 (2005).

[173] A. Van der Ven, G. Ceder, Phys. Rev. B 71, Issue 5, (2005).

[174] E. Vincent, C. S. Becquart, and C. Domain, Nucl. Instrum. Methods Phys. Res. B 228, 137 (2005).

[175] E. Vincent, C. S. Becquart, and C. Domain, J. Nucl. Mater. 359, 227 (2006).

[176] E. Vincent, C.S. Becquart, C. Domain, Nucl. Instr. Meth. Phys. Res. B 255 (2007) 78-84.

[177] E. Vincent, C.S. Becquart, C. Domain, J. Nucl. Mater. 382 (2008) 154-159.

[178] E. Vincent, C.S. Becquart, C. Pareige, P. Pareige, C. Domain, J. Nucl. Mater. 373 (2008) 387-401.

[179] G.H. Vineyard, J. Phys. Chem. Solids 3, 121 (1957).

[180] A. F. Voter, in: Radiation Effects in Solids, edited by K. E. Sickafus et al. (Springer, 2006), p. 1.

[181] W. Wagner, L.E. Rehn, H. Wiedersich, V. Naundorf, Phys. Rev. B 28 (1983) 6780.

[182] H. Wiedersich, P.R. Okamoto, N.Q. Lam, J. Nucl. Mater. 83 (1979) 98.

[183] H. Wiedersich, N.Q. Lam, in: F.V. Nolfi (Ed.), Phase Transformations During Irradiation, Applied Science Publishers, London and New York, 1983, p. 1, Chapter 1.

[184] B.M. Wilamowski and S. Iplikci and O. Kaynak and M.O. Efe, Proceedings of the International Joint Conference on Neural Networks (IJCNN), 2001, pp 1778-1782.

[185] R. O. Williams and H. W. Paxton, J. Iron Steel Inst. 185 (1957) 358.

[186] R. O. Williams, Trans. Metall. Soc. AIME 212 (1958) 497.

[187] B.D. Wirth and G.R. Odette, "Kinetic lattice Monte Carlo simulations of cascade aging in iron and dilute ironcopper alloys", Mat. Res. Soc. Symp. Proc. 540 (1999) 637-642.

[188] B.D. Wirth, G.R. Odette, D. Maroudas, and G. E. Lucas, J. Nucl. Mater. 276 (2000) 33.

[189] G. M. Worrall, J. T. Buswell, C. A. English, M. G. Hetherington, and G. D. W. Smith, J. Nucl. Mater. 148, 107 (1987).

[190] W.M. Young and E.W. Elcock, Proc. Phys. Soc. 89 (1966) 735.

[191] Z. Zainuddin, N. Mahat, and Y. Abu Hassan, World Academy of Science, Engineering and Technology 1 (2005) 79.

[192] E. E. Zhurkin, R. Pereira, N. Castin, L. Malerba, and M. Hou, Mater. Res. Soc. Symp. Proc. 1125 (2009) 121.

[193] This report is apparently not officially accessible anymore, but a summary of it can be found at: http://www.ulb.ac.be/sciences/intra/inforsc_-archives/nrj/ampere/ampere1.html

[194] http://www.efda.org/

[195] http://www.ce2030.be/index.htm

[196] http://www.areva.com/FR/activites-1572/les-atouts-du-reacteur-nucleaire-epr.html

[197] http://www.euratom.org/

[198] http://www.f-bridge.eu/index.php/Project-Description/Objectives.html

[199] http://economie.fgov.be/fr/binaries/rapport_gemix_2009_fr_tcm326-76356.pdf

[200] http://nuklear-server.ka.fzk.de/getmat/

[201] http://unfccc.int/kyoto_protocol/items/2830.php

[202] http://www.mcs.anl.gov/research/projects/mpi/

[203] http://www.open-mpi.org/

[204] https://www.fp6perfect.net/

[205] www.perform60.net/

[206] http://setis.ec.europa.eu/technologies

[207] http://www.snetp.eu/www/snetp/

A Diffusion coefficients and lifetimes for Cu-Vac clusters

This section gives the complete sets of diffusion coefficients and lifetimes that were calculated for vacancy clusters and Cu-Vac clusters in pure Fe matrices. From the diffusion coefficients D. fits with Arrhenius functions provide a value for exponential prefactor D_0 and the migration energy E_m^D:

$$D(T) = D_0 \cdot \exp\left(\frac{-E_m^D}{k_B T}\right) \tag{A.1}$$

From the lifetimes ν. fits with Arrhenius functions provide a value for the dissolution attempt frequency ν_0 and the dissociation energy E_m^ν:

$$\nu(T) = \nu_0 \cdot \exp\left(\frac{-E_m^\nu}{k_B T}\right) \tag{A.2}$$

N	M	500K	560K	640K	750K	900K	D_0	E_m^D
3	1	7.00E-11	3.50E-10	2.00E-09	1.00E-08	8.00E-08	4.07E-04	0.76
3	2	1.00E-11	7.00E-11	4.00E-10	2.50E-09	3.00E-08	4.06E-04	0.81
3	3	3.00E-12	2.00E-11	1.40E-10	1.30E-09	1.35E-08	1.23E-04	0.79
3	4	1.50E-12	9.00E-12	7.00E-11	6.00E-10	5.00E-09	5.70E-05	0.78
3	5	1.00E-12	5.00E-12	4.00E-11	3.00E-10	3.00E-09	6.50E-05	0.81
3	6	5.36E-13	3.00E-12	3.00E-11	2.50E-10	2.00E-09	2.95E-05	0.79
3	7	3.52E-13	2.40E-12	1.60E-11	1.30E-10	1.30E-09	2.96E-05	0.82
3	8	1.70E-13	1.50E-12	1.00E-11	1.00E-10	8.00E-10	5.50E-05	0.87
3	9	9.00E-14	9.00E-13	7.00E-12	7.00E-11	8.00E-10	4.07E-04	0.76
4	1	5.00E-12	7.00E-11	5.00E-10	4.50E-09	3.20E-08	1.82E-03	0.84
4	2	2.00E-12	1.60E-11	1.60E-10	1.60E-09	2.50E-08	2.44E-03	0.91
4	3	1.50E-12	7.00E-12	5.00E-11	4.00E-10	6.00E-09	1.22E-04	0.80
4	4	9.00E-13	4.50E-12	3.70E-11	4.00E-10	6.00E-09	2.65E-04	0.85
4	5	5.00E-13	3.00E-12	2.00E-11	2.50E-10	3.50E-09	1.60E-04	0.86
4	6	2.72E-13	2.50E-12	2.50E-11	1.50E-10	1.40E-09	5.47E-05	0.82
4	7	2.00E-13	1.10E-12	9.00E-12	1.00E-10	1.50E-09	7.65E-05	0.86
4	8	8.00E-14	8.00E-13	6.00E-12	7.00E-11	1.00E-09	9.33E-05	0.90
4	9	4.50E-14	5.00E-13	5.00E-12	4.00E-11	6.00E-10	6.02E-05	0.90
4	12		1.50E-13	1.75E-12	2.30E-11	2.80E-10	6.82E-05	0.96

Table A.1: Diffusion coefficients in cm^2/s calculated with 100 independent AKMC simulations for varying number N of vacancies and number M of Cu atoms in the clusters and varying temperature. The two last columns show the exponential prefactor D_0 (in cm^2/s) and the migration energy E_m (eV) fitted in Arrhenius plots.

N	M	640K	750K	900K	1125K	1500K	D_0	E_m^D
5	2	3.50E-11	4.30E-10	7.00E-09		1.12E-06	1.96E-03	0.98
5	3	1.75E-11	2.25E-10	3.50E-09		6.80E-07	1.26E-03	0.99
5	4	1.75E-11	1.60E-10	2.50E-09		4.00E-07	4.92E-04	0.95
5	6	6.00E-12	8.00E-11	1.00E-09	2.00E-08	1.70E-07	4.63E-04	1.00
5	9	2.00E-12	2.75E-11	4.00E-10	5.00E-09	7.40E-08	1.79E-04	1.01
5	12	8.00E-13	1.30E-11	2.30E-10	2.00E-09	3.00E-08	7.44E-05	1.01
5	15	3.44E-13	8.00E-12	1.20E-10	1.40E-09	2.30E-08	8.39E-05	1.06
6	2	1.20E-11	8.00E-11	1.20E-09	3.30E-08	9.00E-07	2.85E-03	1.10
6	6	3.00E-12	3.50E-11	4.50E-10	7.50E-09	1.70E-07	4.40E-04	1.05
6	10	3.70E-13	8.16E-12	1.70E-10	2.30E-09	5.30E-08	3.23E-04	1.13
6	14	1.30E-13	4.00E-12	9.00E-11	1.50E-09	3.00E-08	3.04E-04	1.18
6	18	5.25E-14	1.50E-12	3.50E-11	1.00E-09	1.70E-08	2.65E-04	1.23
7	2	1.00E-11	1.00E-10	1.00E-09	1.80E-08	5.00E-07	9.69E-04	1.03
7	6	2.00E-12	2.50E-11	3.00E-10	5.50E-09	1.00E-07	2.65E-04	1.04
7	11	1.50E-13	1.00E-11	8.00E-11	2.30E-09	4.50E-08	4.43E-04	1.18
7	16	1.50E-13	3.50E-12	3.00E-11	8.00E-10	1.80E-08	7.91E-05	1.11
7	21		1.00E-12	2.50E-11	5.00E-10	1.20E-08	1.37E-04	1.21
8	3	3.30E-12	3.75E-11	5.25E-10	9.00E-09	2.00E-07	5.69E-04	1.06
8	8	6.00E-13	9.00E-12	1.50E-10	2.40E-09	4.00E-08	1.53E-04	1.07
8	14	1.25E-13	3.11E-12	5.00E-11	9.00E-10	1.50E-08	9.19E-05	1.12
8	19		1.00E-12	2.50E-11	4.00E-10	1.00E-08	8.75E-05	1.18
8	24		5.80E-13	1.30E-11	4.00E-10	7.30E-09	1.10E-04	1.23

Table A.2: Diffusion coefficients in cm^2/s calculated with 100 independent AKMC simulations for varying number N of vacancies and number M of Cu atoms in the clusters and varying temperature. The two last columns show the exponential prefactor D_0 (in cm^2/s) and the migration energy E_m (eV) fitted in Arrhenius plots.

N	M	640K	750K	900K	1125K	1500K	D_0	E_m^D
9	3	3.40E-12	4.40E-11	4.70E-10	6.00E-09	1.00E-07	1.71E-04	0.98
9	7	8.00E-13	1.50E-11	1.50E-10	2.25E-09	4.50E-08	1.15E-04	1.04
9	12	2.70E-13	3.20E-12	5.80E-11	1.10E-09	2.00E-08	8.21E-05	1.09
9	17		1.30E-12	3.00E-11	6.00E-10	1.00E-08	8.39E-05	1.16
9	22		9.00E-13	1.50E-11	3.10E-10	7.00E-09	5.14E-05	1.16
9	27		4.00E-13	1.50E-11	2.50E-10	5.00E-09	6.30E-05	1.21
10	3	2.00E-12	3.00E-11	1.00E-09	4.40E-09	1.00E-07	2.75E-04	1.03
10	4	1.40E-12	2.00E-11	2.90E-10	4.10E-09	7.00E-08	1.99E-04	1.04
10	6	8.10E-13	1.10E-11	1.60E-10	2.50E-09	5.00E-08	1.54E-04	1.06
10	8	3.80E-13	6.80E-12	1.10E-10	1.20E-09	2.40E-08	7.53E-05	1.05
10	10	3.45E-13	4.92E-12	7.91E-11	1.13E-09	1.92E-08	6.14E-05	1.05
10	15	9.10E-14	1.78E-12	3.01E-11	6.59E-10	1.26E-08	8.22E-05	1.14
10	20		1.10E-12	1.88E-11	3.62E-10	7.79E-09	5.24E-05	1.15
10	25		5.33E-13	1.08E-11	2.29E-10	5.84E-09	5.91E-05	1.20
10	30		4.37E-13	8.38E-12	1.78E-10	2.44E-09	1.62E-05	1.12
11	4	1.06E-11	1.06E-11	1.66E-10	2.89E-09	5.20E-08	2.33E-05	0.87
11	5		8.74E-12	1.30E-10	1.82E-09	3.02E-08	9.80E-05	1.05
11	7		4.14E-12	8.15E-11	1.58E-09	2.83E-08	2.00E-04	1.14
11	9		3.33E-12	5.45E-11	9.89E-10	2.00E-08	1.14E-04	1.12
11	11		2.58E-12	5.39E-11	9.39E-10	1.65E-08	1.07E-04	1.13
11	17		1.00E-12	2.12E-11	4.32E-10	8.89E-09	7.88E-05	1.17
11	22		5.93E-13	1.78E-11	3.26E-10	7.47E-09	9.05E-05	1.21
11	28		3.96E-13	9.22E-12	1.58E-10	3.70E-09	3.14E-05	1.17
11	33			6.19E-12	1.44E-10	5.03E-09	1.09E-04	1.30

Table A.3: Diffusion coefficients in cm^2/s calculated with 100 independent AKMC simulations for varying number N of vacancies and number M of Cu atoms in the clusters and varying temperature. The two last columns show the exponential prefactor D_0 (in cm^2/s) and the migration energy E_m (eV) fitted in Arrhenius plots.

N	M	500K	560K	640K	750K	900K	ν_0	E_m^v
3	1	1.41E-05	2.21E-06	3.25E-07	3.98E-08	4.49E-09	2.21E-13	0.78
3	2	1.26E-04	1.11E-05	1.35E-06	1.14E-07	8.53E-09	6.78E-14	0.92
3	3	4.88E-04	4.08E-05	3.34E-06	2.08E-07	2.01E-08	5.72E-14	0.98
3	4	1.47E-03	1.26E-04	6.48E-06	4.95E-07	2.29E-08	2.65E-14	1.07
3	5	2.56E-03	1.88E-04	1.47E-05	5.84E-07	3.83E-08	3.48E-14	1.08
3	6	4.33E-03	2.68E-04	1.98E-05	8.47E-07	5.53E-08	4.16E-14	1.09
3	7	5.98E-03	4.51E-04	2.61E-05	1.51E-06	6.12E-08	4.55E-14	1.11
3	8	5.70E-03	6.75E-04	2.93E-05	1.76E-06	1.02E-07	1.04E-13	1.08
3	9	5.86E-03	7.24E-04	4.54E-05	2.16E-06	1.03E-07	1.29E-13	1.07
4	1	3.76E-04	3.76E-05	1.97E-06	1.39E-07	1.34E-08	2.67E-14	1.01
4	2	1.94E-03	1.17E-04	5.95E-06	4.28E-07	2.31E-08	1.74E-14	1.09
4	3	5.06E-03	3.28E-04	1.57E-05	9.58E-07	3.61E-08	1.68E-14	1.14
4	4	6.51E-03	4.95E-04	2.64E-05	1.01E-06	4.36E-08	1.58E-14	1.16
4	5	1.28E-02	5.99E-04	4.48E-05	1.65E-06	7.54E-08	2.67E-14	1.16
4	6	1.42E-02	1.00E-03	5.89E-05	1.99E-06	8.63E-08	2.81E-14	1.17
4	7	2.19E-02	1.11E-03	8.22E-05	2.46E-06	1.06E-07	2.86E-14	1.18
4	8	2.31E-02	1.92E-03	9.57E-05	3.35E-06	1.34E-07	4.02E-14	1.18
4	9	2.79E-02	1.99E-03	1.41E-04	4.57E-06	1.51E-07	5.46E-14	1.17
4	12		6.00E-03	2.13E-04	7.84E-06	2.68E-07	1.98E-14	1.28

Table A.4: Lifetimes in s calculated with 100 independent AKMC simulations for varying number N of vacancies and number M of Cu atoms in the clusters and varying temperature. The two last columns show the lifetime prefactor ν_0 (in s) and the dissociation energy E_m^v (eV) obtained to by fitting data to Arrhenius plots.

N	M	640K	750K	900K	1125K	1500K	ν_0	E_m^v
5	2	4.62E-05	1.84E-06	6.91E-08			3.60E-15	1.30
5	3	9.59E-05	3.92E-06	1.13E-07		1.15E-10	8.53E-15	1.27
5	4	9.94E-05	4.17E-06	1.17E-07		1.31E-10	8.04E-15	1.28
5	6	1.44E-04	6.40E-06	2.31E-07	6.52E-09	2.53E-10	1.24E-14	1.29
5	9	3.47E-04	1.15E-05	3.73E-07	9.98E-09	3.23E-10	1.03E-14	1.34
5	12	5.51E-04	1.87E-05	4.26E-07	1.38E-08	4.58E-10	1.20E-14	1.36
5	15	1.07E-03	2.03E-05	7.16E-07	1.56E-08	3.85E-10	6.80E-15	1.42
6	2	5.87E-04	1.92E-05	3.74E-07	6.47E-09	1.39E-10	1.57E-15	1.48
6	6	9.20E-04	2.75E-05	5.43E-07	1.42E-08	2.84E-10	4.34E-15	1.45
6	10	2.75E-03	5.32E-05	1.03E-06	1.90E-08	4.33E-10	3.43E-15	1.51
6	14	4.85E-03	7.45E-05	1.43E-06	2.52E-08	5.79E-10	3.63E-15	1.54
6	18	4.39E-03	1.07E-04	1.73E-06	2.65E-08	5.73E-10	3.76E-15	1.54
7	2	7.63E-04	1.83E-05	4.81E-07	9.54E-09	2.00E-10	2.75E-15	1.46
7	6	2.01E-03	6.61E-05	1.27E-06	3.25E-08	5.21E-10	7.80E-15	1.46
7	11	1.02E-03	3.40E-05	6.81E-07	2.07E-08	3.63E-10	6.82E-15	1.43
7	16	3.48E-03	1.04E-04	2.25E-06	4.57E-08	7.08E-10	9.07E-15	1.49
7	21		2.03E-04	2.35E-06	4.85E-08	8.57E-10	3.64E-15	1.59
8	3	1.29E-03	4.04E-05	8.56E-07	1.39E-08	3.51E-10	4.11E-15	1.47
8	8	2.69E-03	6.80E-05	1.49E-06	2.91E-08	5.70E-10	6.54E-15	1.48
8	14	4.48E-03	1.30E-04	2.30E-06	4.57E-08	9.60E-10	1.00E-14	1.49
8	19		1.96E-04	3.31E-06	5.70E-08	1.04E-09	5.35E-15	1.57
8	24		2.04E-04	4.72E-06	6.57E-08	1.04E-09	5.28E-15	1.58

Table A.5: Lifetimes in s calculated with 100 independent AKMC simulations for varying number N of vacancies and number M of Cu atoms in the clusters and varying temperature. The two last columns show the lifetime prefactor ν_0 (in s) and the dissociation energy E_m^v (eV) obtained to by fitting data to Arrhenius plots.

N	M	640K	750K	900K	1125K	1500K	ν_0	E_m^v
9	3	1.89E-03	3.77E-05	9.27E-07	2.73E-08	5.17E-10	7.96E-15	1.44
9	7	2.87E-03	7.11E-05	1.95E-06	2.28E-08	6.84E-10	6.80E-15	1.49
9	12	3.67E-03	1.10E-04	2.53E-06	5.19E-08	8.08E-10	1.12E-14	1.48
9	17		1.45E-04	2.84E-06	6.12E-08	1.06E-09	8.37E-15	1.52
9	22		1.96E-04	3.68E-06	6.73E-08	1.14E-09	6.84E-15	1.56
9	27			5.09E-06	1.02E-07	1.21E-09	4.85E-15	1.62
10	3	2.93E-03	8.25E-05	1.73E-06	3.02E-08	5.89E-10	6.33E-15	1.49
10	4	3.87E-03	8.16E-05	1.52E-06	3.73E-08	5.93E-10	5.80E-15	1.51
10	6	4.34E-03	9.51E-05	2.21E-06	4.07E-08	8.18E-10	8.51E-15	1.49
10	8	7.05E-03	1.27E-04	2.09E-06	4.49E-08	8.78E-10	6.10E-15	1.53
10	10	7.40E-03	1.58E-04	2.25E-06	5.06E-08	9.35E-10	6.56E-15	1.53
10	15	1.29E-02	1.94E-04	3.73E-06	7.48E-08	1.24E-09	8.09E-15	1.55
10	20		2.84E-04	3.97E-06	7.42E-08	1.13E-09	4.72E-15	1.60
10	25		2.94E-04	5.80E-06	8.74E-08	1.40E-09	6.66E-15	1.59
10	30		5.13E-04	7.80E-06	1.29E-07	1.76E-09	6.47E-15	1.62
11	4	1.43E-04	1.43E-04	2.33E-06	5.68E-08	7.46E-10	1.33E-13	1.24
11	5		1.59E-04	2.46E-06	5.76E-08	8.25E-10	5.01E-15	1.56
11	7		1.95E-04	3.60E-06	5.14E-08	8.35E-10	3.49E-15	1.60
11	9		2.29E-04	3.10E-06	6.97E-08	1.13E-09	6.03E-15	1.57
11	11		2.30E-04	4.32E-06	7.75E-08	1.18E-09	6.47E-15	1.57
11	17		2.11E-04	5.06E-06	7.54E-08	1.14E-09	6.38E-15	1.57
11	22		2.43E-04	6.14E-06	1.03E-07	1.44E-09	9.56E-15	1.56
11	28		2.10E-04	8.65E-06	1.51E-07	1.56E-09	1.56E-14	1.53
11	33			1.03E-05	1.38E-07	1.58E-09	3.10E-15	1.70

Table A.6: Lifetimes in s calculated with 100 independent AKMC simulations for varying number N of vacancies and number M of Cu atoms in the clusters and varying temperature. The two last columns show the lifetime prefactor ν_0 (in s) and the dissociation energy E_m^v (eV) obtained to by fitting data to Arrhenius plots.

B Paper I

Use of computational intelligence for the prediction of vacancy migration energies in atomistic kinetic monte carlo simulations

Nicolas Castin[*]

Physique des Solides Irradiés et des Nanostructures (PSIN), Université Libre de Bruxelles (ULB), boulevard du Triomphe CP234, 1050 Brussels, Belgium.

Roberto Pinheiro Domingos

European Centre for Soft Computing, Calle Gonzalo Gutiérrez Quiróz s/n
Mieres, Asturias 33600, Spain

Lorenzo Malerba

NMS, SCK•CEN, Boeretang 200
B2400, Mol, Belgium
E-mail: lmalerba@sckcen.be
www.sckcen.be

Abstract

In this work, we try to build a regression tool to partially replace the use of CPU-time consuming atomic-level procedures for the calculation of point-defect migration energies in Atomistic Kinetic Monte Carlo (AKMC) simulations, as functions of the Local Atomic Configuration (LAC). Two approaches are considered: the Cluster Expansion (CE) and the Artificial Neural Network (ANN). The first is found to be unpromising because of its high computational complexity. On the contrary, the second provides very encouraging results and is found to be very well behaved.

Keywords: Neural Networks, Fuzzy Logic, Cluster Expansion, Vacancy Migration Energy

1. Introduction

Phase transformation in general, and solute precipitation in particular, are spontaneous physical phenomena that may occur during operation in structural materials, e.g. steels, and dramatically modify their mechanical properties, thereby threatening the safety of the affected component. Models reliably describing the kinetics of these phenomena are therefore of importance for the safe exploitation of industrial nuclear power plants. For example, the formation of copper-rich precipitates and nanovoids under neutron irradiation is widely accepted

to be the main cause of hardening and embrittlement of nuclear Reactor Pressure Vessel (RPV) steels during operation[1], as a consequence of their acting as obstacles to dislocation motion. Experimental evidences (see e.g. Refs. 2-4) have highlighted that any model for the prediction of RPV steel hardening versus radiation dose (which is the basic requirement for the RPV lifetime assessment) needs to be able to account as correctly as possible for the build-up of Cu precipitate and Cu-vacancy complex density.

---

[*]Contact address : Studie Centrum voor KerneEnergie•Centre d'études de l'énergie nucléaire (SCK•CEN), Boeretang 200, B2400, Mol, Belgium. ncastin@sckcen.be
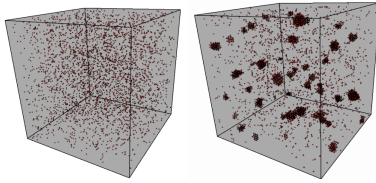
N. Castin et al.

Fig. 1. AKMC simulation box with about 700,000 atoms (1.4% Cu). (left) Initial state with the copper atoms randomly distributed. (right) An intermediate state (after billions of AKMC steps).
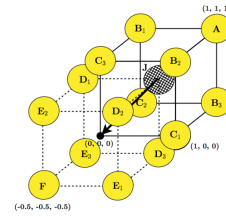

Fig. 2. One AKMC step corresponding to the migration of one vacancy (in reality, it is rather one of its neighbouring atoms that migrates to it). The figure shows the migration of the hatched atom to the vacancy, situated on the front bottom left corner of its BCC cubic cell, in plain lines. The dashed line shows the BCC cell of destination.

Atomistic Kinetic Monte Carlo (AKMC) simulations[5-11] are among the best suited tools for studying the precipitation of Cu in Fe via a vacancy migration mechanism, as a subpart of the general study of RPV steels long-term evolution after decades of operation. AKMC is a compromise between Molecular Dynamics (MD), that considers events at the atomic time and length scale, and coarse-grained tools, such as Object KMC[12] and rate theory[13-14], that are necessary to extend the simulation to the macroscopic scale. AKMC techniques retain the atomic level description, but reduce the number of possible events to the very basic mechanisms of single-defect diffusion and can thus encompass a timeframe (much) larger than MD.

Figure 1 shows an example of AKMC simulation. The cubic box is filled with matrix Fe atoms, and contains a small percentage of Cu. The atoms are arranged in a 3D rigid grid of coordinates that corresponds to the Body Centred Cubic (BCC) crystallographic structure. Several sites are however empty, corresponding to vacancies. At each step of the calculation, one of them is moved as shown on figure 2. Each vacancy has eight possible destinations, corresponding to the eight corners of the BCC cells. One of all candidate jumps, whose number is eight times the number of vacancies present in the system, is selected according to its probability, using the Monte Carlo sampling method[15].

The vacancy jump probability, p_j, is calculated using the classical transition rate theory, i.e. using an Arrhenius-like expression for the jump frequency, that describes the jump as a thermally activated process:

$$p_j = \frac{\nu_{0,j} \exp(-E_j/k_B T)}{\sum_k \nu_{0,k} \exp(-E_k/k_B T)} \qquad (1)$$

Here $\nu_{0,j}$ is a prefactor that is assumed to be constant and of the order of the Debye frequency for Fe (in the present case), k_B is Boltzmann's constant, T is the absolute temperature and E is the Vacancy Migration Energy (VME), which is the key parameter to be assessed and the focus of the present work. A precise definition and description of the latter are given in section 2. For the moment, it is important to know that the VME depends on the Local Atomic Configuration (LAC) and that it can be rigorously calculated with MD-type tools.[†] This is, however, a very time-consuming operation, that cannot possibly be undertaken at every AKMC step. Our project is thus aimed at partially replacing this rigorous calculation by a regression tool, trained to predict the VME on the basis of a limited number of MD calculated examples. Two possibilities have been envisaged and are reported in sections 3 and 4. The objective is to be able to calculate the VME hundreds of billions of times within a reasonable time

---

[†] Note that, technically speaking, it is not full MD that is used for the VME calculation, but algorithms based on the use of an interatomic potential that are easily implemented in an MD code. There exist a number of them (see Ref. 9), but here we shall not enter the detail of these algorithms and will generically speak, for simplicity, of MD calculations. Note also that calculations of the same type, more reliable although much more expensive in terms of CPU-time, can also be performed using *ab initio*, i.e. quantum-mechanics-based, methods.

Artificial Intelligence applied to AKMC

frame, with as few simplifications as possible, because of the crucial role that the VME plays in the time increment of the AKMC, and thus on the prediction of the kinetics of the simulated process.
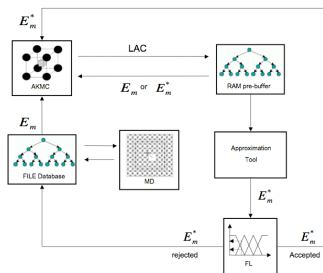

Fig. 3. The evolutionary VME prediction system. E_m is the MD calculated VME, whereas E_m^* is the prediction made by the approximation tool. The pre-buffer and the VME database are implemented under the form of a binary tree.

In order to keep the approximation tool qualities into control, a Fuzzy Logic (FL) based risk assessment method has been developed, too, to determine the probability that a VME prediction is subjected to an unreasonably high error. Such a tool allows the construction of the evolutionary system shown on figure 3. The rejected VME are MD calculated, before adding this new information in the existing database. Once the latter has been filled with a certain amount of new entries, both the approximation tool and its FL risk assessment module can be automatically re-trained. Such a strategy is promising to enhance the VME prediction qualities during the AKMC. The FL module is described in section 5.

Physical considerations suggesting the convenience of this scheme and a few preliminary results, obtained with first rudimentary algorithms, with only a few hints about the architecture of the numerical tools, have been already reported in Ref. 16. Here we focus on the
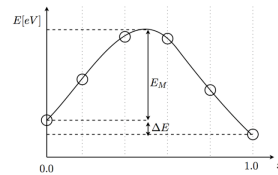
Fig. 4. Vacancy Migration Energy (VME). x is the dimensionless advancement coordinate along the atom displacement path. The y-axis of the figure is the total energy of the system in electron volts. E_M is a cubic spline interpolation of the saddle point in the minimum energy path found by the drag method.

detailed description and discussion of the algorithmic part, presented in its latest form.

2. The Vacancy Migration Energy

The VME is the difference between the largest energy encountered during the process leading to the exchange between a vacancy and a nearby atom (saddle point) and the initial energy of the system, as illustrated on figure 4. The VME can be estimated in a number of ways. Empirical formulas based on the total energy difference (ΔE) (see Ref. 5, 6, 9-11, 17.) are the simplest to apply, but also the most approximate ones, as discussed also in Ref. 16. Rigorous calculations can be undertaken with methods such as "drag", "dimmer" or "Nudged Elastic Band" ones. The interested reader can find a general survey on that topic in Ref. 18. The method we used was a drag refined with cubic splines interpolation. The total energy of the system is calculated with MD performing a quench of the crystal.

The VME varies with the Local Atomic Configuration (LAC), as illustrated on figure 2. The A, B1, ... , F atoms shown in addition to the migrating one are the first nearest neighbors (1nn) of both its initial and final positions. They can be of several chemical types, or even be another vacancy. Depending on their nature, the corresponding VME will be different. The LAC can thus be coded under the form of an array of integers :

$$1\text{nn LAC} = [\ J\ \ A\ \ B1\ \ B2\ \ \ldots\ \ E2\ \ E3\ \ F\ ] \qquad (2)$$

N. Castin et al.

Each entry corresponds to one particular site around the jump and the value it takes depends on the nature of the
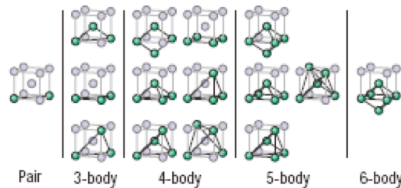

Fig. 5. Types of many-body interactions considered for the 3nn approximation.

object found therein. The length of the LAC string depends on the accuracy of the correlation. For example, the 1nn approximation shown on figure 2 considers 15 atomic sites. Taking further nearest neighbors into account makes this number increase. The list of possible values that the integers may take depends on the physical problem at hand. We talk about the FeAB_Xnn problem when, in addition to the Fe matrix and 1 migrating vacancy, both the A and B species may be encountered and participate in defining the LAC and when the Xnn approximation is used, i.e. the LAC is extended to include the Xth shell of neighboring sites.

Table 1 summarizes the problems of interest for the study of the FeCu binary alloy. It is clear that the number of possible LACs explodes quickly and that a full calculation with MD is totally unfeasible. Note that, in practice, the inherent symmetries of the BCC configuration allow the total number of LACs to be reduced by a factor 6 at the most. This "trick", however, obviously does not remove the underlying complexity mentioned earlier.

Our problem is consequently to correlate the VME with respect to 15 to 39 integer (categorical) type input variables. The output is a smooth real type function that takes values from $0.0^{\ddagger}$ to, say, 1.5 eV, with the largest accuracy possible. The parameters of the drag and

---

‡ Negative values are meaningless as when introduced in (1) they may lead to unreasonably large jump probabilities. Physically, negative values would imply that the transition is not thermally activated, but spontaneous, which means that the initial state is ill-chosen.

Table. 1. Problems of interest for our application. The bold values in the last column are those for which a full calculation with MD and tabulation is feasible. As an order of magnitude, one VME calculation with MD takes about one minute on modern personal computers.

Problem	Num sites	Num LACs
FeCu1nn ; FeVac1nn	15	2^{15} = **32768**
FeCu2nn ; FeVac2nn	21	2^{21} = **2097152**
FeCu3nn ; FeVac3nn	39	2^{39} = 5.50e+11
FeCuVac1nn	15	3^{15} = **1.43e+7**
FeCuVac2nn	21	3^{21} = 1.05e+10
FeCuVac3nn	39	3^{39} = 4.05e+18

quenching methods have been selected as a compromise between speed and accuracy.

3. Cluster Expansion

A cluster expansion for an alloy can be viewed as a generalized Ising Model[19]. Any property of the alloy that only depends on the atomic configuration, the total energy in particular, may be expressed by means of such an expansion. Its application to energy barriers (VMEs) has been proposed e.g. in Ref. 20. More precisely, the occupation variables of the LAC allow a description of the energy barrier as an expansion in terms of polynomials :

$$E_{CE} = J_0 + \sum_{sites} J_i S_i + \sum_{pairs} J_{ij} S_i S_j + \sum_{triplets} J_{ijk} S_i S_j S_k + \dots \quad (3)$$

where S_i are the LAC entries and J_{ij} are the coefficients of the basis functions that can be fitted e.g. to the MD calculated VME's for a variety of different LAC's. Figure 5 shows the types of many-body interactions considered in an expansion, up to the 3nn. The J_{ij} coefficients thus represent the contribution of each of these interactions to the VME.

Of course, only a finite number of interactions can be involved in the expansion. The choice of which

Artificial Intelligence applied to AKMC

interaction is more essential to a particular expansion is an open question.

The fit of the expansion of each LAC to the corresponding MD calculated VME is performed by adjusting the coefficient J_{ij}, minimizing the least-square error between the calculated energy and the predicted energy proposed by the expansion. This minimization can be undertaken with a traditional optimization method like Single Value Decomposition[21] (SVD) based on pseudo-inverse matrix, stochastic methods like Genetic Algorithms[22] (GA), or non-linear parameter fitting like Levenberg-Marquardt[23] (LM).

It was decided, as a first approach, to study the performance of the cluster expansion method for the simple binary 1nn problems. Even for this simple situation it was necessary to impose a truncation on the expansion by considering only many body interaction consisted of pairs or triplets of atoms, in order to avoid the number of terms to explode.

3.1 Binary problems

This section presents our results with the FeCu1nn and FeVac1nn problems (see Table 1), where the S_i occupation variables may thus take only two different values. The number of free parameters to be optimized was $15 + C_{15}^2 + C_{15}^3 = 575$. The percentage of points used for training was 20%. The optimization method applied in this work was GA (SVD and LM gave similar results).

Figure 6 shows the results obtained with the only consideration of pairs. The cluster expansion VME predictions are clearly well behaved for the FeCu1nn problem, even with such a simple model. The additional consideration of triplet interaction, however, allows to reduce the average error committed, as shown on figure 7. On the contrary, the cluster expansion predictions quality is much less satisfying for the FeVac1nn problem.
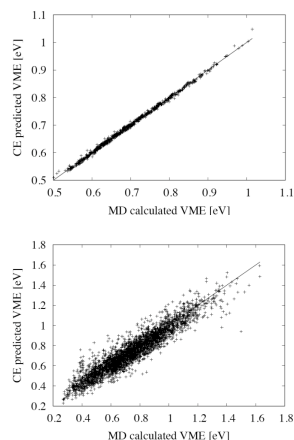
Fig. 6. Cluster expansion VME predictions for the FeCu1nn (up) and FeVac1nn problems (down), with only pair interactions taken into consideration. The average errors are respectively 0.53% and 7.12%.
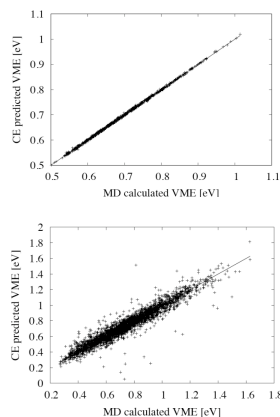
Fig. 7. Cluster expansion VME predictions for the FeCu1nn (up) and FeVac1nn problems (down), with pair and triplet interactions taken into consideration. The average errors are respectively 0.29% and 3.61%.
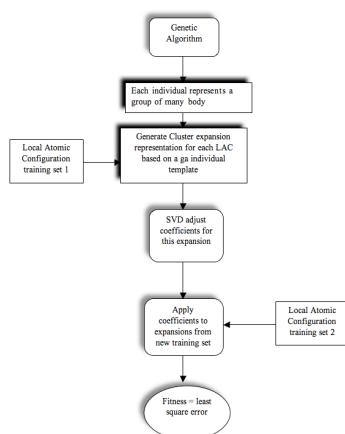
122

N. Castin et al.
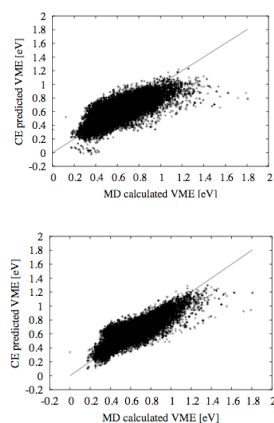

Fig. 8. GA based Cluster Expansion optimization.

Fig. 9. Cluster expansion VME predictions for the FeCuVac1nn with pairs interactions (up) and with triplets interactions (down). The average errors are respectively 15.72% and 11.74%.

3.2. *Ternary problem*

The problem considered here is the more general and more complicated FeCuVac1nn (see Table 1), where both copper atoms and vacancies may be encountered in the LAC.

A solution was imagined to find a way to take into account the possibility for some interactions, in the cluster expansion formulation, to be more important than others[24]. Of course there is no way to identify these interactions on-the-fly during the optimization process, so it was decided to design a GA based intelligent system, depicted on figure 8, in order to identify the relevant many-body interactions of a given problem and to obtain the adjustment of the coefficients for these interactions.

The individuals of the GA population represent different ways to construct a cluster expansion, or different many body interactions to consider. A population of different templates, or possible cluster expansions, is created in the first generation. A training set consisting of local atomic configurations is presented to each individual (training set 1), and these configurations are "translated"

to a cluster expansion according to the scheme of this individual. A SVD algorithm finds the appropriate coefficients, minimizing the least square error on the training set 1. With the adjusted coefficients, a second training set (training set 2) is translated as the cluster expansion proposed by the individual, the least square error is calculated and used as a measure of how good this individual, or this particular set of many-body interactions, is able to produce a good prediction. As the genetic algorithm evolves, only individuals that represent suitable expansions would survive.

The results obtained with the FeCuVac1nn problem are shown on figure 9. The performance of the GA based model is surprisingly good compared to the preliminary results we obtained with the traditional model.

3.3. *Conclusion on CE*

The Cluster expansion approximation tool has been successfully applied to simple VME prediction problems. A GA-based model has been devised to
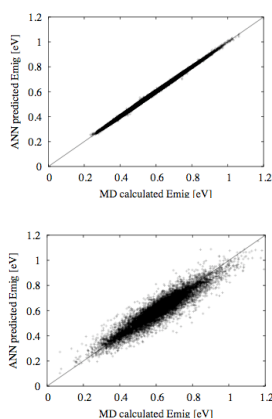
Fig. 10. ANN Prediction qualities for the FeCu3nn (left) and for the FeCuVac3nn (right) (maximum 7 vacancies) problems.

determine the most important interactions to be taken into account, enabling the number of terms in the expansion to be limited.

However, the potential of the cluster expansion seems limited because of the high complexity that the optimization problem takes when the number of atomic sites to be considered is increased. For example, the consideration of the 2nn problems (21 sites in the LAC) requires the determination of around 10,000 free parameters if triplet interactions are introduced in the expansion. An optimization problem of this size is known in GA terminology as Large Parameter Optimization Problem (LPOP) and requires a large population size and many generations to converge. The computational complexity, therefore, quickly explodes. For this reason and considering that, at the same time, we were obtaining better results with a hybrid fuzzy – neural network framework (see following sections), it was decided to abandon this model in favor of a more efficient and robust approach.

4. Artificial Neural Networks

Artificial Intelligence (AI) is the combination of algorithms, data and software used to develop computer systems that can be said to be *intelligent*. Here, the defining feature of intelligence is the capability of learning from past experience and solving problems when important information is missing, so as to be able to handle complex situations and to react correctly to new ones. There are many different computational models that are considered branches of the artificial intelligence field, each one suitable to a different kind of problem. For our particular application, the feed-forwards Artificial Neural Network (ANN) is particularly well suited, as it provides a general framework for representing non-linear functional mappings between a set of input variables and a set of output functions[25]. It is a *universal approximator* in the sense that a Multi-Layer Perceptron (MLP) can approximate any continuous multivariate function to any desired degree of accuracy, provided that a sufficiently large number of hidden neurons are available[26-27].

4.1. *Predictions quality*

Figure 10 shows the ANN prediction qualities for binary and ternary 3nn problems. The average error committed is 0.51% for FeCu3nn and 3.37% for FeCuVac3nn (with maximum 7 vacancies in the LAC). The correlation coefficient r^2 is larger than 0.99 in both cases. The ANN is thus clearly outperforming the cluster expansion, not only because the error committed is much smaller, but also because the training procedure is much less computational time demanding.

The next sections present the experiments we made to study different ANN architectures and training algorithms.

4.2. *Experimental conditions*

Two MLP architectures have been considered. The first is the classical fully interconnected mono-hidden layer network without bypass connections, using sigmoid

N. Castin et al.

activation functions, as widely described in Ref. 25-26 and by many others. It will be denoted as Fixed Architecture MLP (FAMLP) from now on. The second is Fahlman-Lebiere's Cascade Correlation Network (CNN) as described in Ref. 28. It is, contrary to the FAMLP, a constructive algorithm where hidden units are added in successive layers.

Two algorithms have been considered for the FAMLP training. The first one is the steepest-descent Resilient Propagation (RPROP) used in batch mode, as described in Ref. 29. The second one is the Levenberg-Marquardt (LM), described for example in Ref. 25, which is an approximation of the second order Newton method and that does not require the computation of the Hessian matrix. The synaptic connections where initialized at random between $\pm 2.4/F$ (F is the node fan-inn[26] as recommended in Ref. 26).

The CCN training algorithm was changed compared to the original Fahlman-Lebiere's. Instead of proceeding to the addition of a new hidden node in two phases, all synapses linked to it are trained all together with the output ones, with a classical training algorithm (chosen to be LM). The reason is that is seems to us that the original Fahlman-Lebiere training scheme is best suited for classification problems using the 1-of-c coding for the output signal. Furthemore, the QuickProp[30] algorithm originally proposed for CCN training[28] didn't give more satisfactory results than RPROP and LM and is therefore not considered in the present paper. 20 candidate nodes where considered before any new hidden unit addition. The synapses initialization strategy was the same as for FAMLP training, and the activation functions where chosen at random amongst the sigmoid, Gaussian and hyperbolic tangent.

4.3. *Experiment 1*

All architectures and training algorithms have been tested on the FeCu1nn and FeCu2nn problems. The FAMLP was trained with several numbers of hidden nodes. The experiments were run 20 times.

Figure 11 shows the experiments for the FeCu2nn problem. Figure 12 summarizes the results for both the
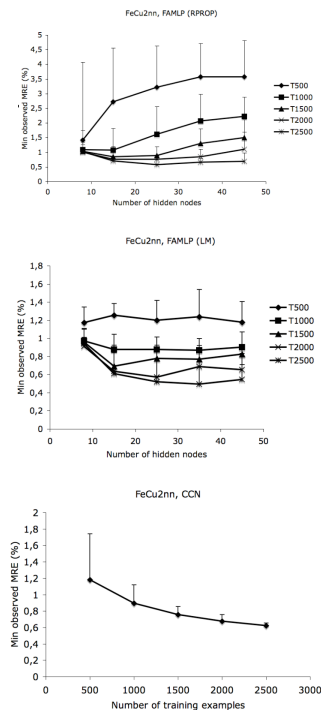

Fig. 11. FeCu2nn trained with FAMLP and CCN. The error bars show the max observed MRE over the 20 trainings performed. T is the number of training examples.

FeCu1nn and the FeCu2nn problems. The following observations can be made :
- The best-suited number of hidden nodes for the FAMLP is not easy to determine, because of the substantial variance of the final MRE. In fact, it is absolutely necessary to run all experiments several times, which is bad news from the computational point of view.

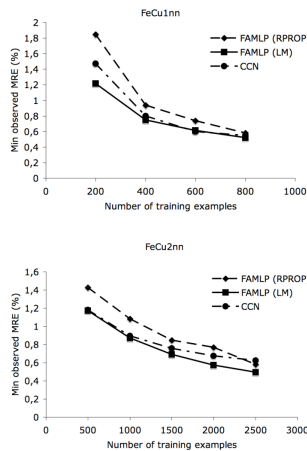*Artificial Intelligence applied to AKMC*

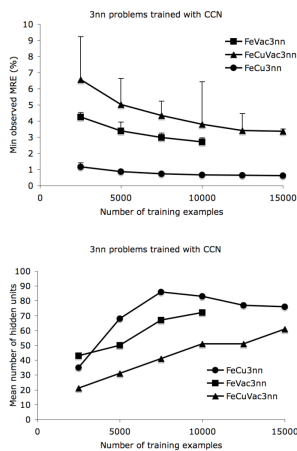Fig. 12. Minimum MRE observed amongst all training experiments performed.

Fig. 13. Experiments with the 3nn problems.

- For FAMLP, the RPROP algorithm does not reach as low MRE as LM, except if the training set is large enough. LM is thus always preferable[§].
- The final MRE variance is unsurprisingly lower for CCN than FAMLP.
- The CCN architecture is in general not capable of reaching the same MRE as the FAMLP, despite its advantage to construct the network automatically. In-depth architectures are thus apparently less appropriate to the VME correlation problem than the use of one single hidden layer.

It is thus clear that reaching a reasonably low MRE is rather easy, whereas fine-tuning is not conceivable without quite a number of training experiments.

4.4. *Experiment 2*

Figure 13 shows the experiments performed with the 3nn problems. Only the CNN architecture was tested. The experiments were run 20 times.

The comparison of figure 12 and 13 shows clearly that the ANN need in training examples to converge to the lowest possible MRE behaves well with the problem complexity. Shifting from 1nn to 2nn or 3nn has not made this number explode. The same observation holds for the comparison of the FeVac and FeCuVac problems to the FeCu one. This is a very important point vis-à-vis the extension of the methodology to more complex problems.

4.5. *Conclusion on ANN*

The ANN is clearly a very promising tool for the VME regression versus LAC. Low mean residual errors of predictions and very good correlation coefficients are indeed very easily obtained. However, ANN fine-tuning is not an easy issue, because quite a lot of training

---

[§] In practice, LM requires much less training epochs to converge than RPROP or other first order methods. However, The actual training time can be reversed for large systems because of the complexity of the LM : ϑ ($W^3 + TW^2$) where W is the number of synapses and T the number of training examples. Anyway, LM was by far the fastest for our application.

N. Castin et al.

experiments are in practice necessary before the best possible ANN performance can be reached.

Consequently, it seems to us that the CCN training scheme is a good starting point to make the first NN training trials on a new problem, in order to determine the number of training points required to reach the lowest possible MRE. Then a long series of FAMLP trainings, with the LM algorithm applied on different network architectures, are to be performed as a second step for fine-tuning.
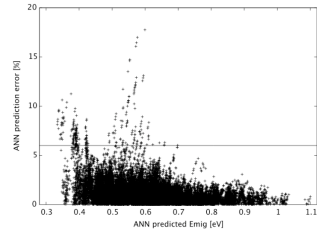
5. Risk assessment on the VME approximation

As already mentioned, a Fuzzy Logic[31] (FL) system has been designed to assess the uncertainty[32] inherent to the use of the ANN and to evaluate the "risk"[33] associated with its use instead of the full calculation, so as to be able to build an integrated system, capable of feedback. Figure 14 shows an example of ANN trained for the FeCu1nn problem and of associated risk assessment scheme. Despite the reasonably low MRE, the error committed for some cases may be rather large, up to 17.5% for this particular example. The objective of the FL would then be to identify, on the basis of both the LAC and the corresponding predicted VME, whether the correlation error is probably larger than a certain Error Rejection Threshold (ERT).

The FL system we developed follows a Sugeno model[34-35] and produces an output that is either 0 or 1, respectively meaning "acceptance" or "rejection" of the ANN predicted VME. The FL inputs are various information extracted from the LAC. For example, the variables for the FeCu1nn problem were :

- The total number N_{Cu} of copper atoms in the LAC.
- The difference ΔC_{Cu} between the number of Cu atoms that are 1nn of the jumping atom and the number of Cu atoms that are 1nn of the jumping vacancy.
- The ANN prediction E_m^* of the migration energy.

Triangular-shaped sets are used for the fuzzification part of the Sugeno model.

Fig. 14. ANN trained for the FeCu1nn problem. The MRE is 1.06%. The ERT is set at 6%.

Fig. 15. FL sets and rules determination by the GA. μ is the relevancy to belong to a certain FL set.

An automated Genetic Algorithm (GA) based learning scheme has been developed in order to generate the FL system definition on the basis of the full VME available database. The GA optimization variables are the central coordinates of the triangular sets on the FL inputs, as shown on figure 15. The rules' conclusions are determined after a passage in the VME database. If at least one point that has an unacceptable error fulfils a certain rule, its conclusion is automatically chosen to be 1 ("rejection"). The conclusion is 0 ("acceptance") otherwise. The GA objective function is :

$$f = (1 + R_1) \cdot (1 + R_2) + 1 \qquad (4)$$

Where R_1 is the proportion of acceptable ANN predictions wrongly rejected by the FL and R_2 is the proportion of unacceptable predictions wrongly

Artificial Intelligence applied to AKMC

accepted. The GA task is thus to select the fuzzy sets as properly as possible, so as to isolate in what conditions precisely is the ANN failing at producing a good VME prediction.

The classical Sugeno inference scheme has been modified in order to improve the FL performance. The FL output O is calculated by :

$$O = round(o) \, ; \, o = \frac{\sum_r H_r \cdot C_r \cdot T_r}{\sum_r H_r \cdot T_r} \qquad (5)$$

Where H_r is the rule relevancy and C_r is its conclusion. The T_r coefficient (named trust index) has been added. It is the maximum relevancy observed by the GA that has led to the conclusion C_r for the rule r. For example, on figure 15, the conclusion C_r of all rules involving the set number 3 on the FL input considered is "rejection". Only one observation leads to that conclusion, but it was lying very close to the set peak. In that case, the "rejection" conclusion for the rule r can be used with a high degree of trust, and so $T_r = \mu_a$ is very close to 1. On the other hand, the conclusion of all rules r^* involving the set number 2 are also "rejection", but with a much lower level of confidence. In this case, $T_{r*} = \mu_C$ is close to 0. The r^* rule has, consequently, fewer chances to induce wrong rejections of acceptable VME predictions.

The complexity of the GA objective function is rather large. Two passages in the overall VME database are indeed required : the first serves to determine the rules' conclusions (as depicted on figure 15) and the second serves to calculate the function f. The application of such a method is very time-consuming if the database is large, which is the case in practice for the 2nn and 3nn problems. In order to reduce that complexity, we have introduced a clustering operation on the FL learning data, depicted on figure 16. Points having a similar value for all FL inputs are removed, as long as they lie on the same side of the ERT. Only one point is left, and is assigned with a weight factor that corresponds to the
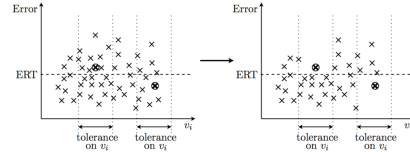

Fig. 16. Clustering operation on the FL learning data, around the two circled points, with respect to 1 FL input. In practice, points are removed only if the illustrated condition is fulfilled for all the FL inputs.
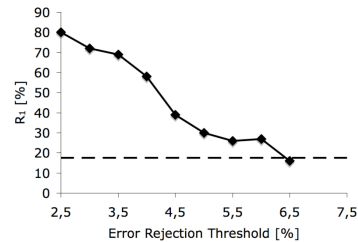

Figure 17 : Evolution with he ERT of the best-obtained R_1 for the problem of Fe-Cu 1nn . R_2 has immediately vanished during the GA optimization in all cases.

number of points removed plus one. This factor is then taken into account when the R_1 and R_2 members of the f function are calculated. In practice, points above the ERT are not clustered (it is affordable since they are not very numerous) and the clustering operation is very fast thanks to the help of a binary tree. The size of the FL learning table is in turn tremendously reduced and the GA optimization is consequently much faster. An appropriate choice of the selection tolerances shown on figure 16 allows the effect of clustering to be made negligible on the real FL abilities to isolate the unacceptable ANN predictions.
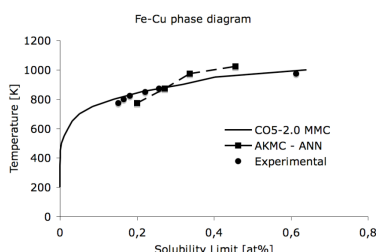
N. Castin et al.

Fe-Cu phase diagram

Fig. 18. AKMC results obtained with the use of the FeCu3nn trained MLP. The CO5 curve shows the Cu solubility limit of the inter-atomic potential used for the energy calculations.

The last undefined feature of the FL system is the appropriate selection of the ERT. In fact, we can see it as a compromise between ANN accuracy and AKMC speed. The ERT is actually not *a priori* chosen, but rather determined as low as possible with the constrain to keep R_1 reasonably limited, as shown on figure 17. The choice of a large ERT is consequently less rigorous in the ANN risk assessment but summons MD calculations less often, and reversely.

6. Example of AKMC results with the use of ANN

Figure 18 shows the results of thermal annealing experiments computed with a small AKMC box containing 1.4% of Cu and 1 single vacancy. The VME was correlated with an MLP having a MRE of 0.5%. No FL Risk assessment was applied. Satisfactory predictions of the Cu solubility limit in Fe were obtained at different temperatures.

7. Conclusion and outlook

In this work, we have reported about our effort to develop a regression tool to partially replace a costly "molecular-dynamics" calculation of the local-atomic-configuration-dependent vacancy migration energy in an atomistic kinetic Monte Carlo scheme, where the local atomic configuration is presented under the form of an array of tens of categorical integers. We have in the first place envisaged a cluster expansion approach.

The latter has however been abandoned for an artificial neural network, that has been proven to be more robust, well behaved and promising for future developments of the project.

Our future objective is to continue the application of this method to more complicated situations. First, the number of atomic sites taken into consideration must be increased for a better description of the physics, and more chemical species are to be included in the model. Secondly, in order to tackle irradiation problems, the atomistic kinetic Monte Carlo model must be able to consider the migration of another type of point-defect, i.e. the self-interstitial. It is a more complicated event than the vacancy migration, mainly because of the more extended and anisotropic strain-field than for a vacancy, which enhances and complicates its interaction with neighboring atoms.

Our future work will thus have to face not only a larger number of artificial-neural-network input variables, but also a more complicated mapping between these inputs and the point defects migration energies that have to be predicted.

Acknowlegments

References

1. G.R. Odette and G.E. Lucas, JOM 53 (7) (2001) 8-22. Mater. Sci. Eng. 6 (1998) 19-28.
2. J.T. Buswell, W.J. Phythian, R.J. McElroy, S. Dumbill, P.H.N. Ray, J. Mace and R.N. Sinclair, J. Nucl. Mater. 225 (1995) 196-214.
3. K. Fukuya, K. Ohno, H. Nakata, S. Dumbill and J.M. Hyde, J. Nucl. Mater. 312 (2003) 163-173.
4. Y. Nagai, Z. Tang, M. Hasegawa, T. Kanai and M. Saneyasu, Phys. Rev. B 63 (2001) 134110.
5. B.D. Wirth and G.R. Odette, "Kinetic lattice Monte Carlo simulations of cascade aging in iron and dilute iron-copper alloys", Mat. Res. Soc. Symp. Proc. 540 (1999) 637-642.

Artificial Intelligence applied to AKMC

6. C. Domain, C.S. Becquart and J.-C. van Duysen, Microstructural Processes in Irradiated Materials, Materials Research Society Symposium Proceedings, Vol. 650, edited by G.E. Lucas, L.L. Snead, M.A. Kirk Jr. et al. (Materials Research Society, Warrendale, PA, 2000), p. R3.25.

7. S. Schmauder and P. Binkele, Comp. Mater. Sci. 25 (2002) 174.

8. Y. LeBouar and F. Soisson, "Kinetic pathway from EAM potentials: influence of the activation barriers" Phys. Rev. B 65 (2002) 094103.

9. E. Vincent, C.S. Becquart and C. Domain, J. Nucl. Mater. 351 (2006) 88.

10. E. Vincent, C.S. Becquart and C. Domain, Nucl. Instr. & Meth. in Phys. Res. B 255 (2007) 78.

11. E. Vincent, C.S. Becquart, C. Pareige, P. Pareige and C. Domain, J. Nucl. Mater. 373 (2008) 387.

12. C. Domain, C.S. Becquart and L. Malerba, J. Nucl. Mater. 335 (2004) 121.

13. C. Ortiz, M.-J. Caturla, C.-C. Fu and F. Willaime, Phys. Rev. B 75, 100102(R) (2007).

14. C. Ortiz and M.-J. Caturla, Phys. Rev. B 75, 184101 (2007).

15. K. Binder (Ed.), "Monte Carlo Methods in Statistical Physics", Springer-Verlag, Berlin (1979).

16. F. Djurabekova, R. Domingos, G. Cerchiara, N. Castin, E. Vincent and L. Malerba, Nucl. Instr. & Meth. In Phys. Res. B, 255 (2007) 8.

17. Young66: W.M. Young and E.W. Elcock, Proc. Phys. Soc. 89 (1966) 735; Kang89: H.C. Kang and W.H. Weinberg, J. Chem. Phys. 90(5), 2824 (1989).

18. G. Henkelman, G. Jóhannesson, and H. Jónsson, "Methods for Finding Saddle Points and Minimum Energy Paths", in Progress on Theoretical Chemistry and Physics, 269-300, Ed. S. D. Schwartz (Kluwer Academic Publishers, 2000).

19. J.M. Sanchez, F. Ducastelle and D. Gratias, Pysica 128A (1984) 334-350.

20. Van der Ven, A; Ceder, G, "Vacancies In Ordered And Disordered Binary Alloy Treated With The Cluster Expansion," Physical Review B, Vol. 71, Issue 5, (2005).

21. Trefethen, L. N., & Bau, D. (1997). Numerical linear algebra. Philadelphia: Society for Industrial and Applied Mathematics

22. D. E. Goldberg, Genetic Algorithms in Search, Optimization, and machine learning, Addisson-Wesley, Reading, Mass., 1989.

23. Gill, P. E. and Murray, W. "Algorithms for the solution of the nonlinear least-squares problem", SIAM J. Numer. Anal. 15 [5] 977-992, 1978.

24. G. L. W. Hart, Volker Blum, Michael J. Walorski, and A. Zunger, "Evolutionary Approach for Determination of First-Principles Hamiltonians," Nature Materials 4, 391-394 (01 May 2005).

25. C.M. Bishop, "Neural Networks for pattern recognition", Clarendon press, Oxford, 1995.

26. S. Haykin, "Neural Networks: A Comprehensive Foundation", New York: MacMillian (1994).

27. K. Hornik, M. Stinchcombe and H. White, "Multilayer feedforward networks are universal approximators", Neural Networks, 2 (1989) 359-366.

28. Scott E. Fahlman and C. Lebiere, "The Cascade-Correlation Learning Architecture", in Advances in Neural Information Processing Systems, volume 2, pages 524-532, 1990.

29. M. Riedmiller and H. Braun, "A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm", Proc. of the IEEE Intl. Conf. on Neural Networks, pages 586-591, 1993.

30. Scott E. Fahlman, "An Empirical Study of Learning Speed in Back-Propagation Networks", Computer Science Technical Report, Carnegie-Mellon University, 1988.

31. L.A. Zadeh, Fuzzy Sets & Systems 100 (1999) 9.

32. D. Dubois and H. Prade, "Unfair Coins and Necessity Measures: Towards a Possibilistic interpretation of Histograms", Fuzzy Sets & Systems 10 (1983) 15.
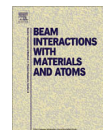
C Paper II

Prediction of point-defect migration energy barriers in alloys using artificial intelligence for atomistic kinetic Monte Carlo applications

N. Castin [a,b], L. Malerba [a,*]

[a] Structural Materials Group, Nuclear Materials Science Institute, Studiecentrum voor Kernenergie Centre d'étude de l'énergie nucléaire (SCK CEN),
Boeretang 200, B-2400 Mol, Belgium
[b] Université Libre de Bruxelles (ULB), Physique des Solides Irradiés et Nanostructures (PSIN), CP234 Boulevard du triomphe, Brussels, Belgium

ARTICLE INFO

ABSTRACT

We significantly improved a previously proposed method to take into account chemical and also relaxation effects on point-defect migration energy barriers, as predicted by an interatomic potential, in a rigid lattice atomistic kinetic Monte Carlo simulation. Examples of energy barriers are rigorously calculated, including chemical and relaxation effects, as functions of the local atomic configuration, using a nudged elastic bands technique. These examples are then used to train an artificial neural network that provides the barriers on-demand during the simulation for each configuration encountered by the migrating defect. Thanks to a newly developed training method, the configuration can include a large number of neighbour shells, thereby properly including also strain effects. Satisfactory results have been obtained when the configuration includes different chemical species only. The problems encountered in the extension of the method to configurations including any number of point-defects are stated and solutions to tackle them are sketched.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

The presence of neutron-irradiation-created point-defects enhances mass transport in alloys and accelerates possible phase changes, which result in material property modifications. The development of predictive tools to model the kinetics of such processes is of great importance, especially for the nuclear sector.

Molecular dynamics (MD) is a reliable tool to study atomic-level processes involving radiation-produced point-defects, their mutual interaction and their interplay with solute atoms, provided that a suitable interatomic potential is available for the material of interest [1]. The timeframe is, however, limited to a few tens of nanoseconds at the most and cannot encompass slow diffusion processes involving vacancies, such as solute precipitation and segregation. Kinetic Monte Carlo (KMC) tools, particularly object [2] or event [3] KMC, where atoms are not explicitly treated and only the kinetics of defects is included, are suitable to extend the timeframe of an atomic-level simulation up to years and to simulate an irradiation process in a realistic way. However, using these methods all the mechanisms involved must be quantitatively known in advance. In addition, the application of these techniques to concentrated alloys is not straightforward. Atomistic kinetic Monte Carlo (AKMC) techniques on rigid lattice are in-between these two extremes [4,5]: they retain the atomic-level description and reduce the number of possible events to the very basic mechanisms of single-defect diffusion, but they can encompass a timeframe (much) larger than MD.

This paper summarises the current state of advancement in the development of a new approach to describe point-defect migration in an alloy, as well as the microchemical evolution it brings about, in an AKMC framework. The approach is based on the idea of exploiting artificial intelligence techniques, namely artificial neural networks (ANN), as advanced regression tools to predict energy migration barriers as functions of the local atomic configuration (LAC), thereby substituting computationally prohibitive on-the-fly calculations. A first attempt, with LACs limited to first and second nearest neighbour shells, had already been reported in [6]. Here we describe how the ANN efficiency and flexibility has been significantly improved, so as to allow the extension to large LAC volumes, thanks to a newly developed ANN training strategy, which is explained in some detail. We show that the achieved regression capability is satisfactory when the LAC is defined by different chemical species only, even for large LAC volumes, thereby including also the effect of strain fields. We also address the more difficult problem of LACs defined by a number of point-defects, sketching possible solutions. The application of the methodology for AKMC simulations of thermal ageing in FeCu and FeCuNi alloys is reported in a more extended companion paper [7].

2. Atomistic kinetic Monte Carlo simulations

The AKMC algorithm models stochastically point-defect migration events as thermally activated processes. The probability for the event i is proportional to the jump frequency $\Gamma_i = \nu_0 \exp(-E_i^m/k_B T)$, where ν_0 is the attempt frequency, E_i^m the migration energy barrier (MEB), k_B Boltzmann's constant and T the absolute temperature. The simulation time is then incremented using a mean residence time algorithm [8]. If the pre-exponential factor ν_0 is assumed to be constant, the only unknown is E_i^m. The latter depends strongly on the LAC, defined by the presence of atoms of different chemical species and also, a priori, of other point-defects, each time distributed in different ways around the migrating point-defect.

Ideally, E_i^m should be properly calculated on-the-fly for each event. In practice, this on-the-fly calculation is generally performed using heuristic approaches, compromise between computational feasibility and physical acceptability [4,5]. Often, the MEB is calculated as a constant barrier, depending only on the chemical nature of the atom exchanged with the point-defect, corrected by the total energy variation due to the migration event, calculated without allowing for relaxation effects, using either pair energies or an interatomic potential [5]. Alternatively, broken-bond methods extended to a few nearest neighbour shells are used for the calculation of both saddle point energy and initial energy, estimating the MEB as difference between the two [4,5]. Either method essentially disregards relaxation effects on the MEB and is limited to small LAC volumes (second nearest neighbour shell). In addition, the former method artificially links a kinetic quantity (the energy barrier) to a thermodynamic quantity (the energy variation).

3. Artificial intelligence approach

Given a LAC, we rigorously calculate the corresponding MEB, using an interatomic potential, with the nudged elastic band (NEB) method [9], thereby including all chemistry and relaxation effects, as much as they can be reproduced by the potential. However, the migration events and the configuration of the AKMC box are still described in a rigid lattice framework, in order to keep the advantageous simplicity of such a model.

Since the systematic on-the-fly use of NEB in the course of an AKMC simulation is unfeasible in practice, because of its inherently huge computing time cost, we rely on an artificial neural network (ANN) [10] as partial replacement of it [11]. Once trained, the latter provides MEB values as functions of the LAC essentially coincident with the NEB calculations, at much lower computational cost. The

LAC is described as a string of integers, i.e. by on-site variables taking different values depending on the chemical species sitting in the corresponding lattice site.

The success of this approach relies completely on the ANN ability to reproduce, with reasonable accuracy and using only on-site variables as input, the result of the NEB calculation tool, even when large LAC volumes are considered. Fig. 1(a) shows one example of satisfactory correlation, including a very large number of atoms in the LAC: although not all ANN predictions are equally accurate, the average error is small and the correlation coefficient is close to 1. The second example (Fig. 1(b)) is less satisfactory from the point of view of the correlation, despite including less lattice sites in the LAC. The reasons for this are discussed in Section 5. Other examples of good correlation are provided in [7].

4. General procedure for the use of ANN in AKMC

The general procedure to be followed to properly train the ANN to calculate the MEB is described in logical order in the next subsections.

4.1. Evaluation of the number of neighbours in the LAC

The number of atomic sites to be included in the LAC for MEB calculation depends on the problem at hand. It must be chosen big enough, so that the MEB does not vary too much if the configuration changes at far enough sites. But it must not be chosen too large, to avoid unnecessary complications and minimise the computational cost. One way to establish the optimal LAC volume is the following: the point-defect is followed in the simulation box in a first AKMC test simulation (for example conducted with a first ANN, preliminarily trained on randomly generated LACs – see Section 4.2) and the encountered LACs (expressed in terms of on-site variables [6]) are listed. For each migration jump, the MEB is then calculated by NEB for increasing LAC volumes. The value corresponding to the largest size is taken as reference to decide starting from which LAC volume the MEB value converges to an essentially constant value: this LAC volume will be the best choice.

4.2. Random databases generation

A reference database of NEB-calculated MEBs, for a first-estimate LAC volume, is generated before the ANN can be trained. In the initial absence of AKMC results, it can be randomly generated. LACs are thus randomly chosen and the MEBs are calculated for all of them. Two types of randomization are possible:
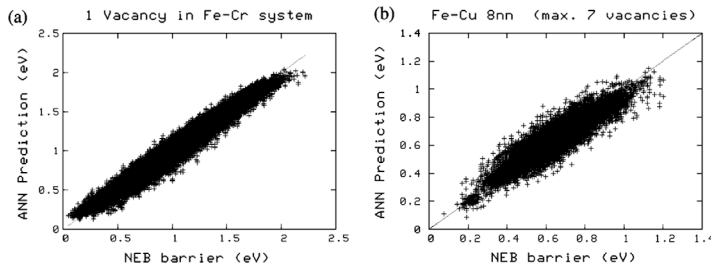

Fig. 1. Quality of ANN predictions of MEBs of a single vacancy (a) in FeCr without and (b) in FeCu with other vacancies (7 max.) in the LAC. The number of LAC atomic sites taken into account are 399 (a), and 143 (b). The mean errors/correlation factor (R^2) are 4.97%/0.99 (a) and 3.5%/0.94 (b).

1. The LAC volume is randomly filled with foreign solute atoms (and/or other point-defects), without any preferential spatial distribution. This corresponds to sampling different configurations in a solid solution.

2. Clusters of foreign solute atoms (and/or point-defects) are randomly created at various distances from the migrating point-defect, and the LAC is generated accordingly. This situation must be included in order to properly describe precipitation (and/or clustering), but it is the result of *a priori* arbitrary choices concerning the size, geometry and composition of the clusters.

The random database must ideally contain examples of both types, in order to be representative of all possible situations that can be encountered during the AKMC simulation where precipitation/clustering is expected. This first training set can be later on complemented by other LACs encountered during further test simulations, if deemed necessary.

4.3. ANN training

ANN training for the regression of MEBs is a computing time consuming operation, especially for large LAC volumes, i.e. with a large number of on-site variables as input [11]. A special ANN training algorithm, shown on Fig. 2, has been developed to alleviate this problem. The algorithm is twice constructive: (i) the input variables are gradually connected to the network, layer of atomic neighbours by layer; (ii) within a level of approximation, nodes are gradually added in the only hidden layer with a constructive algorithm. The fixed architecture training phases use the Levenberg – Marquardt (LM) algorithm, as its superior qualities for this application have been demonstrated [11]. (Further details, including the mathematical aspects, will be published elsewhere.)

The practical interest of this algorithm is manifold: the best network architecture is automatically determined, the algorithm makes good use of multi-processor environments and, finally, most of the training time is spent on the most influencing variables, i.e. the closest neighbouring sites: improving the regression for larger LAC volumes becomes increasingly faster when smaller volumes are already correctly regressed. The use of this LM based training scheme has allowed to take many more atomic neighbours into ac-
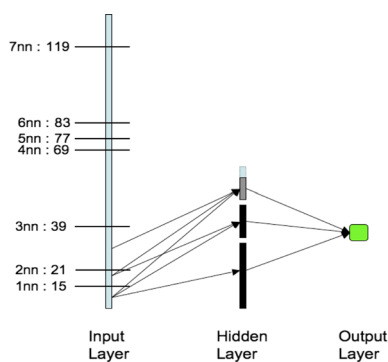
count in the regression compared to our first work, as the ANN trained with classical back-propagation didn't give satisfactory results when more than 21 input variables were taken into account [6].

Note that the complete MEB database is not identically utilized. In practice, less than half of it is used for training, i.e. to adapt the degrees of freedom of the ANN, whereas the remaining part serves as validation set, to measure the actual extrapolation skill on never seen cases.

4.4. Assessment of the ANN predictions in the AKMC

The preliminary ANN is trained on the basis of randomly generated examples. It is thus advisable to verify *a posteriori* that its predictions are close enough to the actual MEB for the cases that are encountered in the course of a simulation, sampled as described in Section 4.1. The error of the ANN on those cases is then measured, while the optimal LAC volume is assessed. The ANN is eventually re-trained if necessary, for a larger LAC volume if needed, including in the training set the newly sampled LACs. This assessment (and re-training, if needed) procedure should be in principle performed after each long AKMC simulation, so that the accuracy of the regression improves with time.

5. Application of the method to the general problem of many point-defects

When the LAC includes not only changing chemical species, but also other point-defects, the problem becomes more involved, because the appearance of significant strain effects must be accounted for by the ANN. This corresponds to higher mathematical complexity and larger number of variables. In this case, in addition, the rigid lattice description of the atomic system implicit in the AKMC scheme is pushed to its limits of validity, since many close-by point-defects, possibly forming a cluster in a fully relaxed configuration, may not be straightforwardly associated with precise rigid lattice sites. Thus, for the moment we restricted ourselves to a limited number of point-defects (e.g. a maximum of seven vacancies, as in Fig. 1). To add complications, it turns out that some AKMC events, although hypothetically possible within the rigid lattice description of the system, lead to unstable final states and should therefore not be considered. Fig. 3 shows an example involving self-interstitials. The problem is that *a priory* no obvious means exist to identify these events on the sole basis of the LAC, without relaxing the crystal to verify the stability. The ANN predictions of the MEBs for these events are thus out of control, since no NEB-calculated value can be provided as example. Moreover, in the presence of point-defects causing important strain
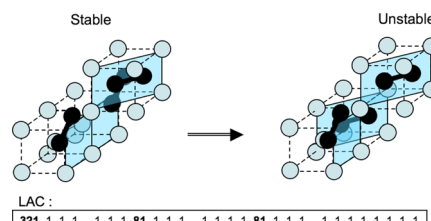
Fig. 2. Constructive ANN training algorithm. The ANN variables that correspond to the second nearest neighbours (2 nn) were already connected in previous steps of the algorithm. The variables of the third level are currently being connected. Subsequently, the effect of further neighbour shells will be included.

Fig. 3. Example of AKMC event that leads to an unstable final state (at least one of the SIA shown on the figure changes if it is relaxed with quenching or conjugate gradients).

effects (self-interstitials), the NEB method itself becomes less reliable: we have observed cases in which the choice of the first guess migration energy path influences the result, thereby determining a situation where more than one MEB value corresponds to the same LAC. Supposing that the latter problem can be solved by devising a procedure to properly choose the initial guess migration energy path, three strategies are currently considered to tackle the problem of many defects in the LAC:

1. A classification ANN can be trained to decide, on the basis of the LAC, whether an AKMC event is to be considered or not.
2. The migration energy ANN can be trained to produce very large barriers for the unacceptable events. It therefore becomes possible to avoid them, by ignoring all events for which the MEB is predicted to be above a certain threshold.
3. It may be possible to associate a reference NEB-calculated MEB with all events if the unstable states are constrained. This method is, however, delicate and risky to apply in practice, because the constraints must be chosen so that they do not significantly alter the MEBs corresponding to non-problematic events.

6. Conclusion

We have presented the advances made in a new approach to predict the migration barriers of point-defects in alloys as functions of the local atomic environment, as described by an interatomic potential, properly allowing for chemical and relaxation effects, of use in atomistic kinetic Monte Carlo simulations on rigid lattice. The main improvements concern the training strategy and the possibility the new strategy offers of taking into account the influence of a large number of neighbouring atoms on the energy barrier, thereby also allowing for strain effects. Satisfactory results have been obtained when the local environment influencing the migration barrier is only determined by the presence of different chemical species. The extension of the method to the general problem of many point-defects requires care, because strain effects become so important that the rigid lattice description is pushed to its limits of validity. Possible solutions have been sketched and will be further investigated in the near future.

References

[1] D. Frenkel, B. Smit, Understanding Molecular Simulation – From Algorithms to Applications, Academic Press, 1996.
[2] C. Domain, C. Becquart, L. Malerba, J. Nucl. Mater. 335 (2004) 121.
[3] C.C. Fu, J. Dalla Torre, F. Willaime, J.L. Bocquet, A. Barbu, Nat. Mater. 4 (2005) 68.
[4] F. Soisson, J. Nucl. Mater. 349 (2006) 235.
[5] E. Vincent, C.S. Becquart, C. Pareige, P. Pareige, C. Domain, J. Nucl. Mater. 373 (2008) 387.
[6] F. Djurabekova, R. Domingos, G. Cerchiara, N. Castin, E. Vincent, L. Malerba, Nucl. Instr. and Meth. Phys. Res. B 255 (2007) 8.
[7] N. Castin, L. Malerba, G. Bonny, M.I. Pascuet, M. Hou, these Proceedings.
[8] W.M. Young, E.W. Elcock, Proc. Phys. Soc. 89 (1966) 735.
[9] H. Jonsson, G. Mills, K.W. Jacobsen, in: B.J. Berne, G. Ciccotti, D.F. Coker (Eds.), Classical and Quantum Dynamics in Condensed Phase Simulations, World Scientific, Singapore, 1998.
[10] C.M. Bishop, Neural Networks for pattern recognition, Clarendon Press, Oxford, 1995.
[11] N. Castin, R.P. Domingos, L. Malerba, Int. J. Comput. Intell. Syst. (IJCIS) 1 (4) (2008) 340.
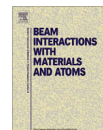
D Paper III

Modelling radiation-induced phase changes in binary FeCu and ternary FeCuNi alloys using an artificial intelligence-based atomistic kinetic Monte Carlo approach

N. Castin [a,b], L. Malerba [a,*], G. Bonny [a,c], M.I. Pascuet [a,d,e], M. Hou [b]

[a] Structural Materials Group, Nuclear Materials Science Institute, Kernenergie Centre d'Etude de l'Energie Nucléaire (SCK CEN), Studiecentrum voor, Boeretang 200, 2400 Mol, Belgium
[b] Physique des Solides Irradiés et des Nanostructures (PSIN), Université Libre de Bruxelles (ULB), Boulevard du Triomphe CP234, 1050 Brussels, Belgium
[c] Laboratory of Theoretical Physics, Universiteit Gent, Proeftuinstraat 86, B-9000 Gent, Belgium
[d] CAC-CNEA, Departamento de Materiales, Avda. Gral. Paz 1499, 1650 San Martín, Pcia. Buenos Aires, Argentina
[e] CONICET, Avda. Rivadavia 1917, 1033 Buenos Aires, Argentina

ABSTRACT

We apply a novel atomistic kinetic Monte Carlo model, which includes local chemistry and relaxation effects when assessing the migration energy barriers of point defects, to the study of the microchemical evolution driven by vacancy diffusion in FeCu and FeCuNi alloys. These alloys are of importance for nuclear applications because Cu precipitation, enhanced by the presence of Ni, is one of the main causes of hardening and embrittlement in reactor pressure vessel steels used in existing nuclear power plants. Local chemistry and relaxation effects are introduced using artificial intelligence techniques, namely a conveniently trained artificial neural network, to calculate the migration energy barriers of vacancies as functions of the local atomic configuration. We prove, through a number of results, that the use of the neural network is fully equivalent to calculating the migration energy barriers on-the-fly, using computationally expensive methods such as nudged elastic bands with an interatomic potential. The use of the neural network makes the computational cost affordable, so that simulations of the same type as those hitherto carried out using heuristic formulas for the assessment of the energy barriers can now be performed, at the same computational cost, using more rigorously calculated barriers. This method opens the way to properly treating more complex problems, such as the case of self-interstitial cluster formation, in an atomistic kinetic Monte Carlo framework.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

The formation of copper-rich precipitates under neutron irradiation is widely accepted to be one of the main causes of hardening and embrittlement in nuclear reactor pressure vessel (RPV) steels during operation [1]. The development of physical models describing the kinetics of this microchemical process is therefore a contribution of primary importance towards the overall goal of producing, in a multiscale simulation framework, tools capable of assessing the service lifetime of nuclear components [2].

In this paper we apply a novel, artificial intelligence based, atomistic kinetic Monte Carlo (AKMC) approach [3–6] to model the microchemical evolution in binary FeCu and ternary FeCuNi alloys driven by vacancy migration. The latter alloy is of particular interest because Cu and Ni are so far considered to be the two elements mainly determining the onset of embrittlement in RPV steels [7]. The novelty of the approach, whose technical details are provided in [6], is that the vacancy migration energy barrier

in the alloy is rigorously calculated allowing for the effects of local chemistry and relaxation, using a reliable interatomic potential and the nudged elastic band (NEB) method [8]. Since the latter is computationally too expensive to be used on-the-fly, it is partially replaced by an artificial neural network (ANN), trained to reproduce, as accurately as possible, the NEB-calculated energy barriers. During the AKMC simulation, at each Monte Carlo step, the ANN provides on-the-fly and on-demand the correct energy barriers seen by the vacancy for the local atomic configuration (LAC) seen by the vacancy at that step. We prove that the use of ANN-generated energy barriers is indeed totally equivalent to the use of NEB-calculated ones. Thus, AKMC studies can now be conducted, at affordable computational cost, with more rigorously calculated energy barriers than hitherto done by means of heuristic formulas [9,10].

2. Methodology

2.1. The AKMC algorithm

In the AKMC algorithm [9,10], point-defect migration events on a rigid atomic lattice, leading to phase changes such as segregation

* Corresponding author. Tel.: +32 14 333090; fax: +32 14 321216.
E-mail address: lmalerba@sckcen.be (L. Malerba).

and precipitation, are described as thermally activated processes. The probability for the event i is proportional to the jump frequency $\Gamma_i = v_0 \exp(-E_i^m/k_B T)$, where v_0 is the attempt frequency, E_i^m the energy barrier, k_B Boltzmann's constant and T the absolute temperature. One out of all possible events is chosen, at each Monte Carlo step, by extracting a random number. The simulation time t^{MC} is then incremented, according to a mean residence time algorithm [11], by a quantity $\Delta t = 1/\Sigma_i \Gamma_i$. If the pre-exponential factor v_0 is assumed to be constant (as is usually done [9,10]: here it is set to $6 \times 10^{12} \text{ s}^{-1}$), the only unknown is E_i^m. The latter depends strongly on the local atomic configuration (LAC), defined by the presence of atoms of different chemical species, each time distributed in different ways, around the migrating point-defect and also, *a priori*, of other point-defects.

2.2. NEB calculation of migration energy barriers

The application of the NEB method [8] to calculate the energy barriers requires first of all an atomic cohesive model (here an interatomic potential is used) and then a good initial guess for the minimum energy path followed by the atom exchanging its position with the vacancy. This is provided for example by the drag [12] method, i.e. by moving the atom towards the vacancy along a straight line. The NEB is then applied to optimize this initial guess for the minimum energy path. In practice, the improvement is generally negligible as compared to the simple drag method, used e.g. in [4]. We opted, however, for the NEB method because of its wider generality and more reliable applicability also for the study of the migration path of other point defects (e.g. self-interstitials [6]) and in the presence of important strain field effects. For example, whenever other point-defects are present and participate in defining the LAC, the correction that the NEB introduces compared to the drag is appreciable.

2.3. ANN training and use

A database of NEB-calculated energy barriers must be generated in order to train the ANN. The ANN 'learns' from the examples that are provided (about 10^5 examples are needed, independently of how many lattice sites are included in the LAC) and 'finds the logic' connecting the input variables (the LAC, expressed in terms of on-site variables, i.e. integers indicating which chemical species or defect occupies a given lattice site within the LAC volume) with the output (the corresponding energy barrier). The underlying criteria applied to build the databases used as training and validation sets for the ANN and to choose how many atomic sites should be included in the LAC (LAC volume) are discussed in [6]. Examples of the application of such criteria are provided in the present work, in Sections 3–5. The details of how the ANN is internally constructed while trained and how its accuracy is tested, are provided in [6].

When the training and validation procedure is completed, the ANN is simply a piece of software that provides the AKMC code, given a LAC, with the corresponding energy barrier, requiring for this operation the time that the computer needs to calculate the value of a mathematical function, which is orders of magnitude shorter than the time required for a NEB calculation.

3. Application to the FeCu system

3.1. Preliminary studies

The interatomic potential used to calculate the energy barriers in the FeCu system is the CO5.20 [13], specifically designed to provide a good description of the thermodynamic properties of FeCu

(phase diagram), as well as a prediction consistent with *ab initio* indications for a few key vacancy migration energy barriers.

In order to define the number of sites to be included in the LAC, and the type of LACs to be used as examples to train the ANN, the specific features of the problem at hand were taken into account. Namely, the FeCu alloys of interest for nuclear applications are fairly dilute. Thus, except nearby precipitates, the probability to encounter Cu atoms near a particular site is on average correspondingly small and, in this situation, the LACs can be randomly generated, without any loss of generality. When Cu clusters or precipitates form, these are essentially pure Cu and take near-spherical shapes already for small sizes. The situation of a vacancy in the matrix nearby a precipitate is therefore the most critical one, but realistic configurations can be easily anticipated. The convergence of the NEB calculation towards a certain value for growing LAC volumes has therefore been evaluated by considering cases of the latter type, with vacancies close to spherical Cu clusters of increasing size, as pictorially shown in Fig. 1. The migrating vacancy was placed at random positions around and also inside it and the corresponding energy barrier was calculated for increasing LAC volumes. The largest size was taken as the reference to define the error committed by taking LACs of smaller sizes. Fig. 1 shows that the LAC should ideally include more than 500 lattice sites, to ensure a good precision if clusters of 300 atoms (0.94 nm), or more, are expected to be formed during the AKMC simulation. It is reasonable, however, to limit the LAC to about 150 neighbours, if the expected precipitate size does not exceed 150 atoms (0.75 nm), i.e. to study the early stages of the precipitation process. Two hundred and twenty five atoms are expected to provide sufficient accuracy also for larger precipitate sizes.

For the present application, the ANN has been trained taking 143 atomic sites into account, which corresponds, in a body-centred cubic (bcc) lattice, to including all atoms up to the 8th nearest neighbours (8nn) of both the initial and the final position of the migrating vacancy. Fig. 2 shows the final quality of the predictions of the trained ANN, as compared to NEB-calculated values, for cases that were never seen before by the ANN. It can be appreciated that the precision of the ANN is very high: the ANN error is much smaller than the error due to the choice of the LAC volume (Fig. 1). Thus, we can consider the use of the ANN as totally equivalent to the use of the NEB method itself.
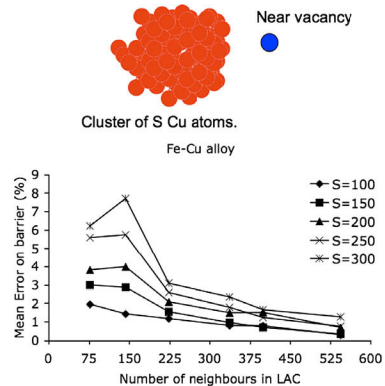

Fig. 1. Evolution of the NEB barrier accuracy with the size of the LAC, for increasing Cu precipitate size (S), expressed in terms of number of Cu atoms.
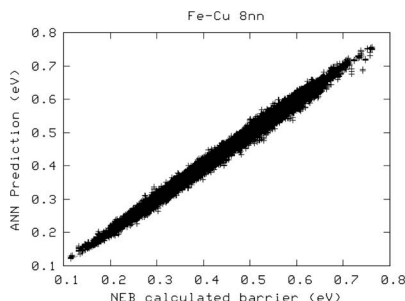
Fig. 2. Quality of the ANN prediction of vacancy energy barriers in the FeCu alloy: the average error is 1.10% and the correlation, R^2, is 0.999.

3.2. FeCu phase diagram according to the model

The CO5.20 interatomic potential has the outstanding feature of reproducing the correct solubility limit of Cu in Fe as a function of temperature, thereby providing a phase diagram comparable with the experimental one [13], at least in the region of low Cu concentration, below the temperature at which Fe undergoes the transformation to the face-centred-cubic (fcc) γ phase. The phase diagram according to the potential has been evaluated using different methods, as shown in Fig. 3.

By using the ATAT code [14,15] for the evaluation of the phase diagram, the solubility limit predicted by the potential in the case of equilibrium between bcc Fe and fcc Cu is in good agreement with the experimental points. ATAT allows also the hypothetical case of equilibrium between bcc-Fe and bcc-Cu to be evaluated. This case cannot be compared to any experimental phase diagram, because bcc Cu is not a stable phase. However, this equilibrium is the one of relevance during the early stages of Cu precipitation, when precipitates are known, also experimentally, to keep a bcc structure [16]. The ATAT calculation reveals that, according to the potential, the solubility limit in the case of bcc-Fe/bcc-Cu equilibrium is higher than in the former case. Metropolis Monte Carlo (MMC) [17] studies with the potential, in which the final configuration is characterised by a certain density of bcc Cu precipitates, confirm this result. Since the AKMC simulations are performed on a rigid bcc lattice (i.e. only coherent precipitation is possible), the bcc-Fe/bcc-Cu solubility limit must be used as reference to
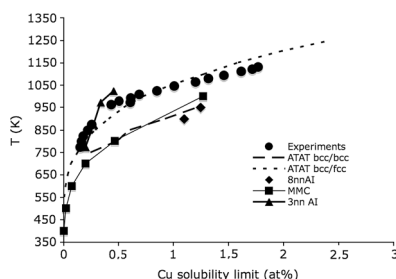

Fig. 3. Fe–Cu phase diagram according to the interatomic potential, obtained using different models and for different equilibria, compared with experimental points from [26]. See text for details.

evaluate the performance of the ANN based AKMC model. It can be seen in Fig. 3 that, indeed, the two points hitherto obtained as a result of ANN-driven AKMC simulations fall close to the MMC points, thereby proving that no significant thermodynamic information is lost when using ANN-generated barriers, instead of those calculated by NEB. To this regard, it should be noted that the thermodynamic information is "hidden" in the barriers in terms of difference between direct and inverse process. A priori, large errors in the prediction of the barrier for the two processes may lead to largely different thermodynamic properties predicted embodied by the kinetic model. Fig. 3 shows that this is not that case and that using ANN-generated barriers is in fact equivalent to using barriers generated directly from the interatomic potential, employing the NEB method for their calculation. It is instructive to note that, if the ANN is trained to LACs extended up to the 3nn shell only, then the solubility limit predicted by the model changes dramatically. This is most likely a chemical effect, because the 3nn approximation ignores the contribution of the 4nn and 5nn that are still in the range of the interatomic potential cut-off. It is however necessary to take even more neighbours than the 5nn into account, in order to provide a proper thermodynamic description, if relaxation effects are to be correctly included.

3.3. Diffusion of small Cu-vacancy clusters

In order to further establish the equivalence between ANN-driven AKMC simulations and the use of NEB-calculated energy barriers, the diffusion coefficient of small Cu-vacancy clusters has been studied. Similar calculations had already been performed in [18], where the method employed to assess the diffusion coefficient is described. The advantage of these simulations is that the possible LACs are limited in number. It is therefore possible to perform the AKMC simulation by taking the energy barriers both from a trained ANN and from tables of NEB-calculated values. Fig. 4 shows the result of the calculation of the diffusion coefficients of two small Cu-vacancy clusters, namely CuV_2 and CuV_3 (V stands here for 'vacancy') using both methods. It can be seen that the ANN error has only very limited influence on the final result.

3.4. Simulation of thermal ageing experiments

A thermal ageing experiment in an Fe-1.34 at% Cu alloy has been simulated at a temperature of 773 K, in a box of $64 \times 64 \times 64$ bcc cells (524, 288 atoms). 5×10^{10} Monte Carlo steps have been performed in about 1.5 CPU-months, using modern mono-processor computers. The simulation has been conducted using energy barriers calculated in three different ways: (a) using the ANN trained for LACs extended up to 8nn, thereby predicting the correct Cu solubility limit in Fe according to the potential and providing an accuracy in the energy barriers acceptable to treat precipitates of up to 200 Cu atoms; (b) using the ANN trained for LACs extended up to 3nn only, known to change the solubility limit in a non negligible way; (c) using for comparison purposes the heuristic formula $E_i^m = E_0 + \Delta E / 2$ [10], where ΔE is the total energy difference between after and before the vacancy jump, calculated on the rigid lattice with the potential. It should be noted that the CPU time required by the three methods is of the same order.

Fig. 5 shows snapshots of the AKMC boxes. Qualitatively, both simulations using the ANN predict a slower coarsening of the initially nucleated clusters than the heuristic formula, after the same number of Monte Carlo steps. In turn, the cluster density is larger with the 3nn ANN, and the clusters are bigger, when compared to the 8nn ANN.

The biggest Cu cluster observed in the 8nn ANN AKMC box counts 204 atoms. Coming back to Fig. 1, we see that the NEB
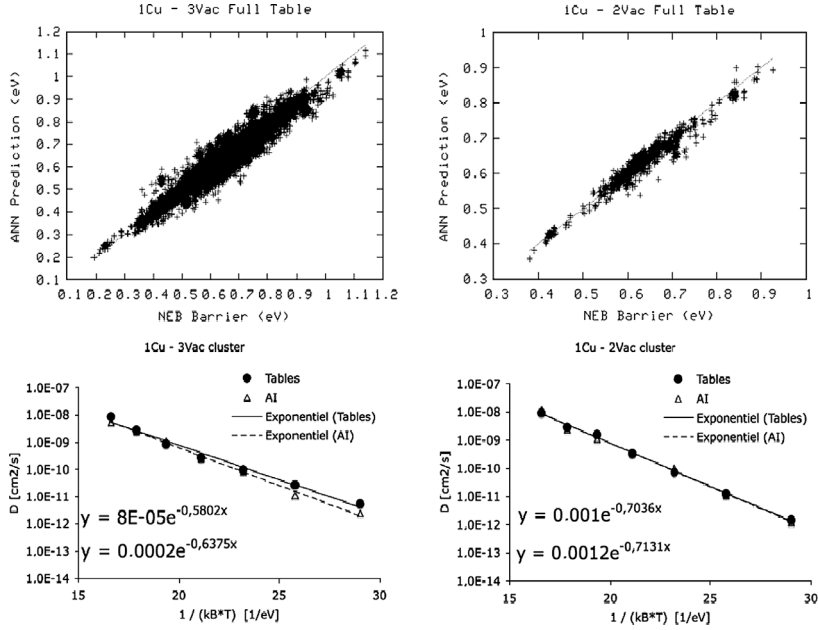
Fig. 4. Determination of the diffusion coefficient versus temperature of two small Cu-vacancy clusters using energy barriers calculated with the ANN or with NEB directly. Top figures: ANN prediction versus the full NEB-calculated table of possible cases (mean errors: 1.77% for CuV$_2$, and 2.81% for CuV$_3$). Figures below: comparison of the diffusion coefficient data points obtained with both methods.

accuracy starts to decrease quickly if this cluster grows further and no more than 150 neighbors are taken into account in the LAC. The ANN should therefore be in principle re-trained to take more neighbours into account, before the simulation is continued, in order to maintain the same accuracy level.

The AKMC time t^{MC} must be rescaled before any comparison with experimental measurements is possible, because the vacancy concentration in the simulation, C_V^{MC}, is in general much larger than the (estimated) real one, C_V^{real}. We have applied the rescaling method proposed in [9]:

$$t^{real} = t^{MC} \cdot C_V^{MC} / C_V^{real}. \tag{1}$$

The real concentration of vacancies is calculated on the basis of the vacancy formation enthalpy, H_V^f (1.71 eV for the potential used):

$$C_V^{real} = \exp\left(\frac{-H_V^f}{k_B T}\right). \tag{2}$$

The vacancy concentration in the box is $1/N$, N being the total number of atomic sites in the box. It is however proposed in [9] to correct it with the fraction of time (Monte Carlo steps) that the vacancy spends surrounded by Fe atoms only (up to 2nn), f:

$$C_V^{MC} = 1/N \cdot f/x_{Fe}, \tag{3}$$

where x_{Fe} is the fraction of Fe atoms. In our simulations, on average, $f \approx 10^{-4}$ when the energy barriers are provided by the ANN, while $f \approx 10^{-2}$ when the heuristic formula is used.

Fig. 6 shows the comparison with experimental data from [19–22] of our AKMC results, with the three methods, after time rescaling according to Eqs. (1)–(3). It is clear that, despite the long CPU time and large number of Monte Carlo steps, the simulation did not go far enough to allow a full comparison with the experimental data points. These correspond already to the coarsening stage and more or less start where the simulation, mainly reproducing nucleation and growth stages, stops. It is difficult to decide whether the decrease of the precipitate density and the increase of the precipitate radius in the simulation are going to show similar trends as the experimental curves. The simulation should be continued for much longer times to draw definitive conclusions. In addition, a bigger simulation box should ideally be considered in order to follow completely the experimental curves. For example, the average precipitate radii reach eventually, in the experiment, a size of 8 nm. Aside from the fact that, when such a size is reached, the precipitates are most certainly incoherent with the matrix [16], and therefore irreproducible with a rigid lattice simulation method, with a Cu concentration of 1.34 at% it also would be necessary to run the simulation in a 190 × 190 × 190 cubic cell box (about 14 million atoms) just to provide enough Cu atoms to create only one 8 nm-size precipitate! This emphasises the need to develop parallel AKMC paradigms and also highlights the inherent limitations of the method. It is important to note, at any rate, that the kinetics predicted with the use of the heuristic formula is, as expected, significantly different from the prediction made with ANN-generated energy barriers. Conversely, the overall difference between using an 8nn or a 3nn trained ANN is negligible, thereby suggesting that,
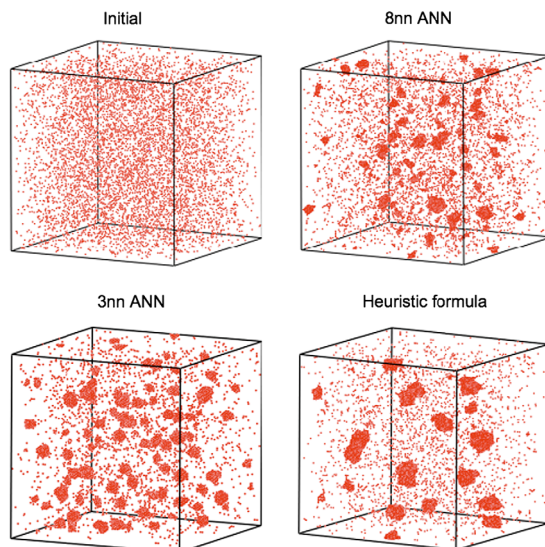
5

Fig. 5. Initial and final snapshots in the AKMC thermal ageing simulation performed using two different ANN (trained to 8nn and 3nn LACs) and using the heuristic formula linking the energy barrier with the energy difference between after and before the vacancy jump.
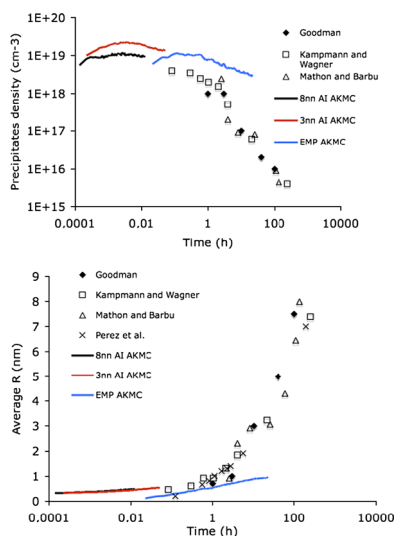

Fig. 6. Comparison of the thermal ageing AKMC simulation results with experimental data, similarly to the comparison performed in [9].

probably, at the end of the day the loss of accuracy associated with the smaller LAC volumes is less important than other factors.

4. Application to the FeCuNi system

The same methodology as for FeCu has been applied to develop a model for FeCuNi alloys. In this case, a ternary interatomic potential [23] has been used, constructed by coupling the CO5.20 for FeCu to a new FeNi potential, equivalently performant from the thermodynamic point of view and capable of reproducing also the CuNi experimental phase diagram [24].

Fig. 7 shows the quality of the ANN predictions with 8nn LACs. The mean error committed is larger than in the FeCu case, most likely because the correlation problem is now more complicated, as a third atomic species must be taken into account.

Preliminary thermal ageing simulations have been conducted at a temperature of 823 K, in a $40 \times 40 \times 40$ bcc cell box containing 1.13% Cu and the 1.36% Ni. For this conditions reference experimental data exist [25]. However, the simulation results here reported concern only the very early stage of the precipitation process, so a direct comparison, such as in Fig. 6 for FeCu, is meaningless. Fig. 8 shows the initial and final snapshots of the AKMC box. A comparison with simulations conducted in FeCu at the same conditions revealed that the cluster density is larger, and the average cluster size smaller, in presence of Ni, thereby suggesting that Ni somewhat enhances Cu cluster nucleation, in qualitative agreement with experimental observations [25]. In addition, an analysis of the precipitate profiles clearly indicated that Ni tends to gather at their periphery, again in qualitative agreements with experiments conducted using atom probe techniques.
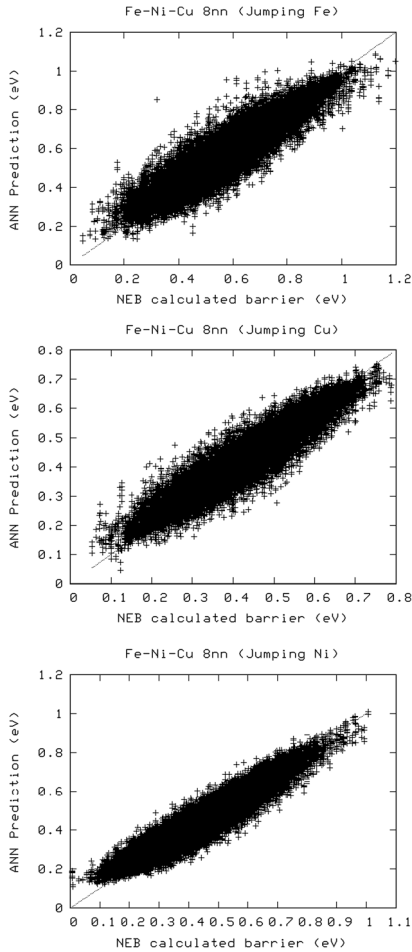
Fig. 7. ANN quality of energy barrier prediction for the FeCuNi system. The mean errors are 4.38% (jumping Fe atom); 4.94% (jumping Cu atom) and 8.17% (jumping Ni atom). The R^2 are all about 0.95.

5. Discussion

The examples provided in the previous sections show clearly that, thanks to artificial intelligence techniques such as ANNs, it is possible to introduce, at least partially, the effect of relaxation and local chemistry in AKMC rigid lattice frameworks, at affordable computational cost. As a matter of fact, the use of a conveniently trained and constructed ANN proves largely equivalent to performing on-the-fly, at each Monte Carlo step, multiple NEB calculations of energy barriers using an interatomic potential and taking into account the effect of the local environment up to distances of
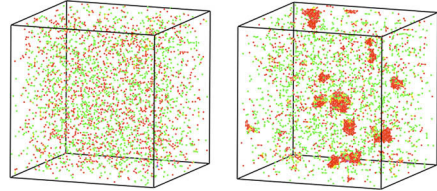

Fig. 8. AKMC initial and final snapshots in a thermal ageing simulation of an FeNiCu alloy (Cu atoms in red, Ni atoms in green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

8nn and beyond. Performing on-the-fly such kind of calculations would be computationally prohibitive.

We have shown that a priori the energy barriers are so sensitive to the environment that convergence to a constant value is achieved only if the effect of about 500 atoms is taken into account, the optimal choice being largely dependent on the problem at hand. The overall effect is a combination of chemical and strain field contributions that are difficult to distinguish in a clearcut way. This casts a priori serious doubts on the acceptability of models where the effect of local chemistry is reduced to two or three neighbour shells only [9,10]. At the same time, we have seen that, when applied to simulating situations comparable with experiments (e.g. thermal ageing in FeCu alloys), the difference between including the effect of many or only a few neighbour shells is not overwhelmingly important, thereby suggesting that, in general, we are authorised to limit ourselves to small relaxation volumes, without significant loss of information. This is somehow good news, as it certainly simplifies the task, independently of which technique is used to estimate the energy barriers in AKMC simulations.

Nonetheless, limiting the LAC to too small volumes may indeed be misleading. One example has been already provided in Section 3.2, in the case of the ANN trained to 3nn LACs, which gave only apparently correct results in terms of predicted phase diagram. Another example is cited here. In the first training of the ANN for the FeCuNi alloys, limited to the 3nn level of approximation, a randomly generated energy barrier database was initially used. Fig. 9(a) shows the corresponding ANN performance, which appears to be fully satisfactory. However, in AKMC simulations using this ANN, Cu precipitation did not occur at all, at any temperature, for Cu contents well above the solubility limit. An appropriate analysis revealed that the ANN predictions were actually good on average for the LACs encountered during the simulation, except some cases, especially when the vacancy was surrounded by a large or small number of Cu and Ni atoms, e.g. in the vicinity of a precipitate (Fig. 9(b)). This suggested the systematic application of a more careful procedure for the choice of the LACs to be used as examples for the training, as discussed in [6], as well as the need to consider larger LAC volumes.

Preliminary attempts at extending the application of the method to LACs including several point-defects have been made. As discussed in [6], this represents a more difficult problem. As a first exercise, we tried to predict with an ANN the energy changes after migration jumps performed by self-interstitials surrounded by other self-interstitials. We limited ourselves to the migration event of a ⟨110⟩ dumbbell that jumps to one of its 1nn positions and becomes oriented ⟨011⟩, including up to five other dumbbells in the LAC. Fig. 10 shows the quality of the ANN prediction of the relaxed energy difference, in the presence of 3–5 dumbbells in the LAC. Clearly, the ANN did, globally, understand the logic between the LAC and the corresponding energy difference. The predictions are, however, sometimes very far from the desired value. Although
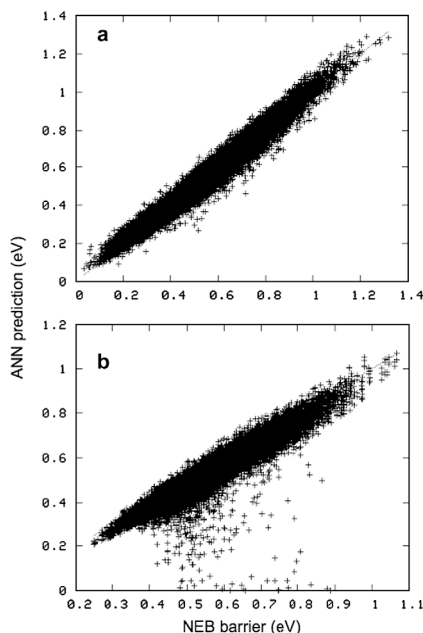
7

Fig. 9. 3nn ANN ill predictions on AKMC encountered cases if the training set is overspecialized to homogeneous distributions of solute atoms.
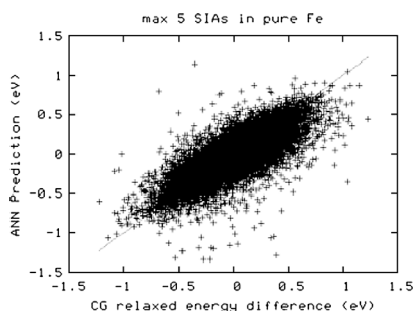

Fig. 10. ANN quality of prediction of the relaxed energy difference before and after SIA migration in the case of five self-interstitials in the LAC ($R^2 = 0.80$).

the result is encouraging, it points to the need of more elaborate strategies in order to reach performances comparable to those obtained for LACs defined by chemical species only.

6. Conclusion

We have shown that artificial intelligence techniques, such as neural networks, can be valuable tools to introduce relaxation and local chemistry effects in atomistic kinetic Monte Carlo models, even remaining formally in a rigid lattice approximation. They can indeed reliably provide the same result as on-the-fly calculations of energy barriers, using the nudged elastic band method with an interatomic potential, at a much lower computational cost. The availability of good interatomic potentials for FeCu and FeCuNi alloys allowed the application of the method with results that are consistent with experimental data concerning thermodynamic properties and thermal ageing processes. The extension of the method to simulate irradiation processes requires, however, the effect of many point-defects to be accounted for. Preliminary efforts in this direction provided encouraging results.

References

[1] G.R. Odette, G.E. Lucas, JOM 53 (7) (2001) 8.
[2] J.-P. Massoud, S. Bugat, J.-L. Boutard, D. Lidbury, S. Van Dyck, F. Sevini, PERFECT: prediction of irradiation damage effects on reactor components, in: Proceedings of the EU Conference FISA 2006, Luxemburg, p. 123.
[3] R. Domingos, G. Cerchiara, F. Djurabekova, L. Malerba, Applied artificial intelligence, in: D. Ruan, P. D'Hondt, P.F. Fantoni, M. De Cock, M. Nachtegael, E.E. Kerre (Eds.), Proceedings of the 7th FLINS Conference, 2006, p. 883.
[4] F. Djurabekova, R. Domingos, G. Cerchiara, N. Castin, E. Vincent, L. Malerba, Nucl. Instr. and Meth. Phys. Res. B 255 (2007) 8.
[5] N. Castin, R. Pinheiro-Domingos, L. Malerba, International Journal of Computational Intelligence Systems 1 (4) (2008) 340.
[6] N. Castin, L. Malerba, these Proceedings. doi:10.1016/j.nimb.2009.06.041.
[7] ASTM Standard E900-02, Predicting Radiation-Induced Transition Temperature Shift in Reactor Vessel Materials, ASTM International, West Conshohocken, PA, <http://www.astm.org>.
[8] H. Jonsson, G. Mills, K.W. Jacobsen, in: B.J. Berne, G. Ciccotti, D.F. Coker (Eds.), Classical and Quantum Dynamics in Condensed Phase Simulations, World Scientific, Singapore, 1998.
[9] F. Soisson, C.C. Fu, Phys. Rev. B 76 (2007) 214102.
[10] E. Vincent, C.S. Becquart, C. Pareige, P. Pareige, C. Domain, J. Nucl. Mater. 373 (2008) 387.
[11] W.M. Young, E.W. Elcock, Proc. Phys. Soc. 89 (1966) 735.
[12] G. Henkelman, G. Jóhannesson, H. Jónsson, in: S.D. Schwartz (Ed.), Progress on Theoretical Chemistry and Physics, Kluwer Academic Publishers, 2000. p. 269.
[13] R.C. Pasianot, L. Malerba, J. Nucl. Mater. 360 (2007) 118.
[14] A. van de Walle, G. Ceder, J. Phase Equilibr. 23 (2002) 348.
[15] A. van de Walle, M. Asta, Model. Simul. Mater. Sci. Eng. 10 (2002) 521.
[16] P.J. Othen, M.L. Jenkins, G.D.W. Smith, Philos. Mag. A 70 (1994) 1.
[17] K. Binder, in: K. Binder (Ed.), Monte Carlo Methods in Statistical Physics, Springer Verlag, Berlin, 1979 (Chapter 1).
[18] F. Djurabekova, L. Malerba, C. Domain, C.S. Becquart, Nucl. Instr. and Meth. Phys. Res. B 255 (2007) 47.
[19] S.R. Goodman, S.S. Brenner, J.R. Low, Metall. Trans. 4 (1973) 2363;
S.R. Goodman, S.S. Brenner, J.R. Low, Metall. Trans. 4 (1973) 2371.
[20] R. Kampmann, R. Wagner, in: C. Janot, W. Petry, D. Richter, T. Springer (Eds.), Atomic Transport and Defects in Metals by Neutron Scatterings, Springer, New York, 1986. p. 73.
[21] M.H. Mathon, A. Barbu, F. Dunstetter, F. Maury, N. Lorenzelli, C.H. de Novion, J. Nucl. Mater. 245 (1997) 224.
[22] M. Perez, F. Perrard, V. Massardier, et al., Philos. Mag. 85 (2005) 2197.
[23] G. Bonny, R.C. Pasianot, N. Castin, L. Malerba Phil. Mag., 2009, accepted for publication.
[24] G. Bonny, R.C. Pasianot, L. Malerba, Modelling Simul. Mater. Sci. Eng. 17 (2009) 025010.
[25] J.T. Buswell, C.A. English, M.G. Hetherington, W.J. Phythian, G.D. Smith, G.M. Worrall, in: N.H. Packan, R.E. Stoller, A.S. Kumar (Eds.), Effects of Radiation on Materials: 14th International Symposium, ASTM STP 1046, American Society for Testing and Materials, West Conshohocken, PA, 1990, p. 127.
[26] G. Salje, M. Feller-Kniepmeier, J. Appl. Phys. 45 (1977) 1833.

E Paper IV

Calculation of proper energy barriers for atomistic kinetic Monte Carlo simulations on rigid lattice with chemical and strain field long-range effects using artificial neural networks

N. Castin[1,2] and L. Malerba[2,a)]

[1]*Université Libre de Bruxelles (ULB), Physique des Solides Irradiés et des Nanostructures (PSIN), boulevard du Triomphe CP234, Brussels 1050, Belgium*
[2]*Studie Centrum voor Kernenergie–Centre d'Etudes de l'énergie Nucléaire (SCK•CEN), NMS unit, Boeretang 200, Mol B2400, Belgium*

In this paper we take a few steps further in the development of an approach based on the use of an artificial neural network (ANN) to introduce long-range chemical effects and zero temperature relaxation (elastic strain) effects in a rigid lattice atomistic kinetic Monte Carlo (AKMC) model. The ANN is trained to predict the vacancy migration energies as calculated given an interatomic potential with the nudged elastic band method, as functions of the local atomic environment. The kinetics of a single-vacancy migration is thus predicted as accurately as possible, within the limits of the given interatomic potential. The detailed procedure to apply this method is described and analyzed in detail. A novel ANN training algorithm is proposed to deal with the necessarily large number of input variables to be taken into account in the mathematical regression of the migration energies. The application of the ANN-based AKMC method to the simulation of a thermal annealing experiment in Fe–20%Cr alloy is reported. The results obtained are found to be in better agreement with experiments, as compared to already published simulations, where no atomic relaxation was taken into account and chemical effects were only heuristically allowed for. © *2010 American Institute of Physics*. [doi:10.1063/1.3298990]

I. INTRODUCTION

Atomistic kinetic Monte Carlo (AKMC) methods are widespread tools to study diffusion-controlled microstructural and microchemical evolution in alloys during thermal aging and under irradiation (see, e.g., Refs. [1–14]). In these models, the atoms of the alloy are located on the positions corresponding to the crystallographic structure of interest, generally on a rigid lattice. The evolution of the system is driven by the migration of point defects, generally vacancies[1–14] and recently also self-interstitials,[8,9,13] whose position is exchanged with nearest neighbor atoms (migration jump).[9,14] The jump to occur is each time selected stochastically, based on the Monte Carlo method.

Simulations conducted with these methods can potentially predict in detail, and with high accuracy, phenomena such as precipitation or segregation,[1–14] which are known to affect the mechanical and chemical behaviors of materials, thereby determining their aging when in use. There are two main limitations to the application of AKMC tools: the first one is their still high computing cost;[5,12] the second one is the physical reliability of the method itself. AKMC models are indeed of straightforward implementation and use, but it is not easy to make them computationally faster; on the other hand most of the physics is contained in the energy barriers associated with the migration jumps of the point defects, E_m, which must embody both the thermodynamics and the kinetics of the system being studied. The accuracy and precision with which these migration energies are determined represent therefore a key issue. AKMC simulations will be the more reliable, the better the influence of the local atomic environment (LAE) on E_m is included in the model in terms of chemistry and strain field. Generally, however, improving the reliability of E_m calculations entails a significant increase in the computational cost of the algorithm. In this paper we describe a method that improves E_m calculation reliability at affordable comptuational cost.

Kang and Weinberg[15] proposed that the energy barrier can be decomposed as $E_m = \varepsilon + \Delta E/2$, where ΔE is the energy change associated with the change in thermodynamic state between after and before the jump, and ε is the *excess* energy to be added to obtain the complete barrier, whose value is *a priori* unknown and will also be, in principle, a function of the LAE. Based on these premises, they also proposed a recipe to assess the energy barrier, given a harmonic potential for the atoms at equilibrium.[15] In practice, such a decomposition has been later most often used assuming ε =constant (e.g., final-initial system energy method in Ref. 10, see also Refs. 6, 9, and 11–14). Somewhat more sophisticated methods, based on broken-bond considerations[1,3,13,14] extended also to the saddle point,[2] have been used as well. In recent times, attempts at increasing the reliability of these heuristic methods have been made, by fitting the pair energy parameters of the model directly to density functional theory (DFT) calculations.[5,6,9,10,13,14] These methods do take explicitly into account the influence of the chemical environment

---

a)Author to whom correspondence should be addressed. Electronic mail: lmalerba@sckcen.be.

132, 074507-1

E. Paper IV

on the energy barrier, but they do so in an oversimplified manner and totally disregard the effect of faraway atoms.

At the other end of the spectrum of possible AKMC models, Henkelman and Jonsson[16] envisaged a way to eliminate not only the rigid lattice approximation, but also the predefined sequence of migration events, given for example an interatomic potential as Hamiltonian. In their scheme, the choice of the migration events, together with the calculation of their corresponding migration barriers, is made on the fly by applying the dimer method,[17] which allows, given the initial state, all possible transition paths to other nearby local minima in the potential energy surface to be found. The advantage of such a method is clearly its flexibility with respect to systems where the rigid lattice approximation would no longer be valid, such as at free surfaces and grain boundaries, or in the bulk of materials containing dislocations or nanostructural features such as nanovoids and dislocation loops. The main drawback is the complexity and high computational cost of such a method, which requires systematic use of the dimer search for possible system transitions.

Recently, methods using sophisticated mathematical techniques to calculate energy barriers, such as cluster expansion,[18] or genetic programming,[19] have been proposed. These approaches keep a rigid lattice description of the system, but calculate the energy barriers between two given states on a nonrigid lattice, then fitting reliable expressions to them with appropriate regression methods. However, so far these approaches have been used based on a limited number of examples (that can practically be calculated with DFT methods) and either little importance has been given to verifying the capability of the obtained mathematical expression to predict never seen cases (local environments),[18] or the total amount of possible cases was any way relatively small.[19]

Along similar ideas, we have proposed in Refs. 20 and 21 a method that constructs a mathematical regression of the energy barrier vs the LAE. Energy barriers are calculated with the nudged elastic band method (NEB) (Refs. 22 and 23), using an interatomic potential as Hamiltonian. The mathematical regression is implemented with an artificial neural network (ANN).[24] All the necessary atomic neighbors are taken into account, as outlined in Sec. II. Our method is not limited to a small amount of possible cases and is inherently fully validated on never seen cases. Once calibrated, it allows fast on-demand and on-the-fly calculation of the migration energies, without the need for a systematic use of the NEB method (that would be unfeasible in practice). In this paper, we take a few steps further in the development of this ANN-based approach compared to Ref. 21, where the fundamental ideas were only shortly explained. Section II supports the importance of using an approach such as ours, by showing that E_m is indeed influenced by the chemical nature of faraway atoms, via not only chemical interactions, but also the strain field that they create. The ANN development and training technique, especially designed for this application, are presented in full detail in Sec. III. The ANN extrapolation skills are then carefully evaluated in Sec. IV. Finally, a few case studies are presented in Sec. V to show that the AKMC simulations performed with our methodology manage to re-
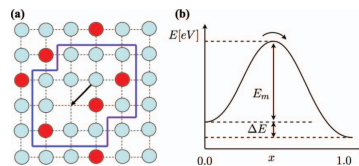

FIG. 1. (a) 2D representation of a vacancy migration event in an AKMC simulation. Different colors denote different chemical species. The LAE considered for the E_m calculation is enclosed by the thick line. (b) E_m is the difference between the saddle point along the minimum energy path followed by the migrating atom and the initial energy of the system.

produce available experimental results satisfactorily, in the especially delicate case of concentrated alloys. We consider here only the case of single vacancies migrating in a chemically changing environment, but the methodology can, at least in principle, be generalized to more complex cases.

II. INFLUENCE OF LONG-RANGE INTERACTIONS ON ENERGY BARRIERS

In the specific AKMC model that we use here, the evolution of the system of atoms is solely driven by the thermally activated migration jumps that a single vacancy takes. The probability associated with the vacancy migration event i is calculated on the basis of the jump frequency, Γ_i, expressed according to the standard transition state theory as

$$\Gamma_i = \nu_0 \exp(-E_{m,i}/k_B T). \qquad (1)$$

Here, ν_0 is an attempt frequency (considered as a constant in first approximation: $\nu_0 = 6 \times 10^{12}$ s^{-1}), k_B is Boltzmann's constant, T is the absolute temperature, and $E_{m,i}$ is the point-defect migration energy corresponding to jump i. Figure 1(a) illustrates schematically the vacancy migration event in a two-dimensional (2D) representation of the rigid lattice of an AKMC model. The actual jump that the vacancy takes at a given point in the simulation time t is chosen stochastically according to the MC algorithm; subsequently, the clock is updated by adding a Δt which is calculated applying the residence time algorithm.[25–27]

For the NEB calculation of E_m, the migrating pair (atom and vacancy with which it exchanges position) is placed at the center of a separate box of atoms, together with its LAE, defined by the xth closest neighbors in the AKMC simulation box. The rest of the box is filled with matrix atoms—Fe atoms in the present study— and periodic boundary conditions are applied. Given the atomic coordinates, the total energy and the atomic forces are calculated with a central force many-body interatomic potential, e.g., Olsson's Fe–Cr potential[28] or Pasianot's Fe–Cu potential.[29] The rigid lattice atomic coordinates corresponding to the initial and final states, before and after the vacancy jump, are first relaxed with the conjugate gradient method[30] to find the nearest local minima in the potential energy surfaces. The problem of calculating E_m is then to find the minimum energy path that the atom exchanging position with the vacancy follows between the relaxed initial and final states, allowing also for relaxation of all the atoms of the box during the migration event
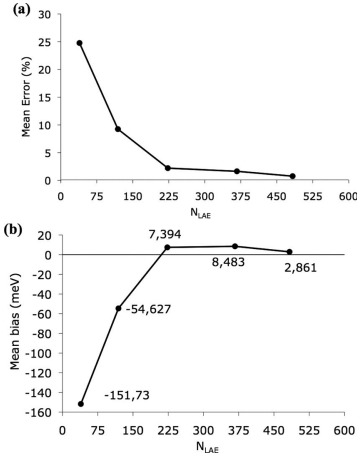
FIG. 2. (a) Average error [Eq. (2)] committed by accepting the NEB calculated E_m value in Fe–Cr alloys corresponding to a given number of atoms in the LAE (N_{LAE}), as compared to a reference value obtained for a very large amount of atoms (615 atoms) in the LAE; (b) Average bias [Eq. (3)] as a function of the same variable. The averages are taken on 2500 randomly selected LAEs.
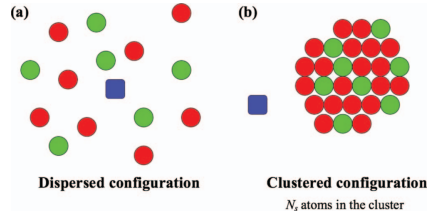

FIG. 3. Generation of LAEs around vacancy migration events. The square represents the migrating vacancy. The circles represent the solute atoms and different colors denote different chemical species. (a) Dispersed distribution of solute atoms. (b) Presence of solute atom clusters.

itself. A first approximation of this minimum energy path can be calculated with the classical drag method,[31] i.e., by simply "pushing" the atom along a straight line between initial and final positions. Then, the NEB algorithm[22,23] is applied to optimize this first guess. The energy barrier E_m is finally obtained by identifying the saddle point along the minimum energy path, as depicted on Fig. 1(b).

Central force many-body interatomic potentials for metallic systems are built with a cutoff (about 5.3 Å for the ones used in this work). Such a cutoff, in body-centered-cubic iron-based alloys, encloses 77 atoms, the neighbors to the migrating pair. These neighbors must thus imperatively be taken into account in the E_m calculation, as they interact chemically directly with the migrating atom before or after its jump. As shown in Ref. 32 as well, however, in reality E_m is sensitive also to the chemical nature of atoms located far beyond the potential cutoff, due to strain field effects. Figure 2(a) shows how the average error committed on E_m, averaging on different LAEs in the Fe–Cr system, scales with the number of atoms defining the LAE, N_{LAE}, while remaining smaller than a reference $\max\{N_{LAE}\}$. This average error is defined as

$$\bar{e}_x = \left\langle \frac{|E_m^x - E_m^*|}{E_m^*} \right\rangle. \tag{2}$$

Here, E_m^x are the migration energies calculated when x neighbor atoms are included in the LAE, and E_m^* are the reference barriers calculated taking $\max\{N_{LAE}\}$ neighbors into account. The interatomic potential used for NEB calculations was

Olsson's.[28] Figure 2(b) shows how the mean residual bias, defined as

$$\bar{b}_x = \langle E_m^x - E_m^* \rangle, \tag{3}$$

scales with N_{LAE}. We see that at least 200 neighbor atoms should ideally be included in the LAE, in order to acceptably converge to the reference value (error $\sim 1\%$), obtained for 615 neighbors.

The Fe–Cr system corresponds to a case of large solubility limit for Cr in Fe,[33] so it makes sense to consider high local concentrations of Cr, even at thermodynamic equilibrium. In the Fe–Cu system, on the other hand, the solubility limit of Cu in Fe is very low and high local concentrations of Cu correspond in practice to the presence of large Cu clusters in a Cu depleted matrix. The difference between the two cases (large number of solute atoms dispersed in the matrix and formation of large clusters) is pictorially illustrated in Figs. 3(a) and 3(b). Figures 4(a) and 4(b) show, respectively, the average errors [Eq. (2)] and the average biases [Eq. (3)] for the E_m calculated for vacancies migrating in the Fe–Cu system nearby a precipitate, as compared to the reference E_m^* value, not only for different numbers of neighbor atoms in the LAE (N_{LAE}) but also for different sizes of the nearby precipitate (N_x). The potential used for the NEB calculations is Pasianot's.[29] It can be seen that, for large Cu clusters, the error committed by not including a sufficiently large number of atoms in the LAE for the NEB calculation is even larger than in the dispersed Fe–Cr system and grows with the size of the precipitate.

We have therefore shown that an accurate calculation of E_m requires that the effect of an extended environment should be taken into account. This cannot be done using heuristic approaches based on broken bonds or on pair interaction energies, not even when the relevant parameters are fitted to highly reliable DFT results as in Refs. 5, 6, 9, 10, 13, and 14.

For this reason, we want to develop a methodology in which a set of NEB calculated values, obtained with the highest accuracy reasonably possible, i.e., for large numbers of atoms in the LAE, is used to produce an equivalent mathematical expression fitted to them. This expression should allow a prediction of E_m values corresponding to LAEs not included in the set of examples, with reasonably small error,
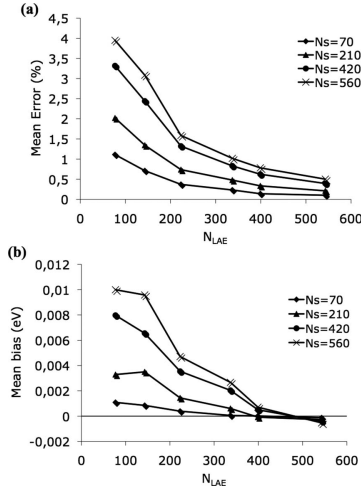
E. Paper IV

FIG. 4. (a) Average error [Eq. (2)] committed by accepting the NEB calculated E_m value in Fe–Cu alloys corresponding to a given number of atoms in the LAE (N_{LAE}) as compared to a reference value obtained for a very large amount of atoms (615 atoms) in the LAE; (b) Average bias [Eq. (3)] as a function of the same variable. The vacancy migrates around or inside isolated spherical clusters composed of N_S Cu atoms. The averages are taken over 1000 randomly selected cases.

so that it can be used as a reliable (and fast) replacement of on-the-fly NEB calculations. This is achieved using the ANN, as is described in the next section.

III. REGRESSION OF THE VACANCY MIGRATION ENERGIES IN ALLOYS WITH ANNs

As stated, in this work we limit ourselves to single-vacancy jumps in chemically changing local environments. The LAE can therefore be always entirely defined by atoms occupying rigid lattice positions and can be coded by assigning an on-site variable to each lattice position included in it. The value of the on-site variable changes depending on the chemical/defect species sitting there, e.g., 0 for vacancies, 1 for Fe atoms, 2 for Cu or Cr atoms, etc. Thus, the independent variables of the function to be fitted are represented by strings of integer numbers, each of them corresponding to the chemical species sitting in the associated lattice position included in the LAE. This definition of the variables is conceptually coincident with the choice that has been made when applying the cluster expansion method for regression.[18]

This section describes the problem of training an ANN to predict, given as input the mentioned strings, the value of E_m that would be obtained by a full NEB calculation. As explained in the previous section, the contribution of many atomic neighbors to the migrating pair (on the order of 200 or more) should ideally be taken into account because of long-range chemical and strain field effects.

To develop the ANN we take the following steps. First, a large table of examples of LAEs is generated at random, and the corresponding E_m are calculated by NEB using the relevant interatomic potential. This table can be as large as computing time allows and is not limited by the nature of the problem (the total number of logically possible LAE, even though countable, can be considered as infinite in practical terms). Next, the table of examples is split up into two non-overlapping sets, called training and validation sets, respectively. At this point, the ANN is trained to minimize the mean square error between its predictions and the actual NEB energy barriers, using only the examples of the training set for this purpose. The validation set is finally used to measure the network extrapolation skills on never-seen cases, and therefore to stop training when no further progress is made. For a successful outcome, the key issue is the choice of the architecture of the ANN and the training strategy, discussed in Sec. III B. In what follows, we spend a few words to explain how the sets of examples are generated.

A. Generation of training and validation sets

For the generation of the databases of examples of E_m, a procedure for the choice of random LAEs must be defined. Such a procedure must ensure that a varied enough spectrum of possible local configurations is explored. Two strategies have been considered, aimed at being representative of any situation that may be encountered:

- Strategy A: Solute atoms are randomly distributed around the migrating pair, with a homogeneous spatial distribution, as depicted on Fig. 3(a). These cases correspond, for example, to the beginning of a thermal annealing experiment, when the alloy is still a solid solution, while at equilibrium they are representative of alloy compositions below the solubility limit. The local concentration of solute atoms should not, however, be constrained to respect the solubility limit value, because the local concentration around the migrating pair in the course of the AKMC simulation changes and can in fact take any value in principle. So, different local concentrations should be explored.

- Strategy B: Clusters of solute atoms are created and the migrating pair is randomly located around them, as illustrated in Fig. 3(b). According to the nature of the studied alloy, if the outcome of the AKMC simulation can be qualitatively predicted, the size and the composition of the clusters, as well as the concentration of solute atoms in the matrix, can be optimized. These parameters can, if needed, be revised later, after some AKMC simulations have been performed.

The database of examples should ideally be created with cases randomly generated following both A and B strategies. The training and validation sets are finally defined by arbitrarily splitting up the table into two subsets. The validation set should be at least as large as the training set, as it is devoted to measuring the ANN error for cases that were not seen during training. The order of magnitudes of these is tens of thousands of examples.

B. Neural network training algorithm

The general issue of training an ANN to solve a given problem exhibits different facets of complexity. Not only the best network architecture is *a priori* unknown, but also, given the architecture, the optimal numerical values for its N_w degrees of freedom (*synaptic weights* or *synapses* in ANN jargon) are to be determined.

In Ref. 20, we compared two types of ANN: the classical single-hidden-layer network, with sigmoid activation functions, and Fahlman–Lebiere's cascade correlation network[34] where, starting from an empty network, nodes are progressively added in successive hidden layers during the training. We found that the quality of the prediction of the former type was generally better than the latter.

If we thus limit ourselves to single-hidden-layer networks, the parameters to be optimized are the number N_H of hidden nodes and the number of examples in the training set, S. In this case, the simplest strategy to develop an optimal ANN is to train different networks independently, with varying N_H, selecting eventually the best one on the basis of the final mean residual error, measured on the validation set. In Ref. 20 we recognized the Levenberg–Marquardt training scheme[35,36] as suitable, in a single-hidden-layer architecture, for a relatively simple 39-variable problem. However, the training time required by this method scales as $\vartheta(N_W^3 + S \cdot N_W^2)$. This approach is therefore unsuitable for problems involving many variables, such as those that we intend to tackle here, especially if different trainings must be performed independently, to be able to choose the best one at the end.

In this context, constructive algorithms[37,38] offer a promising solution. The idea is to construct the ANN progressively, by dividing up the overall training problem into successive subtrainings, on reduced numbers of network synapses. For this purpose, as we focus on single-hidden-layer networks, we modified and generalized the *dynamic node creation* by Ash.[39] Starting from an empty hidden layer, nodes are added progressively, in an iterative way. The degrees of freedom (connections or synapses) associated with the new nodes, together with the output node, are initially trained while keeping frozen the rest of the already existing network (*greedy manner*). This is schematically illustrated in Fig. 5(a). Several independent trainings are performed, changing the initial synapses values, limited to the newly added group of nodes, using the Levenberg–Marquardt algorithm. Then, the whole network is retrained one single time, using the best synapses that were found in the previous step as starting point [Fig. 5(b)]. The number N_n of nodes added at a time should not be too large, to avoid the insertion of unnecessary hidden nodes. At the same time, it should not be too small, to speed up the algorithm. ($N_n = 1$ in the original Ash's algorithm.[39]) Trials have highlighted that this dynamic node creation produces networks that are as accurate as if they had been globally trained with the Levenberg–Marquardt algorithm on fixed architectures, but the CPU time is reduced by more than one order of magnitude.

This training algorithm can be further improved. In the scheme just described, all input variables are simultaneously
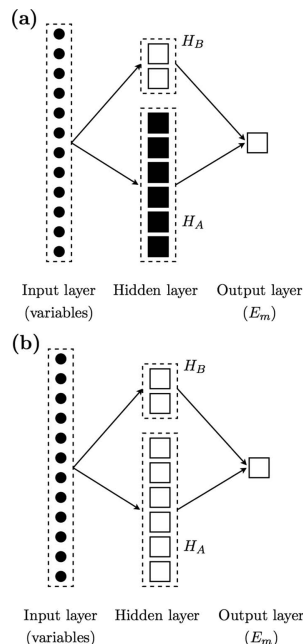

FIG. 5. Node additions to the hidden layer in the dynamic node creation algorithm. The arrows symbolize a whole set of connections (synapses) between groups of nodes. In (a) the black nodes in group H_A are frozen. N_n new hidden nodes (H_B group on the figure) are added and trained together with the output node, using the Levenberg–Marquardt algorithm. The training is repeated several times, with different initial synapses, but only for this group. In (b) the whole network is retrained one single time, using its present synapses as starting point for the Levenberg–Marquardt algorithm.

connected to the newly added nodes. Thus, this algorithm inherently assumes that all the atoms included in the LAE of the migrating pair have the same weight to determine the migration energy. This is clearly untrue because the closest neighbor shells for sure will have a stronger influence on the migration energy than the farthest ones. The training algorithm can thus be improved by taking this information into account. The way to do so is illustrated in Fig. 6. Here, only the input variables corresponding to the first nearest neighbor shell of atoms (denoted as the I_1 group) are first connected and the network is trained without using the other ones. Next, the group of variables I_2 (corresponding to the second nearest neighbor shell) is connected to the network. And so on. We called this training scheme *gradually improving accuracy constructive algorithm* (GIACA). In the GIACA, the connection of each group of variables (say, I_2) proceeds in three phases:

- Phase 1: The new variables are progressively connected to the existing hidden layer, N_n nodes at a time, as in the

E. Paper IV

I_2

I_1

Input layer   Hidden layer   Output layer
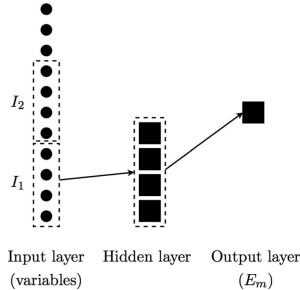(variables)   (E_m)

FIG. 6. Initial network state before the computation of a GIACA step. The variables in the I_1 group have already been connected. The variables in the I_2 group are candidates to be connected to the network.

dynamic node creation. Only the synapses associated with the currently connected group of variables, in the newly connected hidden nodes, are trained, together with the output ones [Fig. 7(a)], in a greedy manner (i.e., keeping the rest of the network frozen). Several independent trainings are required, with random initialization of the new synapses. Then, all the nodes that receive a signal influenced by the I_1 and I_2 variables are fully retrained one single time, starting from the present network state [Fig. 7(b)]. This procedure is iteratively

(a)

I_2

I_1

(b)

I_2

I_1

(c)

I_2

I_1

(d)

I_2

I_1

Input   Hidden   Output
Layer   Layer   Layer
(variables)   (E_m)

Input   Hidden   Output
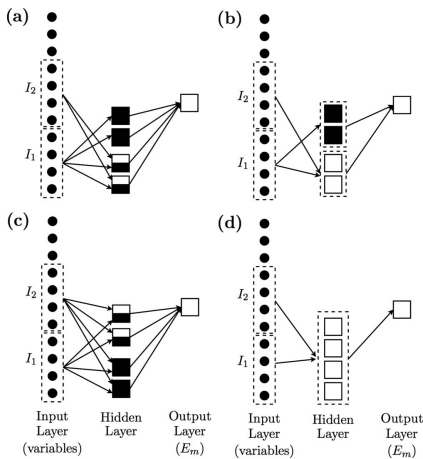Layer   Layer   Layer
(variables)   (E_m)

FIG. 7. First phase for the connection of new input variables in the GIACA. (a) The new inputs (I_2 group) are connected to the N_n first hidden nodes that are not already connected to them. Only the newly created synapses are then trained, together with the output ones, using Levenberg–Marquardt, keeping the rest of the network frozen. (b) All the hidden nodes connected with the inputs of the I_1 and I_2 groups are trained, together with the output node, keeping the rest of the network frozen. (c) The two first steps are iteratively repeated until all I_2 variables are connected to the hidden layer, or until no progress is made. (d) The whole network is retrained one single time.

repeated until the whole hidden layer has been connected with the I_2 inputs [Figs. 7(c) and 7(d)], or is prematurely stopped if no progress is made.

- Phase 2: New hidden nodes are added using the dynamic node creation algorithm, considering together both I_1 and I_2 input variables sets. The existing hidden layer of the network is then totally frozen, even during the dynamic node creation full training phases, in order to limit the number of synapses, as much as possible. The dynamic node creation is again stopped when no further progress is possible.

- Phase 3: The complete network is retrained one single time, using the present synapses as starting point.

Phases 2 and 3 are skipped if not all hidden nodes are connected during phase 1. Training is stopped when the connection of a new group of variables provides no further improvement in the mean error during the first phase.

The main interest of the GIACA, compared to standard dynamic node creation, is that the final complexity of the ANN is limited, while ensuring the highest accuracy of predictions reasonably achievable. This happens because (a) the less influencing input variables, whose connection to the network is not practically desirable with a view to reducing complexity, are spontaneously determined; (b) in the first phase, the connection between the input variables and the hidden layer may remain incomplete, again avoiding the introduction of unnecessary synapses in the network.

The only remaining problem is the determination of the best-suited size of the training set. The latter is a trade-off between having enough learning material for the ANN and avoiding unreasonably long training. However, since there is no *a priori* criterion to choose the training set size, several trials must be made, with increasing training set size, until convergence.

IV. APPLICATION TO BINARY AND TERNARY ALLOYS

The GIACA has been applied to predict with an ANN the E_m of single vacancies in 615 atom LAEs containing both Fe and either Cu or Cr atoms (binary alloys), or simultaneously Cu and Ni atoms (ternary alloys). The training and validation sets were built using the NEB method with Pasianot's potential for Fe–Cu,[29] Olsson's two-band model potential (fitted to projector augmented wave PAW) data for Fe–Cr,[28] and Bonny's potential for the Fe–Cu–Ni ternary alloy.[40,41] (All these potentials have the quality of providing a good reproduction of the thermodynamic properties—phase diagram—of the concerned alloys, as well as a good description of the interaction between point defects and solute atoms in ferritic alloys.) The quality of the ANN prediction is shown in Fig. 8.

The average errors $\bar{e}$ and the average biases $\bar{b}$, obtained by comparing the ANN predictions, o_i, with the corresponding NEB values from the validation set, d_i, are reported on Fig. 8. They were calculated as

$$\bar{e} = \frac{100}{N} \sum_{i=1}^{N} \frac{|d_i - o_i|}{d_i}, \qquad (4)$$
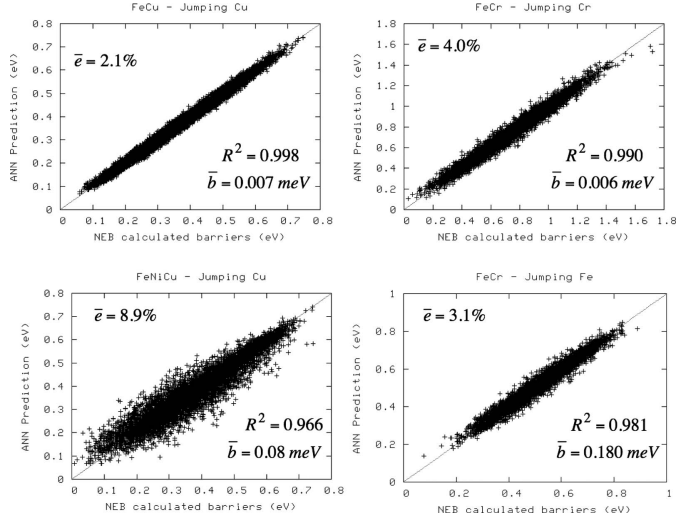
FIG. 8. Quality of the ANN predictions of the vacancy migration energies obtained with the NEB method in three different alloys. The number of examples in the training and reference sets was 30 000 for the Fe–Cu and Fe–Cr alloys; 60 000 for the Fe–Ni–Cu alloy. The average error, $\bar{e}$, and average bias, $\bar{b}$, are calculated using Eqs. (4) and (5), respectively. R^2 is Pearson's product-moment correlation coefficient.

$$\bar{b} = \frac{1}{N}\sum_{N=1}^{N}(d_i - o_i), \qquad (5)$$

where N is the number of examples in the validation set. The ANN predictions for the binary Fe–Cu and Fe–Cr alloys are particularly satisfactory (small error, small bias, and high correlation), so the ANN can clearly be considered as a very good and much faster substitute to the NEB method. The quality of the predictions somewhat decreases with the introduction of a third chemical species in the alloy. However, the ANN error remains unbiased and the correlation coefficient remains close to 1, although the error committed on certain migration barriers is very large (up to 150%).

Figure 9 shows the evolution during training of the mean error committed on the validation set, versus the number of input variables connected to the hidden layer. The number of variables that must be connected in order to have convergence is similar in all cases (about 100), but the final average error depends on the problem. The number of synapses created by the GIACA is similar for both binary alloys, while it is significantly lower in the case of the ternary alloy. As an order of magnitude, training was about 150 CPU h long for all cases, on modern monoprocessor machines.

The final test for the ANN predictive capability is obtained by checking how it handles situations significantly different from those contained in the training set. The following computer experiment has been therefore performed, in the case of the Fe–Cu system. The ANN has been trained either only on configurations with solute atoms dispersed in
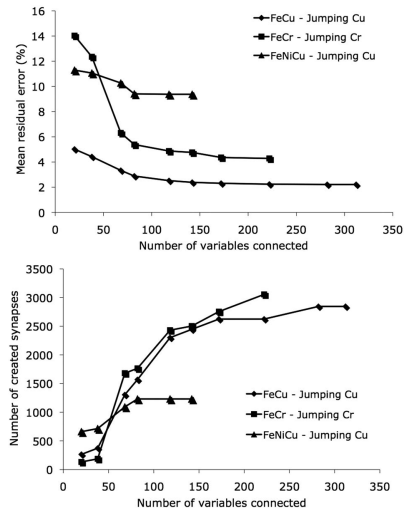

FIG. 9. Evolution of (above) the mean residual error [Eq. (4)] and (below) the number of synapses created in the network during the GIACA training. The number of training examples was 30 000 for the Fe–Cu and Fe–Cr problems; 60 000 for the Fe–Ni–Cu problem.
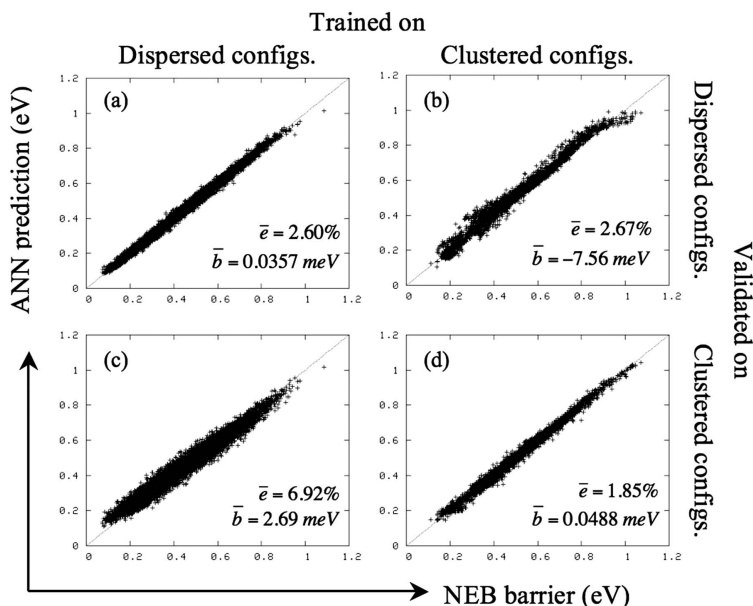
FIG. 10. Quality of the ANN prediction of vacancy migration energies when validated on homogeneous, (a) and (d), and heterogenous, (b) and (c), classes of local atomic configurations, as compared to the training set (dispersed vs clustered configurations). The NEB barriers are in abscissas; the ANN-predicted barriers are in ordinates. The number of training examples was 30 000. The average error, $\bar{e}$, and average bias, $\bar{b}$, are calculated using Eqs. (4) and (5), respectively. R^2 is Pearson's product-moment correlation coefficient.

the matrix [Fig. 3(a)], or only on configurations with solute atoms forming clusters [Fig. 3(b)]. The two differently trained ANNs have been then contrasted versus validation sets containing configuration of either the same type as in the training set, or different. The results are shown in Fig. 10: (b) and (c) correspond, respectively, to the case in which the "dispersed configurations" network is validated on "clustered configurations," and vice versa, the clustered configurations are used to make predictions for dispersed configurations. It can be seen that the predictions of the ANN remain reasonably accurate, even though the average error and bias are higher as compared to (a) and (d), where the ANN is validated on cases homogeneous with those of the training set. This test proves the robustness of the developed methodology.

V. SIMULATIONS OF THE THERMAL ANNEALING OF FE–CR ALLOYS

AKMC methods are often used to simulate thermal annealing processes.[1,2,5,10–12] In these simulations, the alloy studied is initially a random solid solution. Only one vacancy is introduced in the system and periodic boundary conditions are applied. With time, the atoms are progressively rearranged by the diffusing vacancy toward the final equilibrium

for the working temperature. If the model is thermodynamically correct, solute atoms such as Cu and Cr will cluster and eventually precipitate in Fe, if present in concentrations above their solubility limit, in accordance with the corresponding phase diagrams.[42] The ambition of AKMC models is to predict correctly also the kinetics of the precipitation process, which is an important issue in metallurgy, as it plays a role in determining mechanical property changes in the material,[43] e.g., 475 °C embrittlement in high-Cr steels.[44,45]

The ANN-based methodology detailed above has already produced results in qualitative, and partially also quantitative, agreement with experiments in the case of Fe–Cu and Fe–Cu–Ni alloys, that contain relatively small concentrations of solute atoms.[32] Here we apply it to concentrated Fe–Cr alloys, after having trained the ANN to vacancy migration energies obtained by NEB with Olsson's two-band model potential (fitted to PAW data).[28] We conducted simulations at 773 K, for 20% Cr, in boxes of 128 000 atoms. About 1.30 $\times 10^{10}$ AKMC events were computed in approximately 2 CPU months. The quality of the ANN prediction for this problem can be assessed from Fig. 8 (right side) and is very high. However, two aspects that cannot be deduced from Fig. 8 need to be assessed by carrying out the simulation. On the one hand, we need to verify that the accuracy of the E_m
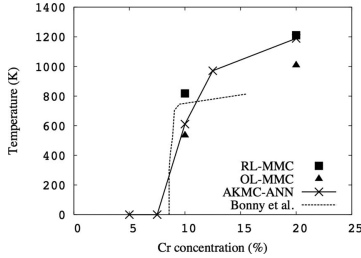
FIG. 11. Fe–Cr phase diagram, in the low Cr concentrations range, as predicted with different atomistic MC methods: rigid lattice Metropolis Monte Carlo (RL-MMC), off-lattice Metropolis Monte Carlo (OL-MMC), and ANN-based AKMC. The method followed to generate this diagram, as well as the MMC predicted phase diagrams, are described in Ref. 47. The dashed line is a recent revision of the experimental diagram (Ref. 33),

prediction by the ANN is sufficient to maintain a correct description of the thermodynamic properties of, in the present case, the Fe–Cr system, which the interatomic potential is known to provide. Second, we want to show that the use of a more accurate evaluation of the migration energies in the AKMC simulation actually improves the description of the kinetics of the process. In order to highlight the latter aspect, the simulation has also been carried out using a Kang–Weinberg-type heuristic formula to calculate E_m as

$$E_m = \varepsilon + \Delta E/2. \qquad (6)$$

Here ΔE is the rigid lattice total energy change between before and after the jump and ε is the excess migration energy. Olsson's interatomic potential[28] was used for the calculation of ΔE, by taking into account the interaction of the migrating pair with all the 77 neighbors encompassed by the potential cutoff (only 39 nearest neighbors were included in Ref. 12). The excess energy ε has been considered constant, equal to 0.65 eV in the case of a migrating Fe atom and to 0.56 eV in the case of a migrating Cr atom, as in Ref. 12. Clusters have been identified and counted by searching for the lattice sites where the local Cr concentration is larger than a given threshold concentration, as proposed in Ref. 11. Their average radius $\bar{R}$ was subsequently calculated on the basis of the average number $\bar{N}$ of atoms in the clusters, assuming spherical shapes

$$\bar{R} = \sqrt[3]{\frac{3\bar{N}}{8\pi}}. \qquad (7)$$

The local concentration was measured taking all neighboring sites encompassed by the potential cutoff into account. The threshold used was 90%Cr.

A. Thermodynamic consistency

Figure 11 shows data points for the Cr solubility limit in Fe, as predicted by various MC schemes, all based on the use of Olsson's interatomic potential[28] as Hamiltonian. The experimental phase boundary, as reviewed by Bonny et al. in

Ref. 33, is also shown. Metropolis MC methods[46,47] are unable to provide the kinetics of the diffusion process leading to precipitation. However, they can correctly account for all contributions to the free energy of the system, as stemming from the used Hamiltonian, including the vibrational contribution. They are therefore more suitable and reliable to trace the phase diagram embodied by a given Hamiltonian than AKMC models. In addition, if the possibility of displacing atoms off lattice is switched off in the Metropolis MC, the vibrational and strain-field contributions to the free energy are switched off too, and the "rigid lattice phase diagram" can be thereby obtained. Both Metropolis MC options have been used to compute the points in Fig. 11.[47] It can be seen that the solubility limit data points that were predicted by the ANN-based AKMC simulations are consistent with the Metropolis MC results. This shows that the good thermodynamic description of the Fe–Cr system provided by Olsson's potential is transferred almost untouched, via ANN regression of migration energy barriers, to the AKMC model. The fact that the ANN-based AKMC data points lie between those obtained by Metropolis MC in a rigid lattice and in an off-lattice approximation can be explained because the ANN-based AKMC allows for relaxation effects, but cannot allow for vibrational entropy effects, visible especially at high temperature. We therefore conclude that our AKMC model reproduces well the thermodynamic properties of the Fe–Cr system.

B. Description of the kinetics of the precipitation process

The kinetics of the precipitation process in Fe–Cr as predicted by the AKMC model is compared to experimental measurements by Jacquet,[48] Novy et al.,[49] and Bley[50] in Fig. 12. For the comparison, the simulation time had to be corrected, to take into account the fact that, due to the smallness of the simulation volume, the concentration of vacancies in the simulation is larger than in reality.[10,51] Thus, the Monte Carlo time t^{MC} should be rescaled proportionally to the ratio between the vacancy concentration in the simulation, C_v^{MC}, and the real vacancy concentration, C_v^{real}, at the considered temperature, namely,

$$t^{real} = t^{MC} \frac{C_v^{MC}}{C_v^{real}} \qquad (8)$$

with

$$C_v^{MC} = 1/N_b, \qquad (9)$$

$$C_v^{real} = 3 \exp\left(\frac{-H_v^f}{k_B T}\right). \qquad (10)$$

In these equations, N_b is the total number of atomic sites in the simulation box and H_v^f is the vacancy formation enthalpy. The numerical value of the latter predicted by the potential has been estimated, using the method described in Ref. 51, to be $H_v^f \approx 1.81$ eV for Fe–20%Cr alloys. However, the experimental value is unfortunately not precisely known ($H_v^f \approx 1.6, \ldots, 2.0$ eV). Since a small change in H_v^f compresses or dilates significantly the rescaled time, because of the ex-
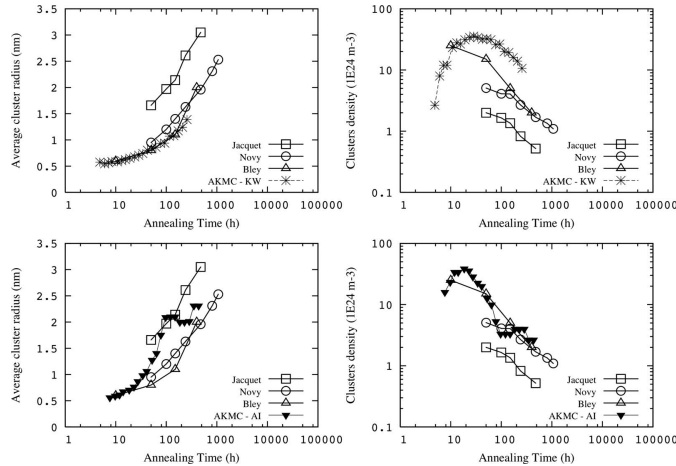
E. Paper IV

FIG. 12. Comparison of experimental average cluster radius (left) and cluster density (right) with the results from Fe–Cr thermal annealing simulations with AKMC, using two different migration energy calculation methods: Kang–Weinberg decomposition (above) and ANN (below). The time on the top figures has been rescaled using $H_{Vf}=1.68$ eV, whereas $H_{Vf}=1.80$ eV has been used for the bottom ones. The experimental values are taken from Refs. 48–50.

ponential function in Eq. (10), following Ref. 10, H_v^f has therefore been fitted in such a way that the first experimental value for the average cluster radius coincides with the first simulation point reaching the same radius. If the model is consistent and correct, this operation should lead to a superposition of all other experimental points with the simulation points, for a reasonable H_v^f value. In practice, the difficulty is also that the experimental data sets are not fully consistent with each other, so different choices are possible. In addition, we have produced two simulation data sets: those obtained with the Kang–Weinberg-type decomposition and those obtained with the ANN method. All the results are given in Fig. 12 and can be so summarized:

(i)   For the Kang–Weinberg driven AKMC, $H_v^f \approx 1.5$ to fit Jacquet's experiments,[48] but $H_v^f \approx 1.68$ eV to fit Bley's[50] and Novy's[49] (this is the case shown in the upper panel of Fig. 12). However, for either choice, while the predicted average cluster radius is in very good agreement with the experiments, the final cluster density is overestimated by almost an order of magnitude.

(ii)   For the ANN-based AKMC, $H_v^f \approx 1.80$ fits the first point of all experimental works. The rate of increase in the average cluster radius is somewhat larger than in the experiments, but the results can be considered very good, especially considering the discrepancy existing between experimental data sets. In addition, the cluster density is in much better agreement with the experiments than using the Kang-Weinberg-type decomposition. (Note that the AKMC curves exhibit a jerky shape because the number of clusters is very small, varying between 3, 4, and 5, at the end of the

simulation. To have better statistics and a smoother curve, a larger simulation box should be used, but the computational cost associated with an increased size becomes prohibitive.)

The improved quality of the ANN-based AKMC results versus experiments, compared to using the Kang–Weinberg decomposition, stems from the combination of two factors: on the one hand, it is the result of the high quality of the two-band model interatomic potential developed by Olsson et al.;[28] on the other, it is the result of having used a methodology, the coupling of ANN to the AKMC simulation, that allows all features of the potential to be exploited and transferred to the AKMC model, including long-range chemical and strain-field effects, neglected in rigid lattice simulation using e.g., a Kang–Weinberg-like decomposition.

Nevertheless, not all results of the simulation compare one-to-one with the experiments. Figure 13 shows for example the evolution versus time of the Cr concentration in the matrix and in the clusters from experiments and according to the ANN-based AKMC. We see that Cr depletion in the matrix is faster in the simulation than in the experiment.[48,49] Consistently, the build up of the equilibrium concentration of Cr in the precipitates in the simulation is totally different from the experimentally observed one and much faster. Essentially, in the simulation Cr clusters are created in thermodynamic equilibrium since the beginning, this implying that the matrix is rapidly depleted and the only readjustement is the emission of some Cr atoms. On the contrary, the experiments suggest that precipitates are initially diffuse and only at the end approach the equilibrium concentration. However, this discrepancy may also partly be due to
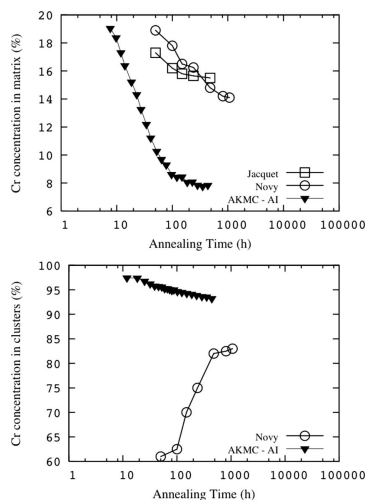
155

FIG. 13. Comparison of atom probe tomography (Ref. 49) and SANS (Ref. 48) study of the experimental Cr concentration in the matrix (above) and in the clusters (below) with the Fe–Cr thermal annealing simulations performed with the ANN-based AKMC. Time has been rescaled using H_{Vf} =1.80 eV.

been trained on examples of significant qualitative difference from those used for the validation. Thus, the methodology we have developed is equivalent to performing an on-the-fly nudged elastic band calculation to determine the frequency of the possible vacancy jumps for use in the atomistic kinetic Monte Carlo, at the same computational cost as using far less precise and reliable heuristic approaches. In particular, the effects of zero temperature relaxation on the migration energies are properly taken into account in our model, despite keeping a rigid lattice description of the system. Thermal annealing simulations in a binary Fe–20%Cr alloy have produced results in better agreement with experiments than those obtained with a Kang–Weinberg-type heuristic approach, while using the same interatomic potential for energy calculations. We therefore believe that the methodology presented here is promising and suitable to be extended to more complex cases. In particular, this methodology is now being extended to the more general problem of point-defect migration, by allowing for more than one vacancy in the atomistic kinetic Monte Carlo box and by introducing self-interstitials. This is not straightforward, as anticipated in Ref. 21, mainly because in these cases not all the logically possible system configurations are stable, but work in this direction is currently in progress.

different definitions of clusters in the experimental analysis as compared to the simulation, as well as to limitations in the precision of the experimental technique (atom probe tomography in this case).

VI. CONCLUDING REMARKS

In this paper we have described in detail how an artificial neural network can be trained to calculate on-the-fly accurate values for vacancy migration energies as functions of the local environment in atomistic kinetic Monte Carlo simulations, used to model the microchemical processes induced by single-vacancy migration in concentrated alloys. We have shown that the vacancy migration energy is determined not only by the chemical nature of close-by atoms, but also by relatively faraway atoms, mainly due to strain-field effects. This obliges, for accurate calculations using, e.g., the nudged elastic band method, large local atomic environments (up to hundreds of atoms) to be taken into account, something that heuristic methods based on the use of broken-bond approximations cannot do and that is often disregarded also in methods based on the Kang–Weinberg decomposition. The procedure to train the artificial neural network to predict the values of energy barriers calculated using the nudged elastic band method has been described and analyzed in detail. A gradually improving accuracy constructive algorithm has been specifically developed for this purpose, as the appropriate tool to handle such a large-size mathematical regression problem. Moreover, the neural network extrapolation skills have been shown to be trustable, even when the network has

[1] S. Schmauder and P. Binkele, Comput. Mater. Sci. **25**, 174 (2002).
[2] Y. Le Bouar and F. Soisson, Phys. Rev. B **65**, 094103 (2002) and references therein.
[3] P. Bellon, *Thermodynamics, Microstructures and Plasticity* (Kluwer Academic, Dordrecht, The Netherlands, 2003), Vol. 108, and references therein.
[4] F. Soisson, J. Nucl. Mater. **349**, 235 (2006).
[5] F. Soisson and C.-C. Fu, Phys. Rev. B **76**, 214102 (2007).
[6] E. Vincent, C. S. Becquart, and C. Domain, Nucl. Instrum. Methods Phys. Res. B **255**, 78 (2007).
[7] A. Chatterjee, D. G. Vlachos, and J. Computer-Aided, Mater. Des. **14**, 253 (2007).
[8] P. Krasnochtchekov, R. S. Averback, and P. Bellon, Phys. Rev. B **75**, 144107 (2007).
[9] E. Vincent, C. S. Becquart, and C. Domain, J. Nucl. Mater. **382**, 154 (2008).
[10] E. Vincent, C. S. Becquart, C. Pareige, P. Pareige, and C. Domain, J. Nucl. Mater. **373**, 387 (2008).
[11] G. Bonny, D. Terentyev, and L. Malerba, Comput. Mater. Sci. **42**, 107 (2008).
[12] G. Bonny, D. Terentyev, L. Malerba, and D. Van Neck, Phys. Rev. B **79**, 104207 (2009).
[13] C. S. Becquart and C. Domain, Phys. Status Solidi B **247**, 9 (2010).
[14] F. Soisson, C. S. Becquart, N. Castin, C. Domain, L. Malerba, and E. Vincent, "Atomistic kinetic Monte Carlo studies of microchemical evolutions driven by diffusion processes under irradiation," J. Nucl. Mater. (in press), and references therein.
[15] H. C. Kang and W. H. Weinberg, J. Chem. Phys. **90**, 2824 (1989).
[16] G. Henkelman and H. Jonsson, J. Chem. Phys. **115**, 9657 (2001).
[17] G. Henkelman and H. Jonsson, J. Chem. Phys. **111**, 7010 (1999).

[18] A. Van der Ven and G. Ceder, Phys. Rev. B **71**, 054102, (2005).

[19] K. Sastry, D. D. Johnson, D. E. Goldberg, and P. Bellon, Phys. Rev. B **72**, 085438 (2005).

[20] N. Castin, R. P. Domingos, and L. Malerba, Int. J. Comput. Intell. Syst. **1**, 340 (2008).

[21] N. Castin and L. Malerba, Nucl. Instrum. Methods Phys. Res. B **267**, 3148 (2009).

[22] H. Jónsson, G. Mills, and K. W. Jacobsen, in *Classical and Quantum Dynamics in Condensed Phase Simulations*, edited by B. J. Berne, G. Ciccotti, and D. F. Coker (World Scientific, Singapore, 1998).

[23] G. Henkelman and H. Jónsson, J. Chem. Phys. **113**, 9901 (2000).

[24] C. M. Bishop, *Neural Networks for Pattern Recognition* (Clarendon, Oxford, 1995).

[25] W. M. Young and E. W. Elcock, Proc. Phys. Soc. London **89**, 735 (1966).

[26] B. Bortz, M. H. Kalos, and J. L. Lebowitz, J. Comput. Phys. **17**, 10 (1975).

[27] K. A. Fichthorn and W. H. Weinberg, J. Chem. Phys. **95**, 1090 (1991).

[28] P. Olsson, J. Wallenius, C. Domain, K. Nordlund, and L. Malerba, Phys. Rev. B **72**, 214119 (2005).

[29] R. C. Pasianot and L. Malerba, J. Nucl. Mater. **360**, 118 (2007).

[30] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).

[31] G. Henkelman, G. Jóhannesson, and H. Jónsson, in *Methods for Finding Saddle Points and Minimum Energy Paths, Progress on Theoretical Chemistry and Physics*, edited by S. D. Schwartz (Kluwer Academic, Dordrecht, The Netherlands, 2000), pp. 269–300.

[32] N. Castin, L. Malerba, G. Bonny, M. I. Pascuet, and M. Hou, Nucl. Instrum. Methods Phys. Res. B **267**, 3002 (2009).

[33] G. Bonny, D. Terentyev, and L. Malerba, Scr. Mater. **59**, 1193 (2008).

[34] S. E. Fahlman and C. Lebiere, *Advances in Neural Information Processing Systems*, The Cascade-Correlation Learning Architecture Vol. 2 (Morgan Kaufmann Publishers, Inc., San Francisco, CA, 1990), pp. 524–532.

[35] K. Levenberg, Q. Appl. Math. **2**, 164 (1944).

[36] D. Marquardt, SIAM J. Appl. Math. **11**, 431 (1963).

[37] T.-Y. Kwok and D.-Y. Yeung, IEEE Trans. Neural Netw. **8**, 630 (1997).

[38] C. Campbell, *Neural Network Systems Technologies and Applications*, Constructive Learning Techniques for Designing Neural Network Systems, edited by C. T. Leondes (Academic, San Diego, 1997).

[39] T. Ash, Connect. Sci. **1**, 365 (1989).

[40] G. Bonny, R. C. Pasianot, and L. Malerba, Modell. Simul. Mater. Sci. Eng. **17**, 025010 (2009).

[41] G. Bonny, R. C. Pasianot, N. Castin, and L. Malerba, Philos. Mag. **89**, 3531 (2009).

[42] T. B. Massalsky, H. Okamoto, P. R. Subramanian, and L. Kacprzac, *Binary Alloy Phase Diagrams* (ASM International, Materials Park, 1990), p. 1273.

[43] D. A. Terentyev, G. Bonny, and L. Malerba, Acta Mater. **56**, 3229 (2008).

[44] R. O. Williams and H. W. Paxton, J. Iron Steel Inst. **185**, 358 (1957).

[45] R. O. Williams, Trans. Metall. Soc. AIME **212**, 497 (1958).

[46] K. Binder, in *Monte Carlo Methods in Statistical Physics*, edited by K. Binder (Springer-Verlag, Berlin, 1979), Chap. 1.

[47] E. E. Zhurkin, R. Pereira, N. Castin, L. Malerba, and M. Hou, Mater. Res. Soc. Symp. Proc. **1125**, 121 (2009).

[48] V. Jacquet, "Effet de l'irradiation sur la démixtion des alliages modèles Fe–Cr autour de 15% de Cr," Ph.D. thesis, Ecole Polytechnique, 2000.

[49] S. Novy, P. Pareige, and C. Pareige, J. Nucl. Mater. **384**, 96 (2009).

[50] F. Bley, Acta Metall. Mater. **40**, 1505 (1992).

[51] D. Terentyev, N. Juslin, K. Nordlund, and N. Sandberg, J. Appl. Phys. **105**, 103509 (2009).

F

Paper V

Modelling phase separation in Fe–Cr system using different atomistic kinetic Monte Carlo techniques

N. Castin [a,b], G. Bonny [a], D. Terentyev [a,*], M.Yu. Lavrentiev [c], D. Nguyen-Manh [c]

[a] Structural Material Group, Institute of Nuclear Materials Science, SCK·CEN, Mol, Belgium
[b] Physique des Solides Irradiés et des Nanostrucutres CP234, Faculté des Sciences, Université Libre de Bruxelles, Bd du Triomphe, B-1050 Bruxelles, Belgium
[c] EURATOM/UKAEA Fusion Association, Culham Science Centre, Abingdon OX14 3DB, United Kingdom

ARTICLE INFO

ABSTRACT

Atomistic kinetic Monte Carlo (AKMC) simulations were carried out to study the α–α' phase separation in Fe–Cr alloys. Two different energy models and two approaches to estimate the local vacancy migration barriers were used. The energy models considered are a two-band model Fe–Cr potential and a cluster expansion, both fitted to ab initio data. The classical Kang–Weinberg decomposition, based on the total energy change of the system, and an Artificial Neural Network (ANN), employed as a regression tool were used to predict the local vacancy migration barriers 'on the fly'. The results are compared with experimental thermal annealing data and differences between the applied AKMC approaches are discussed. The ability of the ANN regression method to accurately predict migration barriers not present in the training list is also addressed by performing cross-check calculations using the nudged elastic band method.

© 2010 Published by Elsevier B.V.

1. Introduction

Conventional and newly developed high-Cr ferritic martensitic steels are promising candidate structural materials for future nuclear power systems. These steels as well as binary Fe–Cr alloys undergo phase separation after thermal annealing (e.g. [1–4]) with the formation of finely dispersed nano-metric nearly-pure-Cr precipitates (α' phase) coherent with the matrix. The latter cause hardening and provoke embrittlement of these steels containing more than 14 at.% Cr after thermal annealing at about 475 °C [1,2,5], as well as, at lower temperature and Cr content, under irradiation, found to accelerate or possibly induce the phase separation [5,6]. Therefore, the development of models to predict quantitatively the kinetics of phase separation is of great importance for the design of complex structural materials relevant to the nuclear sector.

Tracing the kinetics of precipitation in concentrated alloys from its very first stage is not trivial due to the small size and high density of the formed precipitates. Thermal annealing experiments, where nano-metric coherent precipitates can be identified, are complicated and time consuming (years), especially in active specimens. In the literature, only few experimental works providing detailed information about size and density distribution for the α' precipitation in Fe–Cr alloys (containing 20 at.% Cr) are available

up to now [3,4]. Computer simulations seem therefore to be an attractive solution to complement the available experimental data, especially for alloys containing <20 at.% Cr annealed at a temperature below 800 K, where little or no experimental data so far exist.

From the point of view of modelling, the consideration of the diffusion-mediated phase separation requires a description of the system at the atomic level, from which the precipitate composition, its morphology and rearrangement of solutes in the depleted matrix follows naturally. Here, we carried out atomistic kinetic Monte Carlo (AKMC) simulations to study the kinetics of the phase separation in the Fe–20Cr alloy at 773 K, which can be compared to the available experimental data. To model the thermal annealing, two different algorithms to estimate the vacancy migration energy determined by a local atomic environment (LAE) were applied. These are (i) a regression algorithm based on an artificial neural network (ANN) and (ii) the classical Kang–Weinberg (KW) decomposition [7] based on the total energy change due to the vacancy migration. For our simulations, we also considered two different energy models, namely: (i) a cluster expansion (CE) [8] fitted to the formation energy of ab initio calculated intermetallic compounds and (ii) a two-band model interatomic potential (2BMP) from [9]. During our simulations we tracked the average precipitate size and density as a function of simulation time and compared the obtained results between the different methods and energy models and also to experimental data. Based on these comparisons, we discuss the impact of the choice of the energy model and algorithm to estimate the local vacancy migration barrier on the simulation results.

* Corresponding author. Address: Boeretang 200, B2400 Mol, Belgium. Tel.: +32 014 33 21 97; fax: +32 014 32 12 16.
E-mail address: dterenty@sckcen.be (D. Terentyev).

ARTICLE IN PRESS

2

2. Description of the used techniques

Rigid lattice AKMC simulations [10] were performed in a cubic box containing 0.128×10^6 atoms with an initial random distribution of 20% Cr atoms in the Fe matrix at 773 K. The evolution of the system leading to the atomic redistribution is driven by a single vacancy, jumping at a rate evaluated as, $v = v_0 \times \exp(-E_m/k_B T)$, where v_0 is an attempt frequency (set to 10^{13} and 6×10^{12} s^{-1} using the CE and 2BMP models, respectively), k_B is Boltzmann's constant, T is the absolute temperature and E_m is the LAE-dependent vacancy migration energy. Each vacancy jump corresponds to one MC step. The jump to be performed is chosen based on its probability, evaluated in terms of the corresponding jump frequency v. The time between two jumps is calculated according to the mean residence time algorithm i.e. by taking the inverse of the sum of the eight (only 1st nearest neighbour jumps are considered) possible jump frequencies [10]. Therefore, E_m is estimated 'on the fly' during the AKMC run at each step. Accurate methods, such as the drag method [11] or nudged elastic band (NEB) method [12] are time consuming and cannot be effectively used in long-term AKMC simulations.

The Kang–Weinberg decomposition [7] is one of simplest approximations to estimate the LAE-dependent E_m as $E_0 + \Delta E_{f-i}/2$, where ΔE_{f-i} is the total energy change due to the vacancy jump and E_0 is the excess migration energy. ΔE is calculated on-the-fly according to the applied cohesive model. E_0 is taken as the migration barrier of the atomic species exchanging its position with the vacancy (Fe or Cr in our case), in the limit of infinitely dilute solution, as obtained from DFT calculations, namely for the CE E_0(Fe) = 0.64 eV and E_0(Cr) = 0.57 eV [13] and for the 2BMP E_0(Fe) = 0.65 eV and E_0(Cr) = 0.56 eV, according to the prediction of the potential. In this simplified scheme, E_0 is determined only by the migrating species, assuming that it has negligible dependence on the LAE, accounted for only in the ΔE_{f-i} term. It is therefore the total energy difference between equilibrium states that directly influences the kinetics of the phase separation, rather than the actual local vacancy migration barrier.

The ANN regression scheme is trained using an extended dataset of local vacancy migration barriers, calculated with the NEB method for a variety of LAEs. Once trained and provided with a LAE around a vacancy (up to 11th nearest neighbour distance), the ANN can predict the local vacancy migration barrier accurately and with the same speed as the KW decomposition. The only required input is consequently the training data set that could be obtained by applying interatomic potentials or *ab initio* techniques. A detailed description of the method, can be found in [14,15], and is briefly described here. The regression is built using an adequate amount ($\sim 10^5$) of NEB-obtained migration barriers generated for randomly chosen LAEs around a vacancy. The latter inherently includes 0 K relaxation effects and is thus *a priory* more accurate than the KW approach. Given that a database must contain a significant number of barriers, we used the 2BMP to generate it, although *ab initio* methods could also be used for this purpose. The accuracy of the 2BMP to predict E_m for a vacancy in concentrated Fe–Cr alloys was validated by a cross-check comparison with density functional theory (DFT) calculations, as presented in the upper part of Fig. 1, where the migration energies calculated using the NEB method are shown. The 2BMP quite accurately reproduces the DFT data in the Fe-rich region, where E_m is rather low, which is of primary importance for kinetic problem. In the Cr-rich alloys, however, the 2BM values deviate significantly (up to 30%) from the corresponding DFT value. The reason for the discrepancy at high Cr content is most likely related to the complex magnetic behaviour of anti-ferromagnetic Cr atoms embedded in Fe matrix, responsible for the Cr–Cr repulsion and multiple Cr–Cr interaction.
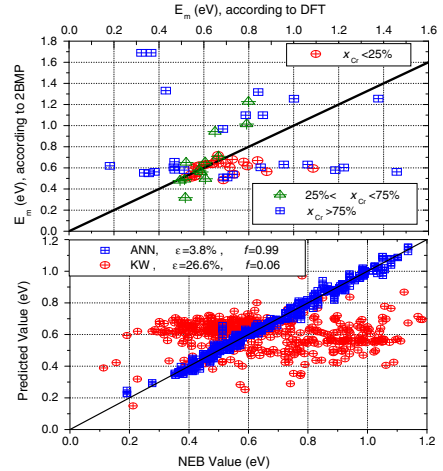

Fig. 1. The local vacancy migration barrier for different LAEs calculated using NEB–2BMP and NEB–DFT methods, shown in the upper part, and NEB–2BMP, ANN–2BMP and KW–2BMP methods, shown in the downer part. x_{Cr}, ε and f denote Cr content, the mean error and correlation factor, respectively.

The performance of both the ANN approach and KW decomposition is also compared in Fig. 1 (see downer part). There, E_m obtained with the latter methods is compared to the NEB obtained values. The mean error (ε) and correlation factor (f) between the NEB and ANN values were calculated to be 3.8% and 0.99, respectively, whereas ε = 26.6% and f = 0.06 when comparing NEB and KW methods. Clearly, the ANN approach outperforms the KW decomposition in reproducing the NEB values and for the simulations performed here the ANN approach is equivalent to NEB.

3. Results, discussion and conclusive remarks

The thermal annealing of Fe–20Cr at 773 K was simulated by three AKMC runs. Two runs involved the 2BMP where both the KW decomposition and ANN approach were applied, while the third run used the CE in combination with the KW decomposition. The CE model does not include atomic relaxation and therefore the ANN was not constructed for the CE. The KW–2BMP simulation is fully described in [16], and the ANN–2BMP simulation is fully described in [15]. Precipitates were identified during the runs following the method proposed in [17], specifically developed to identify non-pure coherent precipitates in a concentrated alloy. The obtained size and density distributions of Cr precipitates versus AKMC time are shown in Fig. 2. From these plots it follows that in all three runs the precipitation process goes via three overlapping stages: nucleation, growth and coarsening. While the three approaches provide a very similar evolution of the precipitate size and density distributions, they give essentially different kinetics. For example, the time evolution with the ANN approach is an order of magnitude faster than with the KW decomposition applied using the 2BMP.

An adequate comparison of the AKMC results with experimental data requires renormalization of the AKMC time (t_{MC}) to ac-
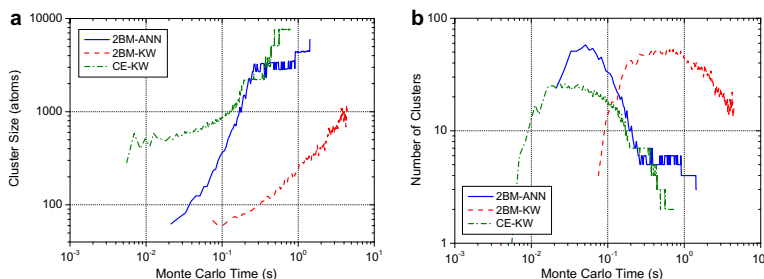
Fig. 2. (a) Average precipitate size and (b) density versus AKMC time calculated using the different AKMC approaches.

count for the discrepancy between the vacancy concentration in the modelled crystal ($C_V^{MC} = 1/128,000$) and in the real specimen during annealing (C_V^{real}). As recently suggested in [16,18], we re-scaled MC time separately for each run so as to fit the first experimental data point for the average precipitate size. The comparison of the obtained size and density distributions with experiment is presented in Fig. 3, which shows that the best overall agreement for both size and density evolution is achieved when using the ANN approach.

Considering the ambiguity of the applied time rescaling procedure, we estimated the vacancy formation energy (E_f^V) corresponding to the obtained time rescaling factors in all three runs. The MC time is related to experimental annealing time as $t_{real} = C_V^{MC}/C_V^{real} \times t_{MC}$ and the thermal equilibrium vacancy concentration is $K \cdot \exp(-E_f^V/k_B T)$, here K is a constant factor approximately equal to 3, due to the configurational entropy [19]. The application of the above-used conversion factors allows to estimate E_f^V to be 1.80 eV and 1.67 eV and 1.92 eV for the ANN–2BMP, KW–2BMP and KW–CE runs, respectively. Thus, in all the cases the calculated E_f^V stays in the narrow range of 1.8 ± 0.2 eV. Note that in pure Fe, different DFT approximations provide a spread for the vacancy formation energy of ±0.2 eV with the mean value of ∼2.05 [20]. Similar spread exists for the vacancy formation in pure Cr, but the average value is about 2.7 eV [20]. The 2BMP, predicts E_f^V in both Fe and Cr metals to be lower by about 0.2 eV. The vacancy formation energy calculated from independent static simulations performed in ran-

dom Fe–20Cr and Fe–10Cr alloys with the 2BMP is 1.88 eV and 1.76 eV, respectively, consistent with the numbers deduced from the time renormalization factors.

Finally, given that the ANN regression was obtained from the database of E_m values calculated in the random alloys, it is important to clarify that it also possesses good accuracy in the alloy that has already undergone the phase separation. Hence, we extracted a set of configurations encountered during the AKMC runs (on the coarsening stage), corresponding to outstandingly high (∼1 eV) or low (∼0.3 eV) values. For these configurations, the barriers were correspondingly recalculated using the NEB method applied with the 2BMP, and the results are shown in Fig. 4. A good correspondence between the ANN and NEB–2BM data can be clearly seen. Thus, all high and low migration barriers spotted in the ANN approach are confirmed by independent NEB calculations.

To conclude, the three different simulations provide qualitatively similar curves for the precipitate size and density distribution, but essentially differ in their time evolution. Comparisons suggest that these results, obtained using the ANN approach (based on the 2BMP), are the most consistent with experimental data compared to KW–CE and KW–2BMP. Additional cross-checks using the NEB method for extreme values of E_m encountered during the simulation, has proven the reliability of the ANN approach. A possible improvement for the AKMC simulation is to apply the ANN approach using a DFT calculated database of migration energies, therefore avoiding the use of a simplified energy model such as
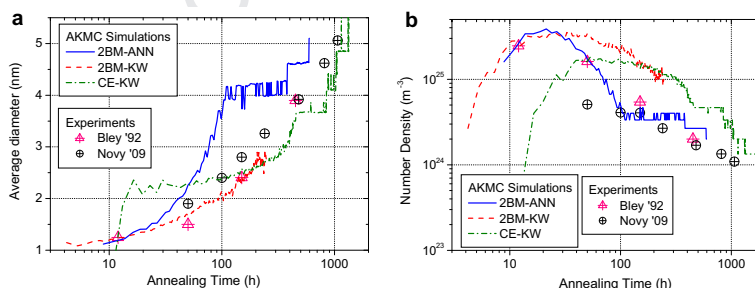

Fig. 3. (a) Average precipitate size and (b) density versus experimental annealing time calculated with the different AKMC approaches and compared with experimental data reported by Bley [3] and Novy et al. [4].
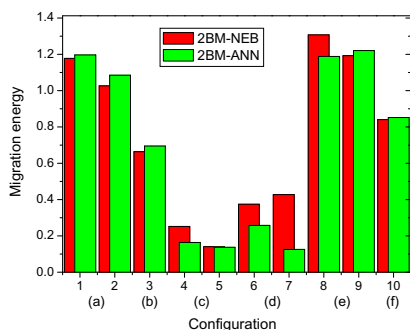
Fig. 4. Cross-check of outstandingly high and low E_m encountered during AKMC run employing ANN. Values of E_m are recalculated with NEB–2BMP and NEB–DFT approaches. (a: conf. 1–2) Cr jumps near clusters (high values). (b: conf. 3) Fe jump near cluster (high value). (c: conf. 4–5) Low migration Cr jumps from the random database. (d: conf. 6–7) Low migration Fe jump from the random database. (e: conf. 8–9) High migration Cr jump from the random database. (f: conf. 10) High migration Fe jump from the random database.

CE or 2BMP, and naturally taking zero temperature relaxation effects into account. This is however not easy, because the automatic calculation of energy barriers with DFT for any LAE is delicate to settle in practice, and finally the amount of examples required to train the ANN ($\sim 10^5$) is such that computing time is hardly affordable. The ANN–2BMP approach therefore remains an interesting alternative to it.

References

[1] R.M. Fisher, E.J. Dulis, K.G. Carroll, Trans. AIME 197 (1953) 690.
[2] P.J. Grobner, Metall. Trans. 4 (1973) 251.
[3] F. Bley, Acta Metall. Mater. 40 (1992) 1505.
[4] S. Novy, P. Pareige, C. Pareige, J. Nucl. Mater. 384 (2009) 96.
[5] P. Dubuisson, D. Gilbon, J.L. Séran, J. Nucl. Mater. 205 (1993) 178.
[6] M.H. Mathon, Y. de Carlan, G. Geoffroy, X. Averty, A. Alamo, C.H. de Novion, J. Nucl. Mater. 312 (2003) 236.
[7] H.C. Kang, W.H. Weinberg, J. Chem. Phys. 90 (1989) 2824.
[8] M.Yu. Lavrentiev, R. Drautz, D. Nguyen-Manh, T.P.C. Klaver, S.L. Dudarev, Phys. Rev. B 75 (2007) 014208.
[9] (a) P. Olsson, J. Wallenius, C. Domain, K. Nordlund, L. Malerba, Phys. Rev. B 72 (2005) 214119;
(b) P. Olsson, J. Wallenius, C. Domain, K. Nordlund, L. Malerba, Phys. Rev. B 74 (2006) 1. E.
[10] W.M. Young, E.W. Elcock, Proc. Phys. Soc. 89 (1966) 735.
[11] G. Henkelman, G. Jóhannesson, H. Jónsson, Methods for finding saddle points and minimum energy paths, in: S.D. Schwartz (Ed.), Progress on Theoretical Chemistry and Physics, Kluwer Academic Publishers, 2000, p. 269.
[12] H. Jonsson, G. Mills, K.W. Jacobsen, in: B.J. Berne, G. Ciccotti, D.F. Coker (Eds.), Nudged Elastic Band Method for Finding Minimum Energy Paths of Transitions, Classical and Quantum Dynamics in Condensed Phase Simulations, World Scientific, Singapore, 1998.
[13] D. Nguyen-Manh, M.Yu. Lavrentiev, S.L. Dudarev, C.R. Phys. 9 (2008) 379.
[14] N. Castin, R.P. Domingos, L. Malerba, Int. J. Compos. Intell. Syst. 1 (2008) 340.
[15] N. Castin, L. Malerba, J. Chem. Phys. 132 (2010) 074507.
[16] G. Bonny, D. Terentyev, L. Malerba, D. Van Neck, Phys. Rev. B 79 (2009) 104207.
[17] G. Bonny, D. Terentyev, L. Malerba, Comput. Mater. Sci. 42 (2008) 107.
[18] E. Vincent, C.S. Becquart, C. Pareige, P. Pareige, C. Domain, J. Nucl. Mater. 373 (2008) 387.
[19] D.A. Porter, K.E. Easterling, Phase Transformations in Metals and Alloys, second ed., CRC Press, 1992.
[20] P. Olsson, C. Domain, J. Wallenius, Phys. Rev. B 75 (2007) 014110.

G Paper VI

Contents lists available at ScienceDirect

Journal of Nuclear Materials

Prediction of radiation induced hardening of reactor pressure vessel steels using artificial neural networks

N. Castin [a,b,*], L. Malerba [a], R. Chaouadi [a]

[a] Studie Centrum voor Kerneenergie – Centre d'Etudes de l'énergie Nucléaire (SCK•CEN), NMS unit, Boeretang 200, B2400, Mol, Belgium
[b] Université Libre de Bruxelles (ULB), Physique des Solides Irradiés et des Nanostructures (PSIN), Boulevard du Triomphe CP234, 1050 Brussels, Belgium

ARTICLE INFO

ABSTRACT

In this paper, we use an artificial neural network approach to obtain predictions of neutron irradiation induced hardening, more precisely of the change in the yield stress, for reactor pressure vessel steels of pressurized water nuclear reactors. Different training algorithms are proposed and compared, with the goal of identifying the best procedure to follow depending on the needs of the user. The numerical importance of some input variables is also studied. Very accurate numerical regressions are obtained, by taking only four input variables into account: neutron fluence, irradiation temperature, and chemical composition (Cu and Ni content). Accurate extrapolations in term of neutron fluence are obtained.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

It is well known that reactor pressure vessel (RPV) steels used in light water nuclear reactors harden and embrittle under neutron irradiation. Embrittlement is customarily measured in terms of increase of the ductile-to-brittle transition temperature (DBTT), measured by means of Charpy tests. Nuclear regulations impose safety margins on this increase, according to rules that may somewhat change depending on the country, as safeguard against RPV failure in both service and accidental conditions [1,2]. Radiation embrittlement of materials depends *a priori* on many variables: not only neutron fluence, flux, and energy spectrum, but also irradiation temperature, chemical composition, and pre-irradiation material history [3]. All these variables must be simultaneously considered to reliably predict pressure vessel embrittlement. However, in order to be able to assess the effect of the different variables on the mechanical response of the steels, each of them should be varied independently of the others in a sufficiently wide range. Such an approach is clearly unrealistic.

Although inadequate to cover all possible conditions, a large amount of data from surveillance capsules and from material test reactors does exist. One of the most important goals for utilities and other nuclear stakeholders is the development, based on "clever" interpolations and extrapolations of the available data, of reliable trend curves, providing estimates of steel embrittlement as a function of the most important among the above-mentioned variables [2].

Artificial intelligence is the combination of algorithms, data and software used to develop computer systems that can be defined *intelligent*. One defining feature of intelligence is the capability of learning from past experience and solving problems when important information is missing, so as to be able to handle complex situations and to correctly react to new circumstances. There are many different computational models which are considered branches of the artificial intelligence field, each one suited to a different kind of problem. Artificial neural networks (ANN), for instance, provide a general framework for representing non-linear functional mappings between a set of input variables and a set of output functions [4]. The list of successful applications of ANN to real-life problems is endless, in sundry domains of interests, e.g. character and image recognition, image compression, stock market prediction, tumor detection in medical image analysis, vehicle piloting, etc. The interested reader can find general information for example in [5,6].

In the field of nuclear materials, ANN have been applied by Kemp et al. [7] to the analysis and prediction of the yield strength increase ($\Delta\sigma_Y$) induced by irradiation in low activation ferritic/martensitic steels, which are candidate structural materials for future nuclear fusion reactors. Their conclusion was very encouraging concerning the ANN capability of analyzing irradiation damage, at least within the range of irradiation parameters and steel composition that are covered in the database used for training. Later on, Windsor et al. [8] have shown that the network can also be used for extrapolating to fluences higher than those included in the training database. In this work, we use a different ANN approach to construct a mathematical regression of the radiation-induced $\Delta\sigma_Y$, as a function of irradiation parameters and steel composition. Differently from [7,8], we compare two different

* Corresponding author at: Studie Centrum voor Kerneenergie – Centre d'Etudes de l'énergie Nucléaire (SCK•CEN), NMS unit, B2400 Mol, Belgium.
E-mail address: ncastin@sckcen.be (N. Castin).

ANN training approaches (classical and Bayesian), and try to identify the most suitable, depending on the purpose of the trained ANN. Moreover, we define and compare two different algorithms to split the available data in training and validation sets, because this aspect of the ANN training problem is very important for applications where the available amount of training data is limited.

The objectives of this work are manifold:

1. To exploit, as effectively as possible, the information contained in the available databases from surveillance and material test reactors irradiations, for both steels and alloys.
2. To identify in a systematic way the variables that appear to be of higher or lesser importance, based on the available data, within the ranges covered, i.e. based on interpolations. For example, the possible existence of a flux effect is addressed.
3. To attempt an extrapolation outside the ranges covered by the databases and evaluate the reliability of these extrapolations, by assessing the capability of the ANN to predict a certain category of data when trained on a different category. For example, prediction of the evolution of hardening for higher fluences will be attempted.
4. To provide a guide to design future irradiation experiments on steels and alloys, in order to better understand specific effects and dependencies.

Fully reliable predictions will only be possible once the important physical mechanisms acting during irradiation have been identified, understood and quantified at all relevant scales, from the atomic scale to the component scale. However, an empirical approach based on advanced regression techniques, such as ANN, can be beneficial for industrial applications within a shorter delay and can even be useful to guide the longer term development of physical models.

In Section 2, we describe in detail the different ANN training approaches that we propose, and provide some theoretical background for the reader unfamiliar with these techniques. Then, we briefly describe, in Section 3, the RADAMO database that is used throughout this paper. In Section 4, training experiments are reported, aimed at identifying which, according to our ANN, are the most influential factors for hardening. Finally, in Section 5 we compare the different algorithms proposed in Section 2 in order to establish which combination is the most suitable for extrapolation under given conditions.

2. Methodology

Artificial neural networks (ANN) are powerful computational models, capable of providing efficient numerical regressions even when many input variables are involved. In this work, we use the classical feed-forwards multi-layer perception [4] with one hidden layer, linear combination functions and hyperbolic tangent activation functions, as depicted in Fig. 1. It is a *universal approximator* in the sense that it can approximate any continuous multivariate function to any desired degree of accuracy, provided that enough hidden nodes are available [9,10]. The *universal approximation theorem*, however, does not provide a theoretical framework for training ANN, but only demonstrates the existence of at least one ideal architecture for any regression problem, without guaranteeing that it can be found by training, and without giving an estimate of the number of training examples that must be provided.

The output of the ANN shown in Fig. 1 can be written as:

$$\Delta\sigma_Y = \tanh\left(w_{00} + \sum_{j=1}^{H} w_{0j}\tanh\left(w_{j0} + \sum_{k=1}^{4} w_{jk}\cdot i_k\right)\right) \qquad (1)$$

where H is the number of hidden nodes, i are the input variables. The fixed coefficients w_{00}, w_{0j}, w_{j0} and w_{jk} are the *synaptic weights* (also often called *synapses*). The advantage of the ANN method is that this generic expression does not require the user to explicitly state how input variables and outputs are related to each others, unlike the usual trend curves [11,12]. All input variables are connected to all hidden nodes, and these play the role of simple processing units that, connected in network, can reproduce complex mappings that are not necessarily visible to the user. The drawback, however, is that no reasonable physical interpretation of the individual terms in Eq. (1) can be given.

ANN training, i.e. the problem of determining the optimal number H of hidden nodes and the optimal numerical value for the synaptic weights in Eq. (1), is in practice solved as a non-linear optimization problem and has many empirical aspects. A database of input/output examples – for this application exclusively coming from either neutron irradiation experiments or nuclear power plants surveillance programs – is used as target for fitting. A good practice is to separate this database into a training and a reference set; the training set is only used to optimize the ANN, and the reference set is used to assess the accuracy of its predictions for new cases. The problem of how to define such sets is addressed in the next section.

2.1. Algorithms for the definition of training and reference sets

A good practice, before training the ANN, is to define training and reference sets from the available database. It is essential to make sure that both sets are equally representative of the domains in the input and output spaces, without overlap, in order to provide a rich training set that contains enough learning material. On the other hand, it is also important to keep enough pertinent and, most

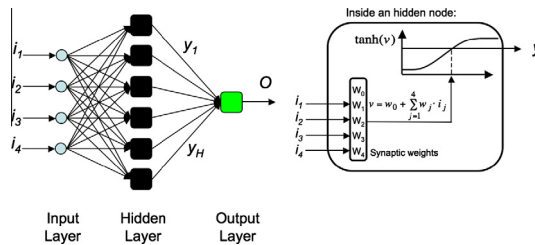**Fig. 1.** Schematic representation of an artificial neural network (ANN) with four input variables (i_1–i_4), one hidden layer with six hidden nodes and one output (O). The input signals are propagated from left to right, in a layer-by-layer fashion, without feedback connection and without layer by-pass. The right part of the figure shows how a node processes its input signals into its own output.

importantly, independent examples for the validation of the trained ANN.

In some applications, the database can simply be split by "shuffling and cutting", like one would do with playing cards. This simple algorithm is, however, unsuitable for this application, for the following reasons:

(A) The distribution of points in the database input/output space is generally non-homogenous, as illustrated in Fig. 2. In particular, some regions are very sparsely populated, usually by just a couple of points. A completely random "shuffle-and-cut" algorithm could therefore lead to a poor sampling of these regions. Measures should be taken to avoid this.

(B) The "shuffle-and-cut" algorithm, even if biased to ensure proper representativity of both sets, inherently assumes that data points are sufficiently independent from each other, under the condition that at least one input variable assumes a different value. For this application, however, the input variables describing the chemical composition of a given steel are likely to operate in synergy and to have a dominant influence on the corresponding ANN output, $\Delta\sigma_Y$. It might thus be preferable to group the data points by steels and build training and reference sets operating on the data taken steel by steel, i.e. without separating data points referring to the same steel.

Based on these considerations, we have used two different algorithms to build training and reference sets from a given database:

- *Algorithm "by independent points"*: Here, we ignore the possibility of dominance of the input variables describing the chemical composition of a given steel and simply apply the "shuffle-and-cut" algorithm. However, we do not apply it in a completely blind (random) way. Instead, we impose the condition that all data points belonging to the sparsely populated regions of the input/output space (Fig. 2), should be equally distributed within the two sets.
- *Algorithm "by steel"*: The data points are grouped depending on the steel they belong to and these groups are then assigned to one or the other set, trying to share equally the distributions of chemical compositions. This is not straightforward because of the nature of the databases.

The former algorithm has the advantage of being simpler and of creating training and reference sets in which all input variables (and corresponding output values) are more equally distributed. The latter algorithm is of more convoluted application, but takes
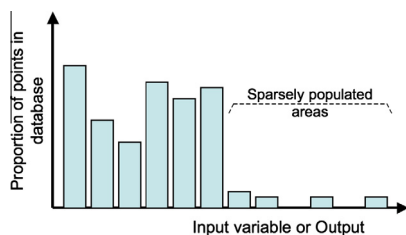
into account the fact that the examples used as reference for the ANN are steels and not independent concentrations of chemical elements.

2.2. Training algorithms

In this section, we describe the problem of training the ANN to reproduce as closely as possible the output data ($\Delta\sigma_Y$ in the present application), as a function of the input variables listed in the introduction.

We use the Levenberg–Marquardt (LM) training algorithm [13,14], *regularized* by *early stopping* [4]. The training set is used to update the synaptic weights in an iterative way, whereas the reference set is used to decide when training should be stopped, by assessing the actual ANN extrapolation skills on "never-seen" cases. In addition to the classical training scheme based on early stopping, we also considered the possibility of applying *node decay*, under a Bayesian training scheme [4], in order to compare the performance of the correspondingly obtained ANNs. The complete objective function *f*, to be minimized on the training set, is:

$$f = \beta \sum_{t=1}^{T} (d_t - O_t)^2 + \alpha \sum_{i=1}^{W} w_i^2 \qquad (2)$$

Here, d_t is the desired ANN output for the training example number t out of T training examples and O_t is the corresponding network prediction. w_i is synapse i out of W. α and β are the *Bayesian hyperparameters*. The second term in Eq. (2) allows node decay to be introduced: it encourages the network to develop small value synapses connections, so as to yield the simplest possible regression. If this term is kept, the application of early stopping becomes theoretically unnecessary. In a classical LM training, with early stopping, we impose $\alpha = 0$ and $\beta = 1$, to turn off node decay. On the contrary, in a Bayesian training, α and β are not imposed, but are iteratively fitted. In this case, a variance on the ANN outputs can be theoretically calculated, to be later used as error bar:

$$\sigma^2 = \frac{1}{\beta} + g^T \cdot A^{-1} \cdot g \qquad (3)$$

Here A is the matrix of the second order derivatives of the function f (see Eq. (2)) with respect to the synapses and g is the vector of the first order derivatives of the ANN output with respect to the synapses.

The reason for comparing the two training schemes is that the Bayesian scheme is expected to create a less complex and therefore more general ANN, presumably more suitable for extrapolation purposes. On the other hand, Bayesian trained networks generally commit larger errors on the reference set compared to networks trained with a classical algorithm, as a consequence of the right-hand side term of Eq. (2).

The input variables are linearly normalized between −0.150 and +0.150 as is common practice in ANN training. Flux and fluences, however, are linearly normalized on a logarithmic scale (i.e. we take their logarithms as input variables and normalize them between −0.150 and +0.150 afterwards), because they vary by several orders of magnitude in the range covered by the database used in this work. This normalization in log-space homogenizes well the distribution of flux and fluences in the database. The output is linearly normalized between −1 and +1, because the output of the ANN is bound to this range, due to the hyperbolic tangent activation function of the output node. The synaptic weights are randomly initialized between −1 and +1. Fig. 3 illustrates how the average error committed on the training and reference sets, after training, evolves with the number H of hidden nodes. The error on the training set always decreases with the addition of new nodes, because: (1) the function f is optimized

Fig. 2. Typical distribution of values for a given input variable in databases of irradiated steels. The variable in question can be for example the content of a certain chemical element or the irradiation temperature; the only output of interest for the RADAMO database is $\Delta\sigma_Y$.
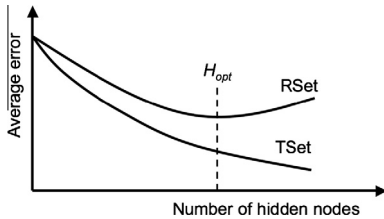
Fig. 3. Illustration of the usual dependence on the number of hidden nodes of the average error committed by the ANN, on the training set (TSet) and reference set (RSet). The dashed line shows the optimal number H_{opt} to retain, for the sake of generality.
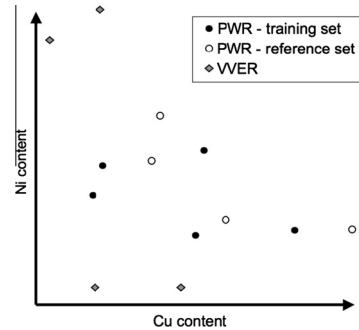

Fig. 4. Distribution of points in the Cu–Ni content space for the nine PWR steels and the four VVER steels contained in the RADAMO database. The PWR points are separated in training and reference sets, defined "by steels".

on this set; (2) the ANN output derived from Eq. (2) is so general that the addition of new degrees of freedom helps improving the accuracy of the prediction. On the contrary, the error committed on the reference set stops decreasing after a certain optimal number H_{opt} of hidden nodes, because the ANN becomes more specialized for the list of examples used for training.

2.3. Use of network committees

A properly trained ANN is expected to provide good predictions for any new set of input variables, at least in the ranges covered by the training and reference sets. However, in practice no single ANN can be fully trusted because the mapping between input variables and output that is constructed during training is fitted "in a mean square sense". The network may have poorly learned the real effect of some input variables under particular sets of conditions that are insufficiently represented in the database. In other words, one should not expect an ANN to be able to predict physical phenomena that are not sufficiently represented in the database of examples used for training: the predictions made in partitions of the input space that are not properly covered by the training database are very likely to vary significantly from network to network, even if trained on the same database. For this reason, it is generally wiser to make predictions using a *committee* of networks, all trained on the basis of the same training and reference sets, rather than using individual networks. The final prediction of a committee of ANN, given a set of input variables, will be the average $\bar{O}$ of the predictions of the individual networks in the committee. In this way, a variance can be calculated:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} \sigma^{(i)2} + \frac{1}{N} \sum_{i=1}^{N} (O^{(i)} - \bar{O})^2 \qquad (4)$$

Here, N is the number of networks in the committee, $\sigma^{(i)}$ is the standard deviation for the prediction of network i (calculated using Eq. (3)), and $O^{(i)}$ is the prediction of network i. Note that all $\sigma^{(i)}$ are 0 if the networks have been trained with the classical LM algorithm, without Bayesian node decay.

3. The RADAMO database

The RADAMO experimental program [15–18] was conducted at SCK•CEN to generate an experimental database covering a large spectrum of irradiation conditions. RADAMO was specifically oriented to measure irradiation effects on the tensile properties of RPV materials. Pressurized water reactor (PWR) and VVER[1] materials (plates, forgings and welds) with various chemical compositions

were irradiated in the BR2 material test reactor under well controlled conditions at two temperatures, $T = 300\ °C$ and $T = 265\ °C$ in a large neutron fluence range from low ($\Phi < 10^{23}\ n/m^2$) to high ($\Phi > 1.5 \times 10^{24}\ n/m^2$, energy > 1 MeV) and various flux levels ($\varphi = 0.2–8 \times 10^{17}\ n/m^2/s$, energy > 1 MeV).

The database contains 346 entries related to PWR materials (nine different steels) and 63 entries related to VVER materials (four different steels). The input variables are: neutron flux, neutron fluence, irradiation temperature, chemical content with respect to several elements (Cu, Ni, P, Mn, Si, etc.) and also the product form (plate, forging or weld). The output is the variation $\Delta\sigma_Y$ of the yield stress measured by tensile tests at room temperature. Fig. 4 depicts the distribution of the 13 different steels in terms of Cu and Ni contents. We see that VVER materials are significantly different from PWR steels, especially in terms of Ni content. At this stage, in this work, we only use the PWR data to train the ANN. VVER data are kept in a separate database, to be used at a later stage to assess the performance of the trained ANNs, when employed to extrapolate to compositions significantly different from those represented in training and reference sets.

Before defining training and reference sets from the PWR data, we also separate a set corresponding to the highest neutron fluences: set PWR_HF, which contains 20 data points, all with neutron fluence higher than, or equal to 1.5×10^{24} n/m². As in the case of the VVER data, this set will be used at a later stage to test the performance of the trained ANNs, when employed to extrapolate to high fluence.

In conclusion, only the 326 remaining data points were used to build training and reference sets. These, as explained in Section 2.1, where defined in two different ways:

(A) *by independent points*, with 162 data points in the training set (PWR_TA) and 164 in the reference set (PWR_RA);
(B) *by steels*, with 174 data points in the training set (PWR_TB), corresponding to five different steels, and 152 in the reference set (PWR_RB), corresponding to four steels.

4. Results

4.1. Identification of the most influential input variables

In this section we describe the differences between ANN committees trained on different sets of input variables, in order to

---

[1] VVER is the abbreviation used to denote Russian-type light water reactors.

identify which are the dominant ones determining the output, $\Delta\sigma_Y$. Neutron fluence and irradiation temperature are unquestionably important and were therefore always included. The training experiments presented here concern the effect of the following variables: neutron flux, product form, and chemical composition.

For these experiments, we exclusively used the training and reference sets defined with the algorithm "by steel", in order to minimize the risk of specializing the ANN for all nine RPV steels of the database. The classical LM training algorithm is the most appropriate, because it allows more precision to be achieved within the given database, as will be clearly shown later.

In order to evaluate the performance of the ANN, we compared the predicted values with the values in the reference set. Better performance corresponds to better correlation between the two values, assessed using Pearson's product-moment correlation coefficient, R^2, as well as by calculating the average error, $\bar{e}$, defined as:

$$\bar{e} = \frac{1}{R}\sum_{r=1}^{R}|d_r - O_r| \qquad (5)$$

Here d_r is the desired ANN output for the reference example r out of R, and O_r is the actual ANN output for the same input variables.

Extensive studies led us to the following conclusions:

- Product form and neutron flux are apparently variables that do not influence the output $\Delta\sigma_Y$, because, all other input variables being identical, their inclusion or exclusion does not change the ANN quality of prediction. Therefore, we conclude that no significant product form or flux effect is numerically discernable, at least in the range of irradiation conditions and chemical compositions covered by the RADAMO database.
- Amongst the chemical composition input variables, Cu content is unquestionably (and unsurprisingly) the dominant one. Fig. 5 shows that, even by considering only the Cu content as chemical composition variable, the correlation between predicted and reference data is very strong.

- The addition of the content of a second chemical element as input variable improved the correlation. However, the improvement achieved is almost the same independently of the choice of the second element (Ni, Mn and Si).

This last point is illustrated in Fig. 5, which shows ANN predictions obtained when neutron fluence, irradiation temperature and contents of two chemical elements are taken into account. We see first of all that the accuracy of the predictions is, in general, very high, and is, on average, of the same order as the experimental uncertainty. We also see that the difference in the quality of the prediction when considering Ni, Si, or Mn content as second chemical composition variable, is almost indiscernible. Taking any of them into account, in addition to Cu, hardly makes any difference in the final ANN accuracy.

We can suggest several explanations for this interesting result:

- Copper has a clear and distinguishable effect on the output $\Delta\sigma_Y$, which is very easily learned by the ANN. The other elements also have a distinguishable influence, separate from the copper effect, but there is a synergy between them, so that it is hard to isolate the individual effects of Ni, Mn and Si. So, the ANN understands the synergic effect without distinguishing the actual role of each element, possibly because of the limited number of steels compositions in the database.
- The ANN is not learning a general logic, but is in fact constructing a non-physical artifact that minimizes the function f in Eq. (2), by somehow "memorizing" the steels of the database, therefore making predictions that are only sound for them. This is however unlikely, because the algorithm "by steel" to define the training and reference sets is specifically aimed at avoiding this.
- The ANN manages to find a correlation between Ni, Mn and Si contents that is not immediately visible to us, but does exist. Such a correlation may exist and be found because the steels
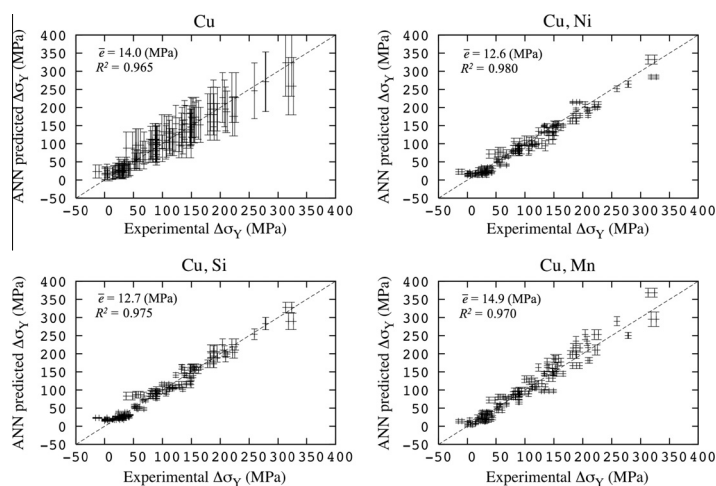
Fig. 5. Comparison of ANN predictions, measured on the reference set by committees of 30 networks (four hidden nodes), obtained when different series of input variables are taken into account. Training and reference sets were defined with the algorithm "by steel" and the ANN's were trained with the classical LM algorithm. In all cases, the input variables are: neutron fluence, irradiation temperature, and the content of the chemical elements indicated on the graphs. Error bars were calculated with Eq. (4), $\bar{e}$ was calculated with Eq. (5), and R^2 is Pearson's product-moment correlation coefficient.

of the database are not sampled in such a way that the chemical content of each element varies independently. On the contrary, the nine steels in the RADAMO database set can uniquely be labeled by just looking at the content of two chemical elements.

A way to differentiate between the importance of Ni, Si and Mn content as variables determining irradiation hardening is to investigate how the ANN is accurate for other steels, whose composition is far from the ranges covered by the training and reference sets. Fig. 6 shows the accuracy of the predictions on the VVER database, obtained by using the same committees of networks as in Fig. 5. We clearly see that: (1) the Cu content variable alone is not enough to ensure accurate extrapolations; (2) in this case Ni content appears the best one to retain, as the second most influential chemical element after Cu, and the one that provides the best extrapolation capabilities.

Finally, our attempt to improve the correlation by taking more than two chemical elements into account failed, because the ANN accuracy was never better than that shown in Figs. 5 and 6. This can either be explained by the fact that a hypothetical synergic effect of Si, Mn and Ni on hardening, plus the effect of Cu, removes the need to take all chemical elements into account, or, more simply, that the database does not contain enough training examples to correctly deal with such a regression problem, if more than four input variables are involved.

To summarize, we have shown in this section that the highest quality ANN predictions can be obtained by taking just four input variables into account: neutron fluence, irradiation temperature, Cu and Ni content. Neutron flux and product form have no significant influence on the RADAMO $\Delta\sigma_Y$ output.

4.2. Comparison between the proposed training schemes

In this section, we retain the four input variables that were selected based on the study reported in the previous section (neutron fluence, irradiation temperature, Cu and Ni contents) and compare the ANN accuracy of predictions after training with either the classical or the Bayesian LM algorithm, as well as defining training and reference sets either "by independent points" or "by steel" (see Section 2.1).

4.2.1. Comparison of the set definition algorithms

In Fig. 7, the performance of ANN committees trained with the classical LM algorithm is shown, defining the training and reference sets either "by independent points" (left side) or "by steel" (right side). In the same figure (lower part), the ANN committee is also tested on the high fluence set, PWR_HF. The biases $\bar{b}$, when shown, were calculated as:

$$\bar{b} = \frac{1}{N} \sum_{i=1}^{N} (O_i - d_i) \tag{6}$$

where d_i is the desired ANN output for the high neutron fluence example i out of N, and O_i is the actual ANN output for the same input variables. We see that the predictions on the reference sets (upper part of the figure) are slightly more accurate when the algorithm "by independent points" is used. The predictions on the high fluence set (lower part) are slightly more biased when the algorithm "by steel" is used. This can be explained by the fact that the ANN, in the first case, is more specialized for the particular steels represented in the database, and therefore manages to perform a more accurate extrapolation on the neutron fluence variable. As an additional illustration, the performance of the same ANN committees is illustrated, in Fig. 8, where predictions are obtained for the steels of the VVER database, i.e. with chemical compositions that are far from those used during training, either in the training or reference sets. We can see that the committee of networks yields slightly better extrapolations when training and reference sets were defined "by steel".

It therefore appears that the use of the algorithm "by independent points" to define the training and reference sets is preferable for the purpose of extrapolation on the neutron fluence variable,
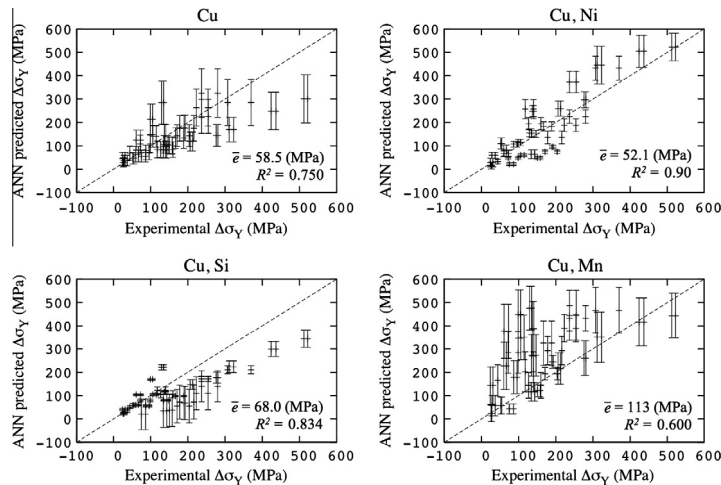

Fig. 6. ANN quality of prediction for the VVER database, using the same committees of networks as in Fig. 5, i.e. trained using the PWR data. (See Fig. 5 caption for the definition of $\bar{e}$ and R^2).

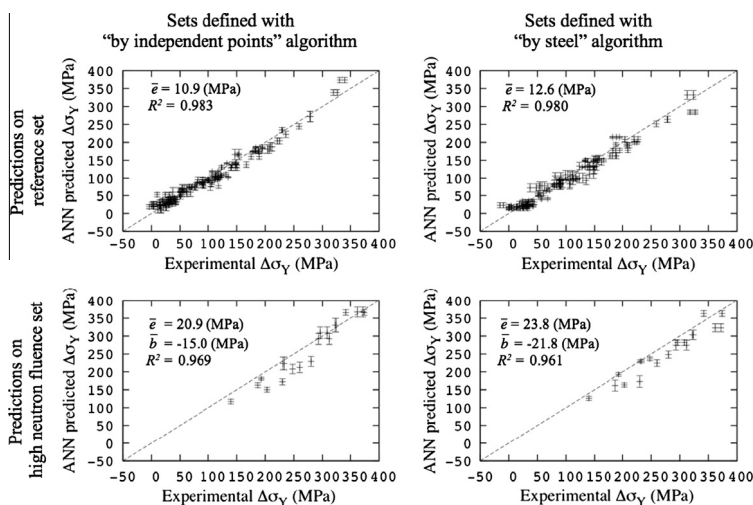*N. Castin et al. / Journal of Nuclear Materials 408 (2011) 30–39*

Fig. 7. Quality of prediction for committees of 30 networks (four hidden nodes), trained with the classical algorithm. Training and reference sets were either defined "by independent points" (left side) or "by steel" (right side). Error bars were calculated with Eq. (4), $\bar{e}$ with Eq. (5), $\bar{b}$ with Eq. (6), and R^2 is Pearson's product-moment correlation coefficient. Upper part: predictions for the reference set; lower part: predictions for the high neutron fluence set.
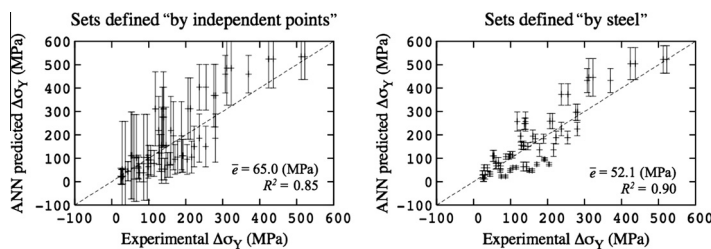

Fig. 8. ANN quality of prediction measured on the VVER database, using the same committees of networks as in Fig. 7, i.e. trained using the PWR data. (See Fig. 6 caption for the definition of $\bar{e}$ and R^2).

considering a steel that is already represented in the database. Otherwise, the use of the algorithm "by steel" is recommended, in order to predict the hardening of steels not included in the reference database.

4.2.2. Comparison between classical and Bayesian training algorithms

Fig. 9 is the equivalent of Fig. 7, for ANN committees trained with the Bayesian LM algorithm, defining training and reference sets either "by independent points" (right side) or "by steel" (left side). In comparison to Fig. 7, we see that the general accuracy of prediction is significantly lower than when classical LM training is used. In particular, the biases of the predictions on the high fluence set are much larger. This can be explained by the introduction of node decay in Eq. (2) (Section 2.2).

The predicted dependence of hardening on fluence for different temperatures, for two steels of the reference database, is shown in

Fig. 10. We see that the Bayesian trained network is accurate for the higher temperature (300 °C), and is in fact closer to experimental data than the other network (trained in a classical way), because the average prediction line touches the bulk of experimental points and the error bands calculated with Eq. (4) encompass all of them. The quality of the predictions for the lower temperature (265 °C), however, is poor compared to the classically trained network. This can be explained by the fact that this temperature is poorly represented in the RADAMO database: only 48 data points, i.e. less than 15% of the database. As the minimization of the function f, in Eq. (2), is solved as a mean-square optimization problem, and as node decay prevents the network from developing a complex structure, the ANN became specialized for the higher temperature, that largely dominates the database.

One possible way to improve the generality of the Bayesian trained ANN, for the irradiation temperature variable, could

Fig. 9. Quality of prediction for committees of 30 networks (four hidden nodes), trained with the Bayesian algorithm. Training and reference sets were either defined "by independent points" (left side) or "by steel" (right side). Error bars were calculated with Eq. (4), $\bar{e}$ was calculated with Eq. (5), $\bar{b}$ was calculated with Eq. (6), and R^2 is Pearson's product-moment correlation coefficient. Top: predictions for the reference set; bottom: predictions for the high neutron fluence set.

Fig. 10. Evolution with neutron fluence of the ANN-predicted $\Delta\sigma_Y$, versus experimental measurements, for two steels. Predictions are obtained using a committee of 30 networks (four hidden nodes). The training and reference sets were generated "on independent points". Dashed lines show the error bands calculated with Eq. (4). Left side: classical training; right side: Bayesian training.
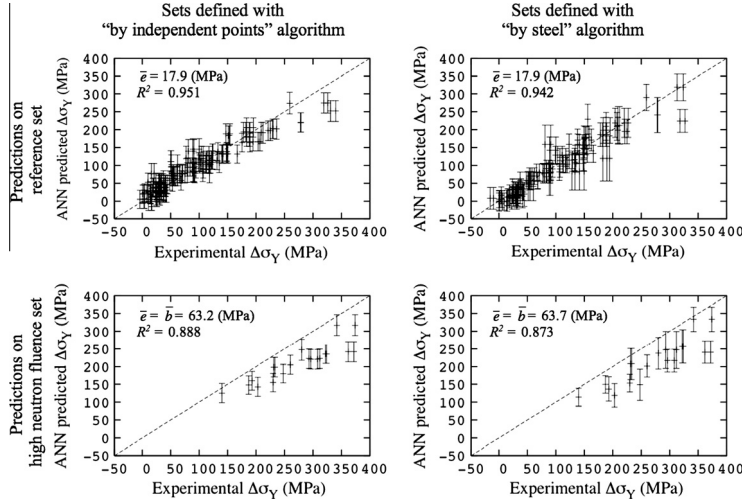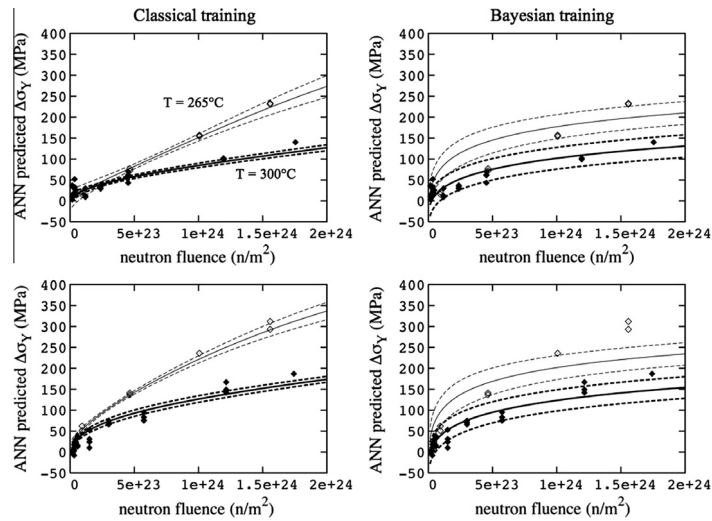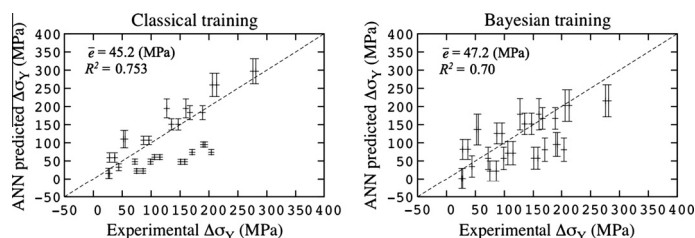
Fig. 11. Quality of prediction for committees of 30 networks (four hidden nodes), trained with the PWR data (training and reference sets were defined with algorithm B), for the VVER data ($T = 300\,°C$). The training algorithm was either classical or Bayesian LM. Error bars were calculated with Eq. (4), $\bar{e}$ was calculated with Eq. (5), $\bar{b}$ was calculated with Eq. (6), and R^2 is Pearson's product-moment correlation coefficient.

therefore be to homogenize the database, by removing several points corresponding to 300 °C, so that each temperature (265 °C and 300 °C) represents about 50% of the database. We thus tried re-training with only 108 data points of the database: nine steels times two temperatures times six data points per steel and per temperature. Unfortunately, even after this homogenization the Bayesian trained ANN accuracy is not improved for both temperatures. This could be the consequence of the limited amount of steels in the database, which makes the application of Bayesian training inappropriate: node decay is apparently too strong a constraint that inhibits the ANN quality of prediction, despite its theoretically higher extrapolation skills. In other words, Bayesian training seems to express its full potential only if a sufficiently homogeneous *and large* database is available for training.

To conclude this study, Fig. 11 compares the accuracy of the predictions provided by the classically trained or by the Bayesian trained ANN committee on the VVER database. Only data points corresponding to 300 °C are shown, since we have already observed that the Bayesian network is inappropriate for the lower temperature. We can see that the accuracy, estimated on the basis of mean error and correlation, is very similar in both cases, although slightly lower for the Bayesian trained ANN committee. We may summarize that, if trained on a database larger than RADAMO, the accuracy achieved with Bayesian training would be higher than with classical training. However, in this specific case, very similar results are obtained in both cases when extrapolating to the different chemical compositions found in VVER steels.

5. Concluding remarks

In this paper, we have shown that artificial neural networks can be used to accurately predict neutron irradiation induced hardening of reactor pressure vessel steels, by taking into account the neutron fluence, irradiation temperature, Cu and Ni contents. The advantage of this numerical regression technique is that no hypothesis about how exactly these input variables influence hardening needs to be explicitly formulated, as the artificial intelligence approach itself takes care of finding non-evident relationships between the input variables and the output. This allowed us to point out the apparently negligible influence of the neutron flux, as well as of the product form, as input variables determining hardening, at least within the range of chemical composition and irradiation temperatures covered by the RADAMO database. At the same time, we have shown that the determination of the most influential chemical elements for hardening based on purely empirical considerations is not straightforward.

We have compared two training algorithms, as well as two methods for defining training and reference sets from the available

database. We concluded that training the artificial neural network with an early stopping regularized algorithm, without the application of node decay, and defining training and reference sets "by independent points", can be recommended in order to train networks that can be accurately extrapolated from an existing database to high neutron fluences.

However, the accuracy in the extrapolation to different chemical compositions (VVER) is not fully satisfactory. In particular, the application of Bayesian node decay as a way to construct better networks provided no significant improvement, probably because of the limited amount of examples in the RADAMO database. In future work, we will further investigate the possibility of extrapolating to different chemical compositions, by extending the database and/or combining several databases of irradiated steels. The application to surveillance data will be our ultimate objective.

References

[1] R. Gérard, Survey of National Regulatory Requirements, AMES Report No. 4, EUR 16305 EN, Brussels, June 1995.
[2] P. Petrequin, A Review of Formulas for Predicting Irradiation Embrittlement of Reactors Vessel Materials, AMES Report No. 6, November 1996.
[3] IAEA Nuclear Energy Series, Integrity of reactor pressure vessels in nuclear power plants: assessment of irradiation embrittlement effects in reactor pressure vessel steels, IAEA Nuclear Energy Series, No. NP-T-3.11, 2009.
[4] C.M. Bishop, Neural Networks for Pattern Recognition, Clarendon press, Oxford, 1995.
[5] F.F. Soulié, P. Gallinari, Industrial Applications of Neural Networks, World Scientific Publishing Co., 1998.
[6] J.R. Rabunal, J. Dorado, Artificial Neural Networks in Real-Life Applications, Idea Group Publishing, 2006.
[7] R. Kemp, G.A. Cottrell, H.K.D.H. Bhadeshia, G.R. Odette, T. Yamamoto, H. Kishimoto, J. Nucl. Mat. 348 (2006) 311–328.
[8] C.G. Windsor, G. Cottrell, R. Kemp, Modell. Simul. Mater. Sci. Eng. 16 (2008) 025005.
[9] S. Haykin, Neural Networks: A Comprehensive Foundation, MacMillian, New York, 1994.
[10] K. Hornik, M. Stinchcombe, H. White, Neural Networks 2 (1989) 359–366.
[11] A.-S. Bogaerts, R. Gérard, R. Chaouadi, Comparison of embrittlement trend curves to high fluence RPV surveillance results, in: International Symposium on Contribution of Materials Investigations to Improve the Safety and Performance of LWRs (Fontevraud 7), Avignon, France, 2010 (26–30 September).
[12] R. Chaouadi, R. Gérard, A.-S. Bogaerts, Assessment of the French and US embrittlement trend curves applied to RPV materials irradiated in the BR2 reactor, in: International Symposium on Contribution of Materials

174

Investigations to Improve the Safety and Performance of LWRs (Fontevraud 7), Avignon, France, 2010 (26–30 September).
[13] K. Levenberg, Q. Appl. Math. 2 (1944) 164.
[14] D. Marquardt, SIAM J. Appl. Math. 11 (1963) 431.
[15] R. Chaouadi, R. Gérard, J. Nucl. Mat. 345 (2005) 65–74.
[16] R. Chaouadi, An engineering radiation hardening model for RPV materials, Belgian Nuclear Research Centre (SCK•CEN) Report, R-4235, September, 2005.

[17] R. Chaouadi, RADAMO – An experimental databank for investigating irradiation strengthening of RPV materials, Belgian Nuclear Research Centre (SCK•CEN) Report, R-3858, April, 2004.
[18] R. Chaouadi, Irradiation-induced hardening of eight RPV steels and welds – RADAMO irradiation, Belgian Nuclear Research Centre (SCK•CEN) Report, R-3676, December, 2002.

H Paper VII

Modeling the first stages of Cu precipitation in α-Fe using a hybrid atomistic kinetic Monte Carlo approach.

N. Castin[1,2,*], M. I. Pascuet[3] and L. Malerba[1]

[1] *Studiecentrum voor Kernenergie – Centre d'Etudes de l'énergie Nucléaire (SCK•CEN), NMS unit, Boeretang 200, B2400, Mol, Belgium.*
[2] *Université Libre de Bruxelles (ULB), Physique des Solides Irradiés et des Nanostructures (PSIN), boulevard du Triomphe CP234, 1050 Brussels, Belgium.*
[3] *Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Av. Rivadavia 1917, C1033AAJ Buenos Aires, Argentina, pascuet@cnea.gov.ar.*
[*] Corresponding author : ncastin@sckcen.be

Abstract

We simulate the coherent stage of Cu precipitation in α-Fe with an atomistic kinetic Monte Carlo (AKMC) model. The vacancy migration energy as a function of the local chemical environment is provided on-the-fly by a neural network, trained with high precision on values calculated with the nudged elastic band method, using a suitable interatomic potential. To speed up the simulation, however, we modify the standard AKMC algorithm by treating large Cu clusters as objects, similarly to object kinetic Monte Carlo (OKMC) approaches. Seamless matching between the fully-atomistic and the coarse-grained approach is achieved again by using a neural network, that provides all stability and mobility parameters for large Cu clusters, after training on atomistically-informed results. The resulting hybrid algorithm allows long thermal annealing experiments to be simulated, within a reasonable CPU time. The results obtained are in very good agreement with several series of experimental data available from the literature, spanning over different conditions of temperature and alloy composition. We deduce from these results and relevant parametric studies, that the mobility of Cu clusters containing one vacancy plays a central role in the precipitation mechanism.

1 Introduction

Copper was inadvertently included as alloying element in nuclear reactor pressure vessel steels of second generation. Later, the formation of Cu-rich precipitates revealed itself as one of the major causes of degradation of mechanical properties in ferritic steels under neutron irradiation [1] (mainly hardening and embrittlement), which is nowadays an important factor limiting the lifetime of nuclear power plants. For this reason, Cu precipitation in α-Fe has been intensively studied during the last decades, both experimentally [2-7], and using theoretical models [8-12] or computer simulations [7, 13-27]. In particular, several attempts have been made in the last couple of decades to simulate Cu precipitation in α-Fe using

atomistic kinetic Monte Carlo (AKMC) techniques [7, 13-25]. In these simulations, a single vacancy is introduced in a volume filled with atoms of different chemical species, located on perfect crystal lattice positions. The vacancy acts as vehicle for atomic species re-distribution, via thermally activated diffusion jumps, i.e. exchanges of position between the vacancy and an atom nearby. The physics of the model is contained in the activation energies of the diffusion jumps (migration energies of the vacancy), which vary significantly depending on the local chemical, or atomic, environment.

Generally AKMC models are subsumed as rigid lattice models, i.e. crystallographic changes cannot be described. This seems to rule out their use to describe precipitation of face-centred-cubic (fcc) Cu in body-centered-cubic (bcc) α-Fe. However, Cu precipitates in α-Fe are experimentally found to be coherent with the bcc-Fe matrix, up to a diameter between 4 and 5 nm [28,29]. Growing further, they first take intermediate crystallographic structures (first 9R, then 3R), and finally become fcc when they reach diameters above 12 nm [6]. This finding was also confirmed using molecular dynamics simulations, which moreover suggested that the stability of Cu precipitates is enhanced by the presence of vacancies inside [27]. Thus, it makes sense to simulate by AKMC the coherent stage of Cu precipitation in Fe, as long as Cu precipitates remain smaller than ~5 nm.

AKMC models differ from each other essentially by the methodology employed to calculate the vacancy migration energies. In most models, they are calculated without allowing for atomic relaxation, using an interatomic potential or fitted pair interaction energies, for example in a broken-bonds formalism, within a range of interaction of first or second nearest neighbour distance. In some cases, the pair interaction energies are fitted to data calculated with first principles methods. For a critical review of most existing approaches, the interested reader is referred to Vincent *et al.* [7]. The conclusion by Vincent *et al.* is, however, that none of the models reviewed can consistently predict Cu precipitation in Fe: depending on the parameterization, either the average precipitate radius, or the precipitate density is correctly predicted, but hardly ever both. In particular, most models tend to overestimate the density.

Soisson and Fu [18] achieved a more satisfactory and consistent prediction. In their AKMC model, the vacancy migration energy was also calculated using a broken-bonds formalism, fitted to energy data calculated by density functional theory methods. However, they were able to predict accurately the evolution with annealing time of both the average precipitate

radius and the precipitate density. The main reason for their success is probably that their model, by incorporating a very strong interaction between vacancies and Cu clusters, predicted that Cu clusters of all sizes were mobile, as a consequence of complex series of vacancy hops at their surface or near the surface. They also observed direct coalescence of clusters, as the result of migration. Their results therefore suggested that the mobility of Cu clusters, as a consequence of strong interactions with vacancies, can play a non-negligible role in the kinetics of Cu precipitation in Fe. Mobility of clusters was also observed in previous AKMC simulations, though limited to the smallest clusters [13, 15]. The diffusion of Cu clusters, on the other hand, had never been considered as mechanism in classical diffusion theory models for Cu precipitation [9-12].

The main limitation of Soisson and Fu's model was a "collateral effect" of the strong binding between Cu clusters and vacancy: the latter remained trapped in the bulk of these clusters for a very large fraction of the simulation time, thereby increasing enormously the CPU cost and therefore limiting drastically the reach of the simulation, which had to be stopped at a very early stage of the coherent precipitation.

In this work, we simulate Cu precipitation in α-Fe with an AKMC computer simulation approach already described in ample detail elsewhere [30]. To summarize briefly, we calculate the vacancy migration energies with the nudged elastic band (NEB) method [31,32], using an interatomic potential, thereby fully allowing for effects of long-range chemical interaction and static relaxation. However, to allow fast estimation of the migration energies during the simulation, the NEB values are in fact provided by a properly trained artificial neural network (ANN) [33]. Key for the physical reliability of the model is the quality of the interatomic potential. The Fe-Cu potential we use here was developed in [34] specifically to fulfill two major objectives: (1) be consistent with thermodynamics, by providing a correct prediction of the experimental Fe-Cu phase diagram; (2) provide an accurate description of the interaction between point-defects in α-Fe and Cu atoms. The potential has been proven to be able to predict the correct final equilibrium for a thermal annealing experiment [23,34], although its ability to fully predict the kinetic path from a random solid solution has not been evaluated yet. This is done in the present work.

In addition, we designed our model to allow *explicitly* for the mobility of Cu clusters of all sizes, while reducing drastically the CPU cost of the simulation, as compared to a full AKMC

180

3

simulation. This was achieved by combining the AKMC approach with a coarse-grained approach, of object kinetic Monte Carlo (OKMC) type [26]. Clusters of Cu atoms *above a certain size* are considered as objects, for which migration and dissociation events are defined, based on specific, size-dependent and thermally-activated frequencies. Seamless matching between the fully-atomistic model used to describe small Cu clusters and the coarser-grain model used to describe larger Cu clusters is guaranteed by calculating the diffusion coefficients and emission probabilities for the object-like clusters based on specific, full AKMC simulations, on which another ANN has been trained. We show in what follows that this "hybrid" AKMC approach proves both computationally efficient and physically very accurate, thanks also to the high quality of the interatomic potential, which is here exploited in the most complete way possible.

The paper is organized as follows. In section 2, we summarize the fundamentals of our ANN-based AKMC algorithm, reporting in detail on the modifications introduced to "hybridize" it with an OKMC approach. In section 3, the hybrid AKMC is used to simulate several thermal annealing experiments in Fe-Cu, for different Cu contents and at different temperatures. Finally, we analyze in section 4 the mechanism of Cu precipitation that stems from our simulations, emphasizing in particular the influence of the diffusion of Cu clusters and precipitates.

2 Hybrid atomistic kinetic Monte Carlo approach

As anticipated in the introduction, in AKMC models atoms are located on the positions corresponding to the crystallographic structure of interest, generally on a rigid lattice. The evolution of the system is driven by the diffusion jumps of vacancies. The jump to occur is each time selected stochastically, according to the jump frequencies:

$$\Gamma = \Gamma_0 \cdot \exp\left({-E_m}/{k_B \cdot T} \right) \tag{1}$$

Here Γ_0 is the attempt frequency, E_m is the migration energy calculated at zero temperature, k_B is Boltzmann's constant and T is the absolute temperature. The Monte Carlo time increment is then obtained using the standard residence time algorithm [35-37].

The simulation volume is a cubic and periodic bcc monocrystal, with Fe as matrix. A number of Cu atoms are introduced, according to the alloy composition, as well as one vacant site. An ANN was trained to predict the vacancy migration energy, as calculated with NEB using the potential from [34], given as input a description of the vacancy local chemical environment (i.e. how Cu atoms are distributed around it). The methodology we followed to train the ANN is described in Ref. [30]. Fig. 1 shows that the predictions are very accurate, therefore the ANN can be considered as an adequate substitute of the computation on-the-fly of energy barriers with NEB. The attempt frequency in Eq. (1) is considered constant and taken to be of the order of Debye's frequency: $\Gamma_0 = 6 \cdot 10^{12}$ (s^{-1}).

Due to the negligible solubility of Fe in Cu, the clusters formed during the simulation are from the beginning completely pure and compact, similarly to the findings by other authors in previous AKMC simulations [7,15,18]. Also consistently with previous work, in a full AKMC simulation with this method, the vacancy is strongly attracted by Cu clusters, and remains trapped inside them for a very large fraction of the AKMC events. The simulation is, therefore, significantly slowed down [7, 18, 23]. To speed-up the calculation, clusters of Cu atoms (henceforth denoted has Cu_N, N being the number of atoms in the cluster) above a certain size ($N \geq N_{min}$) are defined as objects, using the approximations described in what follows. Clusters are defined by groups of atoms that are linked by 1nn or 2nn bonds.

When the vacancy approaches a Cu_N cluster object, as depicted in Fig. 2, the full AKMC algorithm is still applied: possible events are only the migration of the vacancy to a first nearest neighbor position, and the corresponding migration energies are calculated using the ANN shown in Fig. 1. We therefore do not introduce any coarse-grain approximation for the migration path followed by the vacancy in the vicinity of a Cu_N cluster, and – most importantly – for the mechanisms of dragging individual Cu atoms, or small Cu clusters, in the direction of the large cluster.

182

When the vacancy enters in contact with a cluster, many events that would take place in a full AKMC simulation are ignored and replaced by coarse-grain approximations, as shown in Fig. 2. In particular, the most probable outcome of the capture of a vacancy by the cluster is the subsequent inclusion in the cluster of all Cu atoms that are in solution in the matrix, but sufficiently close to the interface with the precipitate to be dragged inside by the vacancy hopping at this interface. To allow for this atomic-level process, we define a spherical absorption radius (radius of the cluster augmented by the 2nn distance) and make the assumption that all Cu atoms within this radius are immediately absorbed by the cluster.

(Fig. 3 about here)

After the vacancy (V) has been absorbed by a Cu_N cluster object, the latter becomes a VCu_N object, and new events are defined in replacement of the vacancy migration in the regular AKMC algorithm:

- Dissociation of the cluster, with a frequency denoted as $\Gamma_N^{(diss)}$, as depicted in Fig. 3. Two dissociation mechanisms are possible: (i) emission of the vacancy from the cluster, or (ii) emission of a VCu_1 pair, as in OKMC simulations [26]. A single frequency for the dissociation is assigned, and the emission of a VCu_1 pair occurs with a probability p_N.

- Migration of the cluster, with a frequency denoted as $\Gamma_N^{(mig)}$. For convenience, we only consider jumps of the central atom of the cluster (dragging the whole cluster) to any of the eight possible 1nn lattice sites (in bcc). After the migration is completed, Cu atoms within the absorption range are added to the cluster.

- Coalescence of clusters, when, after a migration event is chosen, the absorption range of the VCu_N cluster overlaps with the absorption range of another Cu_M cluster.

To summarize, the hybrid AKMC algorithm is a compromise between AKMC and OKMC: the description at the atomic level of the system is retained as long as the vacancy is not trapped in Cu_N clusters. In the latter case, VCu_N clusters are defined as objects, and treated in a similar way as in OKMC methods. Clearly, the application of this algorithm requires the pre-definition of values for $\Gamma_N^{(diss)}$, $\Gamma_N^{(mig)}$ and p_N as functions of the size, N, and temperature, T. These values must guarantee seamless matching with the atomistic description. The procedure to obtain this parameterization is described in the next section.

183

6

2.2 Parameterization

First of all, an appropriate threshold value, N_{min}, above which clusters are considered as objects must be chosen. A sensible choice is e.g. $N_{min} = 15$, because it corresponds, in bcc, to a central atom surrounded by other atoms filling completely its first and second shells of close neighbors (so it is very stable), and it is larger than the critical size for nucleation of Cu clusters in our AKMC. It is also useful to define the maximum allowed size for Cu_N clusters, N_{max}, to remain in a framework of coherent precipitation: $R = 2.5$ nm [28,29] corresponds to $N_{max} \approx 6000$.

Migration and dissociation of Cu clusters are complex processes that are not easily described with simple formalisms, because they are the consequence of a succession of many vacancy jumps at their surface. We therefore use again a numerical approach: cluster migration and dissociation frequencies are estimated with series of independent full AKMC simulations, using the ANN of Fig. 1. The VCu_N cluster is introduced alone in an otherwise pure Fe matrix. The full AKMC algorithm is applied until the vacancy, or a VCu_1 pair, is emitted from the cluster. Repeating this simulation a large number of times, enough statistics can be collected to calculate the cluster diffusion coefficient D_N, and the average lifetime τ_N, following the procedure described in Ref. [38]. The dissociation frequency is the inverse of the lifetime:

$$\Gamma_N^{(diss)} = 1 / \tau_N \tag{2}$$

Neglecting correlations, the migration frequency can be derived from the diffusion coefficient using the relationship:

$$\Gamma_N^{(mig)} = \frac{6 \cdot D_N}{\Delta^2} \cdot \frac{1}{8} \tag{3}$$

where Δ^2 is the square of the jump distance, i.e. of the 1nn distance, equal to $\frac{3}{4}a_0^2$ (a_0 is the lattice parameter, 2.86 Å for dilute Fe-based alloys). The factor 1/8 in is introduced to account for the eight possible destinations of migration.

184

D_N and τ_N are calculated at several very high temperatures (up to 4000 K) and linearly extrapolated to the temperatures of interest via Arrhenius plots: $\ln(D_N)$ or $\ln(\tau_N)$ versus the reciprocal temperature $1/k_BT$. As a matter of fact, high temperatures make the lifetime, and therefore the simulation time, shorter, thereby enabling the collection of statistically relevant quantities of data points, also for very large cluster sizes. A few sets of data points are represented as Arrhenius plots in Fig. 4 and Fig. 5. Fig. 6 shows an example of diffusion coefficient D_N for T = 773K versus size. We see that for $N > 1000$, the value is not monotonously decreasing with N: these oscillations are entirely attributed to extrapolation errors from high temperature, as indicated by the error bars. In order to fit a function providing diffusion coefficients and lifetime of clusters as functions of both cluster size and temperature, we sought for smooth mathematical expressions to these relationships:

(Fig. 4 about here)

(Fig. 5 about here)

(Fig. 6 about here)

$$\ln\left(D_N(T)\right) = f_1\left(N, 1/k_BT\right) \tag{4}$$

$$\ln\left(\tau_N(T)\right) = f_2\left(N, 1/k_BT\right) \tag{5}$$

These functions were constructed by fitting to the data points shown in Fig. 4 and Fig. 5, using ANNs again. In the case of D_N at 773K, Fig. 6 shows both reference data points and the regression obtained using the ANN.

(Fig. 7 about here)

Finally, the probability p_N that the VCu_N cluster dissociates by the emission of a VCu_1 pair is simply calculated as the ratio between the number of times the clusters dissolved by this mechanism, and the number of times the vacancy alone was emitted. To obtain enough statistics, a number of simulations as large as 100 000 was necessary. These calculations could therefore only be performed at high temperature ($T \geq 1125$ K). The values we calculated are shown in Fig 7. Similarly to the above case, we designed a general fitting function based on ANN to extrapolate p_N to lower temperatures and to any size N:

$$p_N(T) = f_3(N, 1/k_B T) \qquad (6)$$

Fig. 7 shows the regression for 2000 K and its extrapolation to 773 K.

2.2 Time rescaling

Simulations of thermal annealing experiments with AKMC are conducted with the introduction of one vacancy in the simulation volume. The equilibrium vacancy concentration in real materials is, however, much smaller, by several orders of magnitude. For this reason, the Monte Carlo time t_{MC} must be rescaled before a comparison with experimental results is possible. The vacancy concentration in the real material can be calculated from the vacancy enthalpy of formation, h_v^f:

$$C_v^{(real)} = A \cdot \exp\left(-h_v^f \Big/ k_B \cdot T\right) \qquad (7)$$

Here, A is a constant number that depends on the alloy composition ($A \approx 1$ for dilute Fe-based alloys). The enthalpy of formation h_v^f also depends on the alloy composition, and is not accurately known, either experimentally, or by first principle calculations. In this work, we use the value predicted by the interatomic potential used for pure Fe, namely, $h_v^f = 1.7$ eV. The vacancy concentration in the simulation box is given by the ratio between the number of vacancies (1 in our case) and the number of atoms N_{at} in the volume. Le Bouar and Soisson proposed in Ref. [15] to correct this concentration, while rescaling time, to take into account the evolution of the local vacancy formation energies during the process of Cu precipitation. The vacancy concentration in the box is then calculated as:

$$C_v^{(MC)} = \frac{f_V}{N_{at} \cdot X_{Fe}} \qquad (8)$$

Here, X_{Fe} is the proportion of Fe atoms in the box ($X_{Fe} \approx 1$), and f_V is the fraction of time spent by the vacancy in a pure Fe environment (up to the 2nn distance).

Not all authors obtained satisfactory results using this procedure for rescaling time [7]. This can be attributed to the uncertainty on the exact value of the formation enthalpy h_v^f, the

186

consequences of which are exacerbated by its appearing in an exponential function in Eq. (7). Other uncertainties can also have significant effect, such as for example the value of the vacancy jump attempt frequency Γ_0 in Eq. (1), or the concentration-dependent constant A in Eq. (7). For these reasons, other authors choose the value of h_v^f in order to obtain the best possible fit between the results of the simulation and the experimental data [7]. This approach can be considered equivalent to applying a global correction factor f_g, and the real time is thus calculated as:

$$t_{real} = t_{MC} \cdot \frac{C_v^{(MC)}}{C_v^{(real)}} \cdot f_g \qquad (9)$$

If the model is globally consistent and correct, given a Cu content in the alloy, a single value of f_g can be found that leads to a complete superposition of the predictions of the model with all experimental data. Ideally, the value of f_g should be as close as possible to 1. We fitted values that vary between 0.1 and 1.30 (see next section). This limited variation no doubt accounts for all uncertainties regarding the concentration dependence of the parameters entering the above equations. It is important to emphasize, however, that for a given Cu content in the alloy, a unique factor is fitted under the game rule that this unique value must allow reproduction of all experimental data at all temperatures for that alloy (average precipitate size, but also density).

3 Simulation of thermal annealing experiments

In this section, we report on the use of the hybrid AKMC approach described in section 2 to simulate several thermal annealing experiments in Fe-Cu, as summarized in Tab. 1. The results of our simulations are summarized in Tab. 2 and Fig. 8.

(Tab. 1 about here)
(Tab. 2 about here)
(Fig. 8 about here)

In order to choose the most appropriate simulation volume, as a trade-off between reasonable CPU time and statistical accuracy of the prediction, we consider that the peak of precipitate density is, at 773K with 1.34%Cu, of the order of 4×10^{24} m^{-3}. The density then decreases to

187

10^{23} m^{-3} at the end of coherent precipitation stage. A sufficient size for the AKMC simulation box is thus 64x64x64 units cells (524 288 atoms), similarly to other authors' choices [7, 18]. The peak of density, if correctly predicted by the model, would then be reached by a number of 24 precipitates in the box, whereas if only one large precipitate remains at the end of the coherent stage the density will be correctly sampled. Nonetheless, to obtain more statistics some simulations were also conducted in a bigger 128x128x128 unit cells box (4 194 304 atoms). Another strategy adopted to optimize the CPU time has been to start the simulation with a smaller box and continue the same simulation in a box eight times bigger when the densities become too low for the small box to provide enough statistics: this was achieved by duplicating the small box before restarting the simulation, an operation consistent with the use of periodic boundary conditions.

Fig. 8 shows at a glance that the predictions of the hybrid AKMC model are in very good agreement with experimental data. For example, the increase of the average precipitate radius versus time in Fe-1.34at%Cu is very closely reproduced at all three temperatures investigated (773K, 873K and 973K). In the case of the experiment at 773K the measured evolution of the density of precipitates is also provided and the model very nicely predicts nucleation (density increases), growth (density reaches a peak and remains temporarily constant while the radius keeps increasing) and coarsening (density decreases while the radius keeps increasing, because large precipitates grow at the expense of smaller ones). The curves obtained in the simulation box with sides of 64 lattice parameters are jerky because, especially when the coarsening stage is reached, only a few precipitates remain in the box and the disappearance of a small one to make a big one produces significant oscillations in the overall density. In particular, step-like increases/decreases are observed: this is a clear indication of the fact that the mechanism leading to the density decrease and radius increase is the coalescence of two mobile precipitates. Simulations conducted in the larger box (side of 128 lattice parameters) allow the jerks to be *damped*, thanks to better statistics (larger number of precipitates in the box). In all cases, the simulations finished with one single cluster in the box, of varying size, consistently with the increasing solubility limit with temperature. Finally, time was rescaled using the same factor f_g in Eq. (9) for all temperatures: $f_g = 0.60$.

In the simulation of the annealing of a Fe-1.1at%Cu alloy at 823K we observe that the first two experimental points are not predicted by our model. The good agreement with experiments obtained in all other cases, however, gives us sufficient confidence to believe

that those two experimental points are probably affected by large uncertainty, possibly as a consequence of the limited resolution of the experimental technique used, i.e. small angle neutron scattering, which is not sensitive to precipitates below 1 nm in diameter and therefore can be supposed to have overestimated the average size at the early stage of the precipitation, especially because in the experiment there was no support from any other complementary technique. This is confirmed by the fact that the third experimental point, still within the limit of coherent precipitation, is correctly reached by the model, and that a visual extrapolation of the curve will lead to reach the fourth point as well, even though this lies well into the regime where crystallographic transformation must have started, i.e. strictly speaking outside the range of validity of the model. Our model, as most experimental data, is roughly consistent with a dependence of the radius on a ½ power of time during growth that decreases to a dependence on a $^1/_3$ power of time during coarsening, as should be expected [6,7] (in logarithmic scale this is a roughly linear dependence, though with gradual change of slope), while no regression interpolating the four experimental points from [4] will respect such a law. Finally, time was rescaled using a different factor compared to the Fe-1.34%Cu alloy: f_g = 1.3 in Eq. (9). This can be explained by the fact that comparison is here made with a single set of experimental data. The difference in the optimal f_g factor can thus be attributed to the usual scatter between sets of experimental data, as one can see in the top panel in Fig. 8.

Finally, the annealing of Fe-0.6at%Cu at 773K leads to a precipitate density significantly smaller than in the above case, at all stages, consistently with the halved solute concentration: the simulation could only be meaningfully performed in a 128x128x128 unit cells box with ~4 million atoms. These simulations were particularly demanding in terms of CPU time, also because the Cu concentration is lower and the acting thermodynamic force correspondingly weaker. Nonetheless, good agreement with experimental data is achieved. The optimal time rescaling factor (f_g = 0.1), however, is significantly smaller compared to the other alloys, because the MC time went too fast compared to the experimental one by about a factor 5. Scatter in experimental data cannot explain such a large difference; neither can an eventual change in the vacancy formation energy h_v^f due to the decrease of the Cu content. Instead, a possible explanation could be the incorrectness of the assumption that the attempt frequency Γ_0 in Eq. (1) can be considered as strictly constant. For example, Soisson and Fu used constant values in [18], but varying with the chemical nature of the jumping atom, being large for Fe. Such a difference in the attempt frequency could have compensated the sudden

decrease of f_g, because the proportion of Fe-vacancy exchanges is logically expected to increase if the Cu content is reduced.

To summarize, the hybrid AKMC model proposed here is able to make relevant predictions of the kinetics of Cu precipitation in α-Fe, in an affordable CPU time on standard work-stations. We thus believe that our model includes all important mechanisms of the investigated physical-chemical process and that these are satisfactorily parameterized. This achievement is of course also closely connected to the quality of the interatomic potential used, from which all parameters are obtained, either directly or indirectly.

4 Analysis of the Cu precipitation mechanism

(Fig. 9 about here)

Our results of the previous section strongly support the idea that the mobility of Cu clusters and even precipitates plays a significant role in the process of precipitation in Fe, similarly to the conclusion of Soisson and Fu [18]. Fig. 9 shows the evolution with annealing time, in our simulations of Fe-1.34%Cu at 773K, of: (1) the ratio between the number of Cu atoms absorbed in big clusters after their migration, and the number of Cu atoms dragged to the cluster absorption range by the vacancy; (2) the ratio between the number of times clusters density was reduced because of big clusters merging and because of clusters dissociation by the emission of CuV_1 pairs. These ratios are significantly larger than 1, and constantly increasing with time, confirming that indeed Cu clusters, in our simulations, grew mainly as a consequence of their migration and subsequent inclusion of Cu atoms, and that the clusters density grows mainly due to the coalescence of mobile clusters. To further highlight this conclusion, we have conducted additional simulations (Fig. 10), in which some events where deliberately prohibited, namely: in one, the emission of VCu_1 pairs from VCu_N clusters was suppressed ($p_N = 0$); in the other, the migration frequency $\Gamma_N^{(mig)}$ was artificially modified in order to progressively inhibit the migration of the biggest VCu_N clusters (for $N > 100$). Compared to Fig. 8, in the former case the average precipitate radius is unsurprisingly somewhat larger and the average precipitate density almost unaltered, the results remaining in good agreement with experiments. In the latter, however, the results deviate significantly

190

from the experimental data: from a certain annealing time on, the average precipitate radius ceases to increase, and the clusters density ceases to decrease, thus remaining higher than the experimental one. Therefore, it is only by allowing large clusters to be mobile that experimental results can be matched by the model. The one-by-one emission of VCu_1 pairs, on the contrary, is not a sufficiently efficient mechanism to enable coarsening as observed in experiments.

(Fig. 10 about here)

5 Conclusion

We have simulated the coherent stage of Cu precipitation in α-Fe during thermal annealing with a novel hybrid atomistic kinetic Monte Carlo simulations. The vacancy migration energies used to parameterize the model were calculated using a suitable interatomic potential, taking into account long-range chemical interactions and static relaxation, by exploiting the capabilities of artificial neural networks to interpolate and extrapolate the results of nudged elastic band calculations. This algorithm was hybridized with an object kinetic Monte Carlo approach, by treating Cu precipitates as objects above a certain size. The seamless matching between the atomistic and the coarse-grain approximation was ensured by calculating all parameters governing object behavior from atomistic simulations and again by exploiting the regression capability of artificial neural networks for extrapolation. This allowed the simulation CPU time to be reduced by orders of magnitude and enabled complete thermal annealing experiments to be simulated, up to the end of the coherent precipitation stage, finding in addition very good agreement with experimental data, both in terms of mean size and density of precipitates. This achievement proves: (1) the suitability of the interatomic potential used, which describes correctly not only the thermodynamic properties of the Fe-Cu system (phase diagram), but also the kinetics of precipitation; (2) that the model includes all important mechanisms driving the precipitation of Cu in iron. In particular, the mobility of Cu precipitates containing even several thousands of atoms turns out to be the dominant mechanism leading to growth and coarsening. This is consistent with what was proposed by Soisson and Fu, though could not be fully proven because of the lesser numerical efficiency of their model [18]. Our hybrid model paves the way to addressing more complex phenomena, such as radiation-enhanced, or even induced, precipitation, in Fe-Cu and also in more complex alloys.

191

14

Acknowledgements

References

[1]   G.R. Odette, G.E. Lucas, JOM **53** (7) (2001) 18–22.
[2]   S.R. Goodman, S.S. Brenner, J.R. Low, Metall. Trans. **4** (1973) 2363.
[3]   R. Kampmann, R. Wagner, in: C. Janot, W. Petry, D. Richter, T. Springer (Eds.), Atomic Transport and Defects in Metals by Neutron Scatterings, Springer, New York, 1986. p. 73.
[4]   J.T. Buswell et al., Effects of Radiation on Materials: 14th International Symposium (Volume II), ASTM STP 1046, N.H. Packan, R.E. Stoller, and A.S. Kumar, Eds., American Society for Testing and Materials, Philadelphia, 1990, pp. 127 – 153.
[5]   M.H. Mathon, A. Barbu, F. Dunstetter, F. Maury, N. Lorenzelli, C.H. de Novion, J. Nucl. Mater. **245** (1997) 224.
[6]   M. Perez, F. Perrard, V. Massardier, et al., Philos. Mag. **85** (2005) 2197.
[7]   E. Vincent, C.S. Becquart, C. Pareige, P. Pareige, C. Domain, J. Nucl. Mater. **373** (2008) 387-401
[8]   A. Deschamps, C. Genevois, M. Nicolas, F. Perrard, and F. Bley, Philos. Mag. **85** (2005) 3091.
[9]   S. I. Golubov, Y. N. Osetsky, A. Serra, and A. V. Barashev, J. Nucl. Mater. **226** (1995) 252.
[10]   S. I. Golubov, A. Serra, Y. N. Osetsky, and A. V. Barashev, J. Nucl. Mater. **277** (2000) 113.
[11]   F. Christien and A. Barbu, J. Nucl. Mater. **324** (2004) 90.
[12]   M. H. Mathon, A. Barbu, F. Dunstetter, F. Maury, N. Lorenzelli, and C. H. de Novion, J. Nucl. Mater. **245** (1997) 224.
[13]   F. Soisson, A. Barbu, and G. Martin, Acta Mater. **44**, 3789 (1996).
[14]   S. Schmauder, P. Binkele, Comput. Mater. Sci. **25** (2002) 174
[15]   Y. Le Bouar, F. Soisson, Phys. Rev. B **65** (2002) 094103, and references therein.
[16]   P. Bellon, Thermodynamics, Microstructures and Plasticity **108** (2003) 395-409, and references therein.
[17]   F. Soisson, J. Nucl. Mater. **249** (2006) 235
[18]   F. Soisson, C.C. Fu, Phys. Rev. B **76** (2007) 214102.
[19]   E. Vincent, C.S. Becquart, C. Domain, Nucl. Instr. Meth. Phys. Res. B **255** (2007) 78-84.

[20]   A. Chatterjee, D.G. Vlachos, J. Computer-Aided Mater. Des. **14** (2007) 253–308.
[21]   P. Krasnochtchekov, R. S. Averback, and P. Bellon, Phys. Rev. B **75**, 144107 (2007)
[22]   E. Vincent, C.S. Becquart, C. Domain, J. Nucl. Mater. **382** (2008) 154–159.
[23]   N. Castin, L. Malerba, G. Bonny, M.I. Pascuet, M. Hou, Nucl. Instrum. and Meth. B **267** (2009) 3002.
[24]   F. Soisson, C.S. Becquart, N. Castin, C. Domain, L. Malerba, E. Vincent, J. Nucl. Mater. **406** (2010) 55, and references therein.
[25]   C.S. Becquart, C. Domain, Phys. Status Solidi b **247** (2010) 9.
[26]   C. Domain, C.S. Becquart, L. Malerba, J. Nucl. Mater. **335** (2004) 121.
[27]   J.J Blackstock, G.J. Ackland, Philos. Mag. A **81** (2001) 2127.
[28]   P.J. Orthens, M.L. Jenkins, G.D.W. Smith, Phil. Mag. A **70** (1994) 1-24.
[29]   S. Pizzini, K.J. Roberts, W.J. Phythian, C.A. English, G.N. Greaves, Philos. Mag. Lett. **61** (1990) 223-229.
[30]   N. Castin, L. Malerba, J. Chem. Phys. **132** (2010) 074507.
[31]   H. Jónsson, G. Mills, K. W. Jacobsen, in: Classical and quantum dynamics in condensed phase simulations, B. J. Berne, G. Ciccotti and D. F. Coker Eds.,World Scientific, Singapore, 1998.
[32]   G. Henkelman and H. Jónsson, J. Chem. Phys. **113** (2000) 9901.
[33]   C. M. Bishop, Neural Networks for pattern recognition, Clarendon press, Oxford, 1995.
[34]   R.C.Pasianot and L. Malerba, J. Nucl. Mater. **360** (2007) 118.
[35]   W.M. Young, E.W. Elcock, Proc. Phys. Soc. 89 (1966) 735.
[36]   B. Bortz, M. H. Kalos, and J. L. Lebowitz, J. Comput. Phys. **17**, 10 (1975).
[37]   K. A. Fichthorn and W. H. Weinberg, J. Chem. Phys. **95**, 1090 (1991).
[38]   M.I. Pascuet, N. Castin and L. Malerba, J. Nucl. Mater., accepted.

Figures and tables caption

Fig. 1 – Quality of the ANN predictions of the vacancy migration energies obtained with the NEB method in Fe-Cu alloys. The input variables are a description of the vacancy local atomic environment up to the 11th nearest neighbors. The average error of predictions is 2.0%, and Pearson's product-moment correlation coefficient R^2 is 0.998.

Fig. 2 – Schematic representation of the absorption of a vacancy in a Cu_N cluster in the hybrid AKMC algorithm. Dashed lines represent the radius of the cluster, and the absorption radius. On the left side, the vacancy is approaching the cluster, and is considered absorbed when a migration event is chosen involving one atom of the cluster. After absorption of the vacancy, Cu atoms situated within the absorption range are immediately absorbed in the cluster, and the new radiuses are increased accordingly, as shown on the right side.

Fig. 3 – Schematic representation of the dissociation event for a VCu_N cluster in the hybrid AKMC algorithm. The frequency of occurrence is denoted $\Gamma_N^{(diss)}$. When applied, the vacancy alone is emitted outside the absorption range (delimited by the dashed lines on the figure). A Cu atom is occasionally also emitted, with a probability denoted as p_N.

Fig 4 – Diffusion coefficients D_N of different VCu_N clusters, versus reciprocal temperature ($1/k_BT$), measured with AKMC simulations. Plain lines show interpolation and extrapolation on the reciprocal temperature using f_1 defined in Eq. (4).

Fig. 5 – Lifetimes τ_N of different VCu_N clusters, versus reciprocal temperature ($1/k_BT$), measured with AKMC simulations. Plain lines show interpolation and extrapolation on the reciprocal temperature using f_2 defined in Eq. (5).

Fig. 6 – Evolution of the diffusion coefficient D_N of VCu_N clusters with the number N of Cu atoms, at 773K. Dots show interpolation/extrapolation from AKMC data (Table 1) achieved with Arrhenius plots. The plain line show interpolation/extrapolation achieved using function f_1 (Eq. 4).

Fig. 7 – Probability p_N for a VCu_N cluster to dissociate by the emission of a VCu_1 pair. Dots show the values calculated with 100 000 independent AKMC simulations, at different temperatures. The dashed line show interpolation using function f_3 (Eq. 6) at 2000 K, and the plain line shows extrapolation using function f_3 at 773K.

Fig. 8 – Comparison of the results of our hybrid AKMC simulations (plain lines) with experimental data (diamonds, squares, triangles and circles) described in Tab. 1. For the Fe-1.34%Cu case at 773K, the (x) mark indicates the results obtained in a 128x128x128 unit cells box, whereas the unmarked one was obtained in a 64x64x64 box.

Fig. 9 – Evolution with annealing time, during simulation with the hybrid AKMC, for a Fe-1.34at%Cu alloy at 773K, of (1) the ratio between the total number of Cu atoms admitted in Cu clusters after their migration and the number of Cu atoms dragged by the vacancy from the matrix to the clusters vicinity; (2) the ratio between the number of VCu_N clusters objects merging and dissolutions.

Fig. 10 – Thermal annealing experiment for a Fe-1.34at%Cu alloy at 773K. Parameters in the model were changed compared to Fig. 8: (a) the probability p_N for VCu_N clusters objects to dissolve by the emission of a VCu_1 pair is set to 0; (b) The VCu_N migration frequency $\Gamma_N^{(mig)}$ is modified to inhibit the migration of clusters bigger than 100 atoms.

Tab. 1 – Summary of the sets of experimental data used in this work, focusing on the coherent stage of precipitation (the average clusters radius $\overline{R}$ < 3 nm). In the experimental technique column, APT stands for atom probe tomography, SANS for small angle neutron scattering and SAXS for small angle X-rays scattering.

Tab. 2 – Summary of the results obtained with the hybrid AKMC simulations performed in this work. N_{Cl} denotes the number of Cu clusters, and $\overline{R}$ the average cluster radius. The last simulation (Fe-1.1%Cu at 823K) was first started in a 64x64x64 unit cells box, then interrupted and continued in a 128x128x128 when the number of clusters was lower than 10. The number of AKMC events is proportional to the CPU time.

195

Table 1

Cu content [at%]	T [K]	Number of points	Max. clusters density [m^{-3}]	Clusters density [m^{-3}] at $\overline{R}$ = 3 nm	Exp. technique	Reference
1.34	773	3	1E24	~ 1E23	APT	Goodman et al. [2]
1.34	773	6	4E24	~ 1E23	SANS	Kampmann et al. [3]
1.34	773	3	2.5E24	~ 2E23	SANS	Mathon et al. [5]
1.34	773	> 10	-	-	SAXS	Perez et al. [6]
1.34	873	> 10	-	-		
1.34	973	> 10	-	-		
1.1	823	4	~ 1.5E23	~ 1E23	SANS	J.T. Buswell et al. [4]
0.6	773	6	2.7E23	~ 7E22	APT	Vincent et al. [7]

Table 2

Cu content [at%]	T [K]	AKMC box size (lattice units)	Num AKMC events (*1E9)	N_{Cl} at peak of density	N_{Cl}	$\overline{R}$ [nm]	f_V Eq. (8)	f_g Eq. (9)
1.34	773	64	1.4	47	1	2.55	1.5E-5	0.60
1.34	773	128	4.2	375	11	1.76	1.5E-5	0.60
1.34	873	64	3.5	27	1	2.32	1.0E-4	0.60
1.34	973	64	6.3	9	1	1.9	8.1E-4	0.60
1.1	823	64 /128	1.5 / 15.3	22 (in box 64)	6	2.17	4.4E-5	1.30
0.6	773	128	30.2	26	4	1.81	2.7E-5	0.10

Figure 1
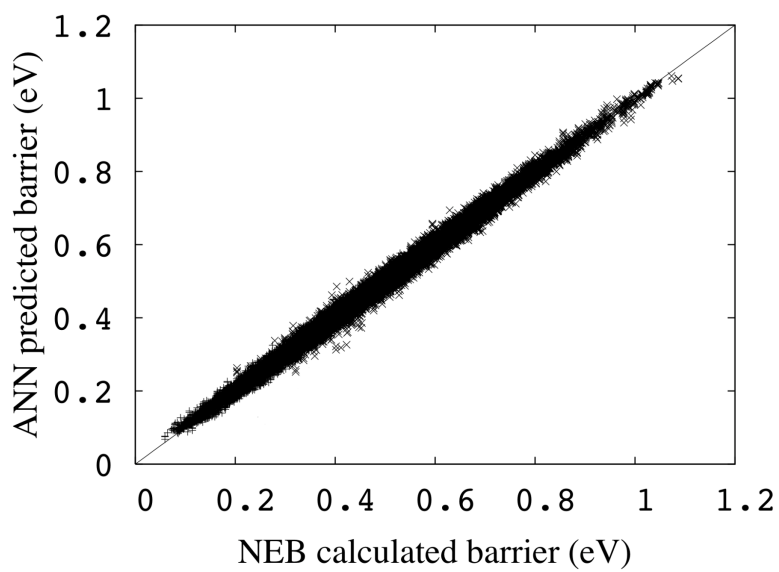
Figure 2
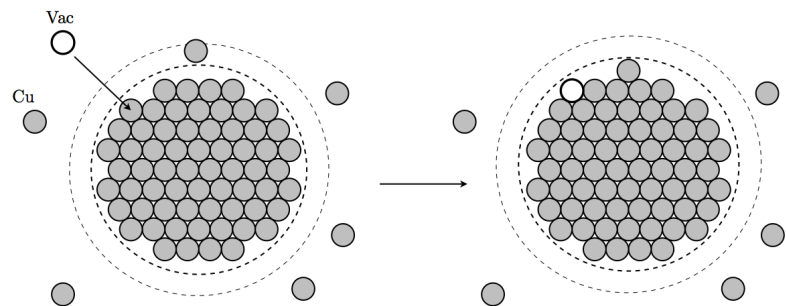
Figure 3
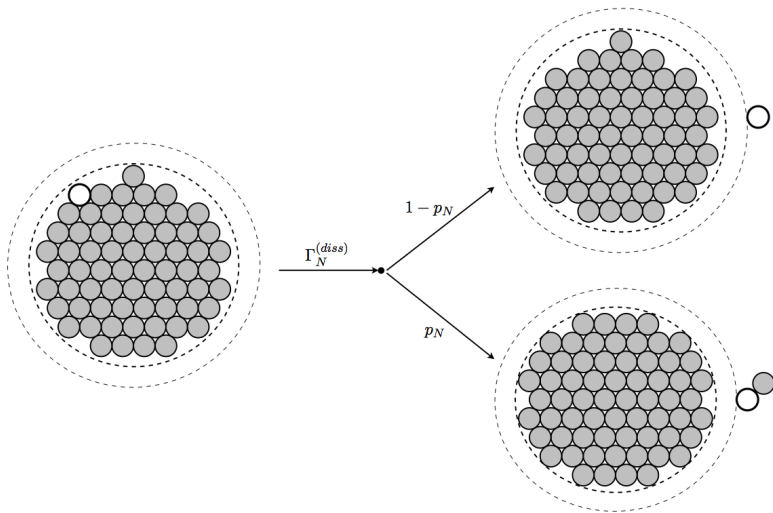
Figure 4
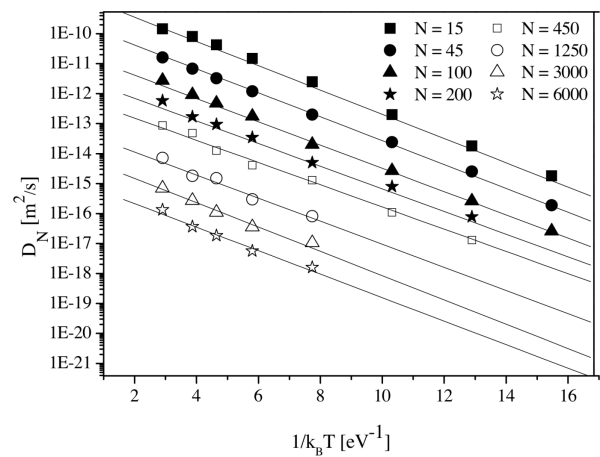
Figure 5
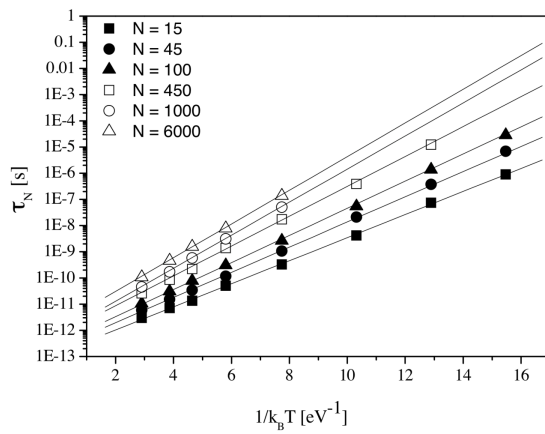
Figure 6
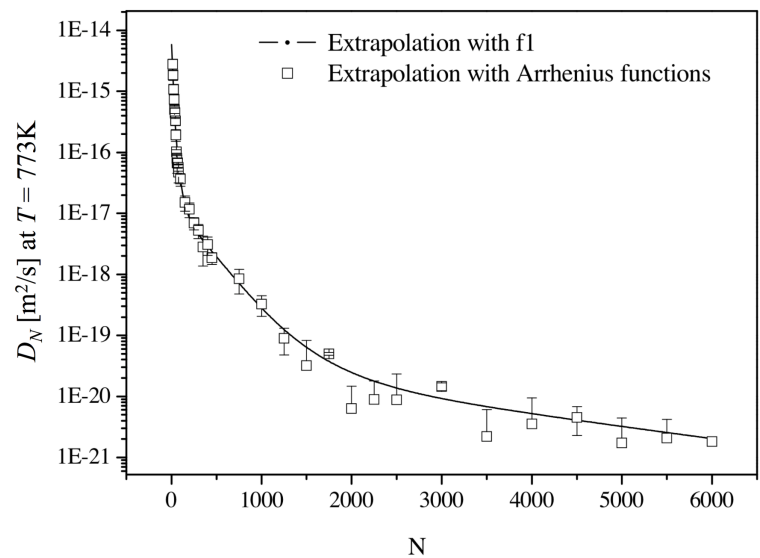
Figure 7
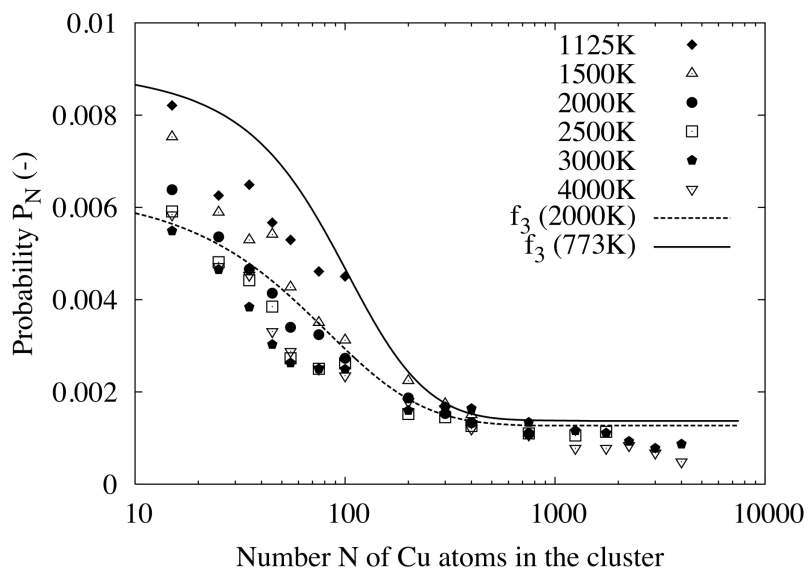
Figure 8
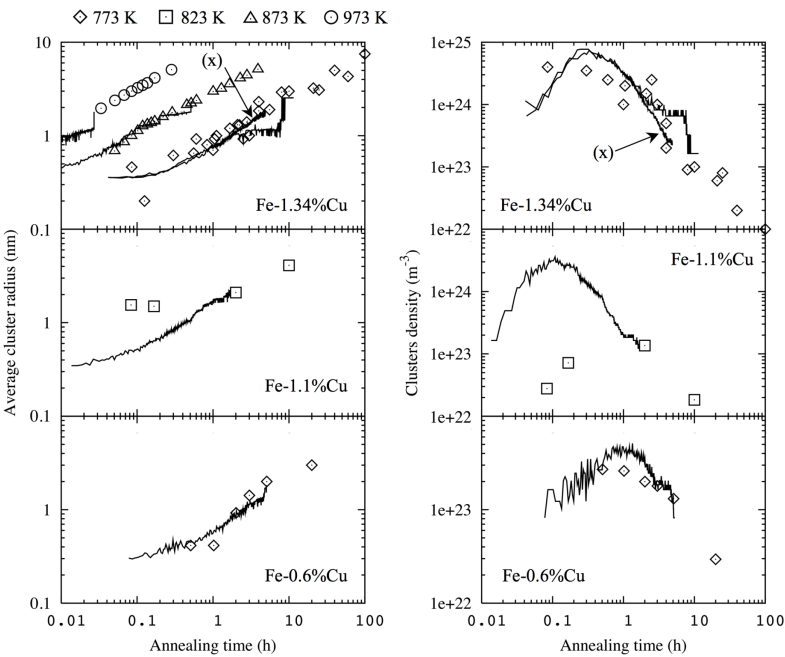
Figure 9
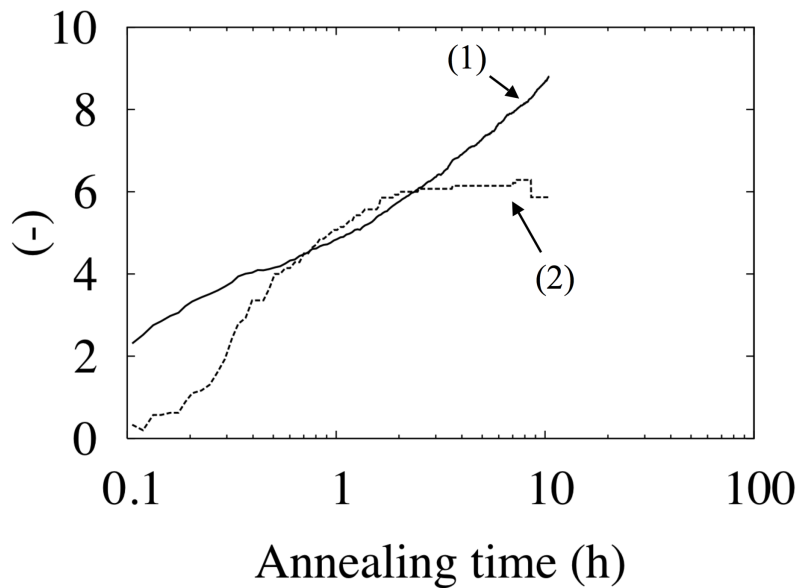
Figure 10
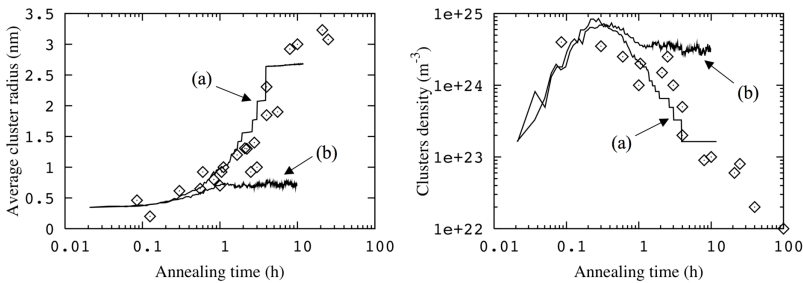
I

Paper VIII

Stability and mobility of Cu–vacancy clusters in Fe–Cu alloys: A computational study based on the use of artificial neural networks for energy barrier calculations

M.I. Pascuet [a], N. Castin [b,c], C.S. Becquart [d], L. Malerba [b,*]

[a] CONICET, Avda. Rivadavia 1917, (1033) Buenos Aires, Argentina
[b] Structural Materials Group, Nuclear Materials Science Institute, Studiecentrum voor Kernenergie (Centre d'Etude de l'Energie Nucléaire (SCK-CEN),
Boeretang 200, (2400) Mol, Belgium
[c] Physique des Solides Irradiés et des Nanostructures (PSIN), Université Libre de Bruxelles, Boulevard du Triomphe CP234, (1050) Brussels, Belgium
[d] Laboratoire de Métallurgie Physique et Génie des Matériaux, UMR 8517, Université Lille-1, F-59655 Villeneuve d'Ascq Cédex, France

ABSTRACT

An atomistic kinetic Monte Carlo (AKMC) method has been applied to study the stability and mobility of copper–vacancy clusters in Fe. This information, which cannot be obtained directly from experimental measurements, is needed to parameterise models describing the nanostructure evolution under irradiation of Fe alloys (e.g. model alloys for reactor pressure vessel steels). The physical reliability of the AKMC method has been improved by employing artificial intelligence techniques for the regression of the activation energies required by the model as input. These energies are calculated allowing for the effects of local chemistry and relaxation, using an interatomic potential fitted to reproduce them as accurately as possible and the nudged-elastic-band method. The model validation was based on comparison with available *ab initio* calculations for verification of the used cohesive model, as well as with other models and theories.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

The formation of copper-rich precipitates under irradiation is accepted to be among the main causes of hardening and embrittlement of reactor pressure vessel (RPV) steels during operation, because they act as obstacles to the motion of the dislocations [1]. Positron annihilation studies on FeCu model alloys show that these precipitates are likely to contain a considerable amount of vacancies and to be formed as a result of the diffusion of mobile complexes containing both Cu atoms and vacancies [2–4] (henceforth Cu–vacancy clusters). Consistently, Cu–vacancy clusters are predicted to be relatively stable by *ab initio* calculations [5], supporting the idea that these clusters should be able to migrate as a whole. This contention has been also qualitatively proven by performing high temperature molecular dynamics studies [6]. Thus, models describing Cu precipitation under irradiation should explicitly include a mechanism of formation based on the diffusion of Cu–vacancy clusters.

Kinetic Monte Carlo (KMC) models based on the residence time algorithm [7,8] are suitable to simulate precipitation and also segregation processes [9,10], in acceptable trade-off between accuracy and computing time [11]. Two main classes of KMC models have

been used to describe Cu precipitation under thermal ageing and irradiation: "atomistic" KMC (AKMC) models [12–20] and "object" KMC (OKMC) models [14,21]. In OKMC simulations point-defects, point-defect clusters and mixed clusters are treated without including the detail of their atomic-level configuration and the technique is suitable to simulate irradiation processes in a fairly realistic way, up to a timeframe of years. However, an OKMC model requires as an input the knowledge of all the parameters defining the mobility and stability of the objects included in it. Notably, in order to introduce the mechanism of formation of Cu precipitates via migration of small Cu–vacancy clusters, the diffusion coefficients of these clusters must be known in advance. These are not quantities that can be experimentally measured, though. Obtaining them from *ab initio* calculations, although *a priori* possible by calculating all migration energies for all relevant migration paths, is a very heavy task, which can only be applied in a few simple cases. Molecular dynamics simulations with interatomic potentials could in principle be a solution, but are in practice not applicable, because of the relatively slow migration of vacancies. Either very high temperatures must be simulated (e.g. [6]), thereby making the trajectory of the clusters too short to be statistically significant with a view to deriving their diffusion coefficient from standard techniques [22], or unaffordably long simulations would be required. In contrast, AKMC models spontaneously treat the diffusion of clusters containing solute atoms and point-defects in terms of

* Corresponding author. Tel.: +32 14 333090; fax: +32 14 321216.
E-mail address: lmalerba@sckcen.be (L. Malerba).

I. Paper VIII

migration jumps of single point-defects on an atomic lattice. In these models each timestep corresponds to a point-defect jump, so the calculation becomes computationally affordable, allowing a precise determination of trajectories and extraction of diffusivity data, with very good statistics. Given the migration energies and the diffusion mechanism, the model automatically explores all possible migration paths. Thus, the AKMC method is the most suitable to estimate the diffusion coefficients of Cu–vacancy clusters, thereby allowing the parameters for OKMC simulations to be produced. The main shortcoming resides in the fact that entropic effects on the migration barriers are not taken into account, although other entropic contributions, such as configurational entropy, are inherently included in the model.

The application of the AKMC algorithm requires the *a priori* knowledge of the migration energies of, in this case, a vacancy, as a function of the local atomic environment (LAE), which varies because Cu atoms and other vacancies are in each case differently distributed around it. These energies are customarily estimated using heuristic approaches, such as linear relationships with the total energy variation between before and after the jump [12–14,17,18,20], or broken bond methods limited to bonds with the first and second nearest neighbour shell [9,10,15–16,19]. However, these approaches are insufficient to describe reliably the complexity of the dependence of the activation energies for a vacancy jump on the LAE, as shown for example in [23].

In [24] the problem was solved by pre-calculating, with a suitable interatomic potential [25,26], all the energy barriers corresponding to all possible LAEs encountered during migration by the vacancies in the Cu–vacancy cluster (limited, however, to the 3rd nearest neighbour shell) and by storing them in tables. This is indeed doable if the LAEs are limited to a sufficiently small amount of atoms and, in general, if the total number of LAEs remains reasonably small. In such a case, the production of the tables requires a long, though still affordable, amount of computing time, but then fast search algorithms can find rapidly the proper value in the tables during the AKMC simulations.

In this work, we want to improve the reliability of the same type of calculations performed in [24], by taking into account the effect on the migration energy of a LAE more extended than the 3rd nearest neighbour (3nn) shell. We also want to study larger clusters, for which the tabulation of all possible barriers for all possible LAEs is unfeasible. In order to do this, we resort to the method developed in [23,27], in which a properly trained artificial neural network (ANN) is used to calculate, on-demand and on-the-fly, during the AKMC simulation, the migration energies as functions of the LAEs.

The paper is organised as follows. In Section 2 we summarise the fundamentals of our artificial neural network-based AKMC simulation. In Section 3 we present our results for a number of Cu–vacancy clusters. In the first place, we provide the statically-calculated formation and binding energies, as these are the quantities traditionally used in OKMC simulations to estimate stability parameters [21]. The calculations are performed using the same interatomic potential as employed for the energy barrier calculations and the results are compared with *ab initio* calculations, in order to assess the reliability of the potential. Then we apply the previously-sketched AKMC method to study mobility and stability of small clusters (up to six elements) versus temperature. Whenever possible, we compare our results with previous ones, obtained in [24]. Subsequently, we present some results on the diffusivity of large copper clusters, containing up to 150 atoms and only one vacancy, and compare the diffusion coefficients obtained with those estimated with other methods. Finally, in Section 4 we analyse and discuss the reliability of the use of the ANN to replace a tabulation of rigorously calculated migration energy values, as well as, more generally, the reliability of the results obtained.

2. Computational model

2.1. The AKMC algorithm and the problem of the evaluation of energy barriers

According to standard transition state theory, the frequency of a thermally activated event, such as a vacancy jump j in an alloy, can be expressed as:

$$\Gamma_j = \nu \exp(-E_j/k_B T) \tag{1}$$

where ν is an attempt frequency, that can be considered constant in first approximation (on the order of the Debye frequency; here $\nu = 6 \times 10^{12}\ \mathrm{s}^{-1}$), k_B is Boltzmann's constant, T the absolute temperature and E_j the LAE-dependent activation energy of the jump (denoted here as energy barrier, or migration energy). These frequencies are used in an AKMC framework to assign a probability to all possible vacancy migration events: the 1nn migration of vacancies, one at a time. The simulation time is incremented following the mean residence time algorithm [7,8].

Given a suitable interatomic potential (e.g. [25,26] for the Fe–Cu alloy), the energy barriers can be rigorously calculated, for any LAE, using the nudged-elastic-band (NEB) method [28]. Atomic-level techniques of this type take implicitly into consideration the influence of the nature of the surrounding atoms on the vacancy jump barrier, as well as strain field effects. However, the use of a NEB calculation at each AKMC timestep to evaluate all Γ_j from Eq. (1) would require a prohibitively long CPU time. As anticipated in the introduction, when the total number of LAEs to be considered is reduced, the production of tables of energy barriers for all possible cases, to be fed to the AKMC code, is possible (e.g. [24]). However, if the clusters to be studied are not-so-small, the size of the table to be constructed grows very fast. In these cases, the only alternative to tabulating is to find patterns in the dependence of the energy barriers on the LAE and use a model instead of tables. For this purpose, we resorted to an artificial intelligence method [27].

2.2. Artificial neural network as regression tool for energy barrier calculation

ANNs are considered as universal approximation tools, capable of learning from experience and to find non-evident dependencies between data. In this work, they are used as a very fast and powerful numerical regression method that can reproduce the complex relation existing between the vacancy migration energy and the LAE. We use ANN of the feed-forward multilayer perceptron type, with one hidden layer, using linear combination functions and sigmoidal non-linear activation functions [27].

The development of an ANN requires first the production of examples, i.e. a table of LAEs and corresponding migration energies calculated with the NEB method. This table must be divided into two non-overlapping sets: a training set and a validation set. The ANN is trained to reproduce the former and validated on the latter, by definition made of examples never seen before by the ANN, i.e. not included in the training set. Given these sets, a complete description of the methodology that is followed in order to design the ANN can be found in [27].

Fig. 1 shows the accuracy reached by the ANN after training, when up to seven Cu atoms and seven vacancies are included in the LAE (FeCuVac system). The LAE encompasses the 5th shell of nearest neighbours around the migrating vacancy and its destination (5nn approximation). The mean relative error is in this case 5.67%, and the correlation factor R^2 is 0.95. The mean bias, $\sum (O - d)/N$, is 0.0024 eV and the mean error, $\sum |O - d|/N$, is 0.033 eV (O is the ANN output, d the desired target, and N the number of points). The correlation is globally good, although not perfect, because ANN predictions and NEB calculated barriers do
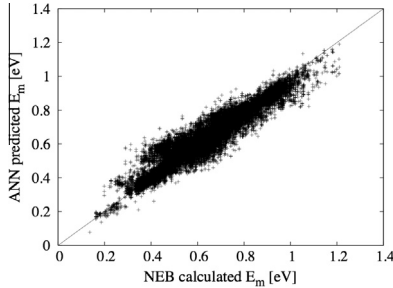
Fig. 1. Predictive capability of the ANN developed for the FeCuVac system (seven Cu atoms and seven vacancies max.), with the LAE extended to 5nn shell, as compared with NEB calculated values from the validation set.
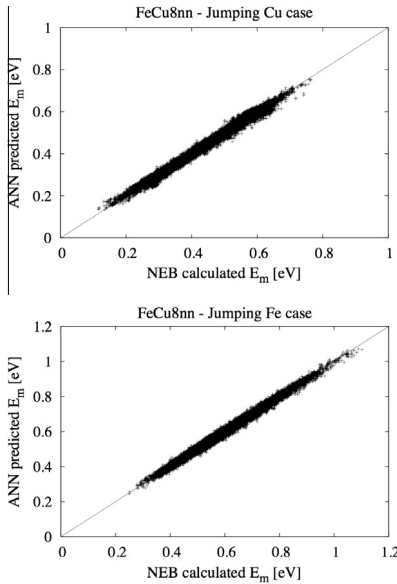

Fig. 2. Predictive capability of the ANN developed for the FeCu system, with LAE extended up to 8nn shell: upper panel, case of a migrating Cu atom; lower panel, case of migrating Fe atom.

not always compare one-to-one. As discussed in Section 4, the quality of the performance of this ANN was *a posteriori* observed to vary significantly from cluster to cluster, so that in principle, via proper retraining, its improvement is possible. Nevertheless, in Section 4 it is also shown that, probably due to the absence of significant biases in the error committed by the ANN, the results are globally acceptable when compared with results obtained directly from tabulations.

In the case of the FeCu system (no extra vacancies in the LAE), two separate ANNs were developed, namely, one for cases when

the vacancy exchanges its position with a Cu atom, and another one for exchanges with Fe atoms. The LAE included up to the 8th nearest neighbour shell (8nn approximation). Fig. 2 shows that in this case the ANN accuracy is extremely good for both Cu and Fe migrating atoms. The mean relative error in the migrating Cu case is 1.13% and 1.10% in the migrating Fe case. The correlation factor R^2 and the mean bias are in both cases 0.99 and 0.001 eV, respectively. The mean error is 0.0051 eV for the migrating Cu case and 0.007 eV for the migrating Fe case.

The ANNs of Figs. 1 and 2 were used to assess E_j in Eq. (1) at each timestep in the AKMC simulations performed to study the diffusivity of, respectively, small Cu–vacancy clusters (FeCuVac system) and Cu clusters containing only one vacancy (FeCu system). The methodology used to extract information from the simulations for the estimation of the diffusion coefficients is described together with the results in the next section.

3. Results

3.1. Formation and binding energies of small clusters

The nanostructural features of main interest for the present investigation are the small Cu–vacancy clusters. As a first characterisation of their stability, static calculations of the relevant formation and binding energies have been performed, using the Fe–Cu interatomic potential developed in [25,26]. The latter has been fitted taking care for a correct description of both the thermodynamic properties of the FeCu system (phase diagram, solubility limit especially) and the interaction between Cu atoms and vacancies in Fe. In particular, migration energies of vacancies in presence of Cu atoms were fitted following indications from available *ab initio* calculations, which suggested that Cu atoms would be dragged by vacancies, as a consequence of the formation of Cu–vacancy pairs migrating together as a whole, as observed also in molecular dynamics simulations [6]. The binding energies between Cu atoms and vacancies were fitted targeting the values that were found to provide the best agreement between experiments and AKMC simulations in [17,29]. Here we calculate the formation and binding energies obtained with this Fe–Cu potential and compare them with *ab initio* calculations. The latter were performed with the density functional theory (DFT) code (VASP) [30], within the Generalised Gradient Approximation (GGA) of Perdew and Wang, PW91 [31] and using fully non-local ultra-soft pseudo-potentials (USPP) of the Vanderbilt type [32] to describe electron–ion interaction. The pseudo-potentials were taken from the code library. The supercell approach with periodic boundary conditions was used to simulate point-defects as well as pure phases. Brillouin zone sampling was performed using the Monkhorst and Pack scheme [33]. The plane wave cut-off energy was 240 eV in order to get converged results. 54 atom supercells with 125 k-points as well as 128 atoms with 27 k-points were used to check the convergence of the calculations with supercell size. Only 128 atom results are reported, which are known from previous experience to be already size-converged (i.e. calculations with more atoms would provide the same results).

Formation energies of vacancy clusters, copper clusters and copper–vacancy clusters, respectively, are obtained using the following formulae:

$$E_f(N_V) = (N_0 - N_V) \cdot [E_c(N_V \, in \, bcc \, Fe) - E_c(bcc \, Fe)]$$
$$E_f(N_{Cu}) = N_0 \cdot E_c(N_{Cu} \, in \, bcc \, Fe) - [(N_0 - N_{Cu}) \cdot E_c(bcc \, Fe) + N_{Cu} \cdot E_c(fcc \, Cu)]$$
$$E_f(N_V + N_{Cu}) = (N_0 - N_V) \cdot E_c(N_V + N_{Cu} \, in \, bcc \, Fe)$$
$$- [(N_0 - N_{Cu} - N_V) \cdot E_c(bcc \, Fe) + N_{Cu} \cdot E_c(fcc \, Cu)]$$
$$\tag{2}$$

109

Table 1
Formation (E_f) and binding (E_b) energies of vacancy–vacancy, Cu–Cu and Cu–vacancy pairs at different mutual distances: interatomic potential (IAP) and *ab initio* (DFT-USPP) calculations using 128 atom supercells.

	E_f (eV)		E_b (eV)	
	IAP	DFT-USPP	IAP	DFT-USPP
V	1.71	2.00	–	–
V–V (1nn)	3.30	3.84	0.13	0.16
V–V (2nn)	3.19	3.70	0.24	0.30
V–V (3nn)	3.45		−0.02	−0.02
V–V (4nn)	3.39		0.03	0.09
Cu	0.43	0.55	–	–
Cu–Cu (1nn)	0.79	0.94	0.08	0.16
Cu–Cu (2nn)	0.79	1.05	0.08	0.05
Cu–Cu (3nn)	0.87		∼0	∼0
Cu–Cu (4nn)	0.87		∼0	
Cu–V (1nn)	2.05	2.39	0.10	0.16
Cu–V (2nn)	2.06	2.34	0.09	0.21
Cu–V (3nn)	2.15		∼0	

Table 2
Formation and binding energies (eV) for some small Cu-vacancy clusters, as calculated with the interatomic potential and using *ab initio* methods. The enclosed element is the one for whose removal the binding energy is evaluated.

	E_f (eV)	E_b (eV)	
	IAP	IAP	DFT-USPP
1	4.62	0.28	0.52
2	5.78	0.57	0.70
3	6.09	0.24	0.31
4	2.31	0.19	0.36
5	2.40	0.10	0.22
6	2.40	0.10	0.18
7	2.31	0.18	0.35
8	2.37	0.21	0.42
9	2.39	0.19	0.33
10	2.69	0.08	0.19
11	2.49	0.28	0.46

Here, N_0 is the total amount of atomic lattice sites in the box, N_V is the number of vacancies in the cluster, and N_{Cu} is the number of copper atoms in the cluster; $E_c(N_V$ in bcc Fe) is the relaxed energy

per atom of a bcc Fe matrix containing a cluster of N_V vacancies in its lowest energy configuration, $E_c(N_{Cu}$ in bcc Fe) is the same energy for a system containing a cluster of N_{Cu} copper atoms, and $E_c(N_V + N_{Cu}$ in bcc Fe) is again the same energy containing a cluster of N_V vacancies and N_{Cu} copper atoms; finally, $E_c($bcc Fe) is the cohesive energy of pure bcc Fe and $E_c($fcc Cu) is the cohesive energy of pure fcc Cu. It should be noted that the formation energies in Eq. (2) are independent of the choice of N_0, provided that the energies per atom of the defective system used in such equation to calculate them correspond to the value of N_0 used in the same equation.

Once the formation energies are known, the binding energy of a vacancy (V) to a cluster of certain size can be obtained according to the following expression:

$$E_b(V) = E_f(cluster) + E_f(V) - E_f(cluster + V) \qquad (3)$$

where $E_f(cluster)$ is the formation energy of the cluster, $E_f(V)$ is the formation energy of the vacancy alone, and $E_f(cluster + V)$ is the formation energy of the cluster plus the vacancy.

The results from the interatomic potential and from *ab initio* calculations, for a number of cluster configurations, are given in Tables 1 and 2. In particular, Table 1 shows the formation and binding energies for the Cu–vacancy, Cu–Cu and vacancy–vacancy pairs. Table 2 lists the same quantities for a number of clusters (results obtained with the interatomic potential for more clusters, up to six elements, are reported in [34]). It can be seen that the potential generally underestimates the strength of the Cu–vacancy binding, as compared to DFT data. At the same time, DFT binding energy values for Cu–vacancy pairs could not be used as they were in [17,20,29] in order to provide acceptable results in AKMC
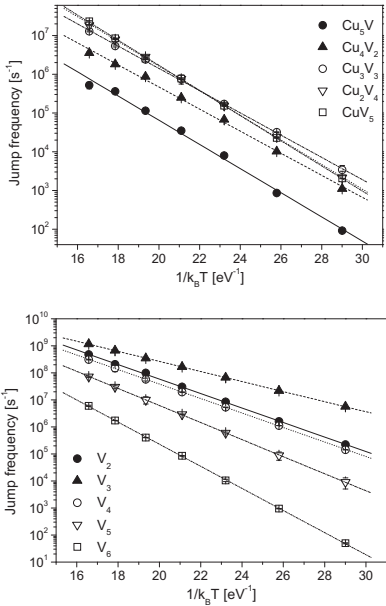

Fig. 3. Jump frequencies of Cu–vacancy clusters with six elements (above) and of vacancy clusters (below).

simulations of Cu precipitation in Fe. The potential provides pair binding energies very close to those chosen for the AKMC model we refer to. Thus, the overall performance of the potential can be judged acceptable, although probably biased on the side of under-estimating the strength of the Cu–vacancy interaction in Fe. This possible bias will have to be taken into account when discussing the results.

3.2. Mobility and stability of small clusters

The mobility and stability of small clusters were investigated using the AKMC model driven by the ANN, described in Section 2. The clusters studied were complexes formed by Cu atoms (Cu) and vacancies (V) up to six elements, namely CuV, CuV$_2$, Cu$_2$V, Cu$_2$V$_2$, CuV$_3$, Cu$_3$V, CuV$_4$, Cu$_2$V$_3$, Cu$_3$V$_2$, Cu$_4$V, CuV$_5$, Cu$_2$V$_4$, Cu$_3$V$_3$, Cu$_4$V$_2$ and Cu$_5$V, as well as vacancy clusters of up to six elements: V$_2$, V$_3$, V$_4$, V$_5$ and V$_6$. The clusters were created in their expected lowest energy configuration at the centre of a box containing an otherwise pure bcc Fe matrix and the system was let to evolve according to the AKMC scheme. The energy barriers were provided in all cases by the FeCuVac ANN, except for the Cu$_5$V cluster, for which the FeCu ANNs were used instead. The simulation box contained $20 \times 20 \times 20$ cubic cells, i.e. 16,000 atoms. Following the methodology applied in [24], the mobility was studied at different temperatures, between 400 K and 700 K, by tracing the successive positions of the centre-of-mass of the cluster (arbitrarily and conveniently assigning the same mass to Cu and V). The latter was considered to be a cluster so long as all elements forming it remained at a mutual distance shorter than the 2nn distance: as soon as this condition was not fulfilled any more, the simulation was stopped and the cluster lifetime in the specific simulation was recorded. For each temperature and cluster, 100–300 simulations were performed, in order to have enough statistics (see Fig. 3).

3.2.1. Jump frequencies

From the analysis of the successive positions and the corresponding time, in each series of simulations for a given cluster at temperature T, the jump frequency could be deduced, using the equation:

Table 3

Attempt frequency (in units of 10^{13} s^{-1}), migration energy (eV) and 1nn average distance for cluster centre-of-mass jumps (in units of lattice parameter, a_0), for all small clusters studied, as obtained from the jump frequency study. For prefactor and migration energies the error bar is also given.

N	Complex	v_0		E_m^v			Dist
2	V$_2$	1.14	±0.18	2.12*	0.62	±0.005	0.62*	0.92
3	V$_3$	0.14	±0.01	0.2*	0.43	±0.001	0.49*	0.92
4	V$_4$	0.89	±0.05	2.7*	0.62	±0.003	0.72*	0.92
5	V$_5$	1.25	±0.08	4.4*	0.72	±0.003	0.88*	0.90
6	V$_6$	3.43	±0.04	62.8*	0.94	±0.005	1.11*	0.89
2	CuV	0.61	±0.26	0.95*	0.67	±0.019	0.65*	1.00
3	CuV$_2$	0.68	±0.13	1.55*	0.65	±0.008	0.68*	0.94
3	Cu$_2$V	0.36	±0.05	1.19*	0.65	±0.007	0.68*	0.94
4	CuV$_3$	0.52	±0.10	0.17*	0.64	±0.008	0.55*	0.92
4	Cu$_2$V$_2$	0.65	±0.05	72.9*	0.69	±0.003	0.86*	0.92
4	Cu$_3$V	0.41	±0.01	8.53*	0.71	±0.012	0.78*	0.92
5	CuV$_4$	0.601	±0.09		0.71	±0.007		0.90
5	Cu$_2$V$_3$	0.181	±0.04		0.65	±0.011		0.90
5	Cu$_3$V$_2$	0.103	±0.03		0.68	±0.012		0.90
5	Cu$_4$V	0.007	±0.01		0.67	±0.018		0.90
6	CuV$_5$	0.502	±0.06		0.75	±0.005		0.89
6	Cu$_2$V$_4$	0.367	±0.04		0.73	±0.005		0.89
6	Cu$_3$V$_3$	0.078	±0.01		0.66	±0.008		0.90
6	Cu$_4$V$_2$	0.023	±0.01		0.66	±0.015		0.89
6	Cu$_5$V	0.011	±0.01		0.72	±0.022		0.90

Data denoted by * correspond to calculations from [24] and are added for comparison.
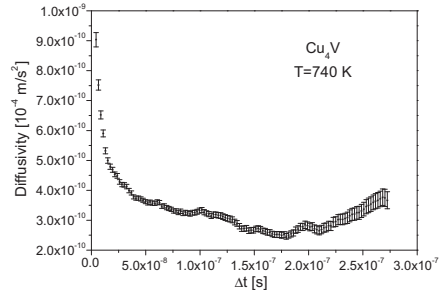

Fig. 4. Example of convergence of the diffusion coefficient as a function of the length of the time interval Δt chosen when applying Eq. (7) ($T = 700$ K, cluster Cu$_4$V).

$$v(T) = \frac{1}{N^{sim}} \sum_{i=1}^{N^{sim}} \frac{n_{jumps}^i}{\tau_{life}^i}(T) \qquad (4)$$

Here, n_{jumps}^i is the number of jumps equal to, or larger than, 1nn distance, performed by the centre-of-mass during the lifetime of the cluster, τ_{life}^i. The migration energies of the clusters were obtained using classical Arrhenius exponential functions to interpolate the data-points obtained for the different temperatures:

$$v(T) = v_0 \exp(-E_m^v/k_B T) \qquad (5)$$

The resulting v_0 and E_m^v for all studied clusters are given in Table 3. For complexes of up to four elements, our results could be compared with those obtained in [24], where energy barriers were calculated with the same interatomic potential as here, but limited to LAEs extended only to the 3nn distance. For visual illustration, the corresponding Arrhenius plots for six element Cu–vacancy clusters and for all vacancy clusters studied are given in Fig. 3 (all plots for all Cu–vacancy clusters can be found in [34]). In Table 3 we provide also the actual average jump distance for each cluster, in units of lattice parameters. In all cases this distance is longer than the 1nn distance and is as large as the 2nn distance for the CuV cluster.

3.2.2. Diffusion coefficients

The diffusion coefficient can be estimated similarly to the jump frequency using the following equation [24]:

$$D(T) = \frac{1}{N^{sim}} \sum_{i=1}^{N^{sim}} \frac{R_i^2}{6\tau_{life}^i}(T) \qquad (6)$$

where R_i^2 is the square of the total displacement of defect i during its lifetime. This scheme is similar to the one originally applied by Guinan et al. for the study of the diffusivity of self-interstitials [35] and amply discussed in [22]. The sampled time, however, is here in each case dictated by the cluster lifetime and is therefore not the same for each run. Eq. (6) corresponds in fact to an adaptation of the general Einstein equation [22,36]:

$$D(T) = \frac{\langle R^2 \rangle(T)}{6\Delta t} \qquad (7)$$

where $\langle R^2 \rangle$ is the mean square displacement within the time interval Δt of a population of random walkers. However, as discussed in [22], the accuracy of the estimated diffusivity with adaptations of Eq. (7) is sensitive to the choice of the time interval length. To analyse in each case (each cluster and temperature) the dependence of the diffusion coefficient and its variance versus isochronal sequences, all trajectories that were independently simulated were

214

eventually chained to one another. The long trajectory thereby obtained was then decomposed into time segments of (approximately) equal length, Δt. The average of the square distances covered by the defect within each interval of length Δt was used

Table 4
Diffusivity prefactor (in units of 10^{-8} m^2/s) and migration energy (eV) for all small clusters studied, including uncertainties, as obtained from the diffusion coefficient study.

N	Complex	D_0			E_m^D		
2	V_2	22.60	±6.24	29.8*	0.63	±0.012	0.63*
3	V_3	0.21	±0.07	1.11*	0.44	±0.015	0.46*
4	V_4	9.77	±2.51	27*	0.62	±0.011	0.71*
5	V_5	9.76	±1.80	45.5*	0.71	±0.009	0.88*
6	V_6	46.30	±5.16	21.6*	0.95	±0.005	1.06*
2	CuV	2.78	±0.73	4.67*	0.64	±0.012	0.63*
3	CuV_2	8.59	±4.74	9.76*	0.69	±0.025	0.67*
3	Cu_2V	5.20	±1.77	6.7*	0.72	±0.015	0.66*
4	CuV_3	2.59	±0.85	1.78*	0.65	±0.015	0.56*
4	Cu_2V_2	2.05	±0.66	2200*	0.68	±0.014	0.89*
4	Cu_3V	1.89	±0.96	65.8*	0.74	±0.023	0.77*
5	CuV_4	1.44	±0.78		0.67	±0.024	
5	Cu_2V_3	0.69	±0.36		0.63	±0.024	
5	Cu_3V_2	0.49	±0.20		0.66	±0.018	
5	Cu_4V	0.75	±0.42		0.75	±0.025	
6	CuV_5	2.08	±0.82		0.72	±0.018	
6	Cu_2V_4	0.98	±0.49		0.68	±0.022	
6	Cu_3V_3	0.36	±0.20		0.64	±0.024	
6	Cu_4V_2	1.09	±0.38		0.72	±0.015	
6	Cu_5V	0.71	±0.45		0.79	±0.028	

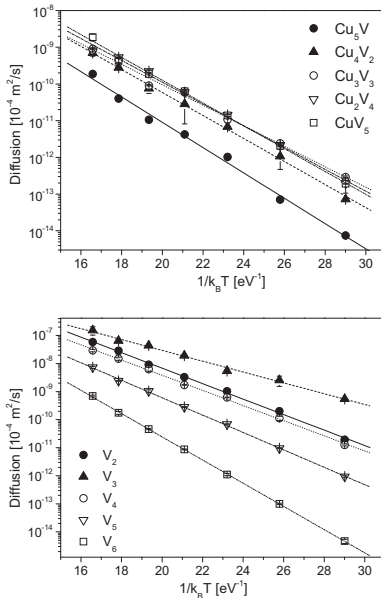Data denoted by * correspond to calculations from [24] and are added for comparison.

Fig. 5. Diffusion coefficients of Cu–vacancy clusters with six elements (above) and of vacancy clusters (below).

to estimate the mean square displacement, which corresponds to the (almost) exact application of the method originally proposed in [35]. By varying Δt in a significantly large interval (possible after chaining trajectories), the converged asymptotical value of the diffusion coefficient is found, as described in what follows.

One example of plot of $D(\Delta t)$ for $T = 740$ K and its variance (error bar) is shown in Fig. 4, for the Cu$_4$V cluster. It can be seen that indeed they depend on the choice of Δt. Short time intervals provide larger statistics, but also larger errors bars, because correlation effects are not correctly sampled. For longer time intervals a convergence to an (almost) Δt-independent D value is observed, due to a more correct sampling of correlation effects in each interval. However, by choosing too long time intervals, the error bars increase again, because the number of intervals over which the average is taken decreases. Hence, the D value cannot be simply obtained as the asymptot for large Δt. Instead, it must be a trade-off between correctness of the sampling and accuracy. For each cluster and temperature, therefore, the diffusion coefficient was estimated as the average between the maximum and the minimum value in the convergence zone. The uncertainty in the calculation of the diffusion coefficient was accordingly estimated as the difference between this average and the largest value in this zone.

Finally, in order to estimate the migration energy and the diffusivity prefactor for each cluster, the data-points for different temperatures were interpolated using the Arrhenius expression:

$$D(T) = D_0 \exp(-E_m^D/k_B T) \tag{8}$$

The resulting D_0 and E_m^D for all studied clusters are provided in Table 4. For complexes up to four elements, the results obtained in [24] are added for comparison. For visual illustration, the corresponding Arrhenius plots for six element Cu–vacancy clusters and for all vacancy clusters studied are given in Fig. 5 (all plots for all Cu–vacancy clusters can be found in [34]).

3.2.3. Correlation factors and cluster lifetimes
The migration energy obtained from the diffusion coefficient, E_m^D, may differ from the homologous value obtained from the jump frequency, E_m^v, because in the former case correlation effects are naturally allowed for, but not in the latter. These effects are related to jumps or series of jumps that, while accounted for in the determination of the jump frequency via Eq. (4), do not contribute to the diffusion coefficient (e.g. back and forth jumps, or loops in the trajectory). In general, the relationship between diffusion coefficient and jump frequency for three-dimensionally migrating species can be expressed as:

$$D(T) = f_c(T) \frac{v_j(T)\Delta^2}{6} \tag{9}$$

where Δ is the jump distance (between 1nn and 2nn distance in the present case) and f_c is the correlation factor, that carries the abovementioned effects. This factor can be temperature dependent. In the case of all clusters studied in this work, at any rate, and for all temperatures, the correlation factors are always close to unity [34]. Consistently, the migration energies obtained with the two interpolations are extremely similar.

The dissociation energy, E_{dis}, and the lifetime prefactor, τ_0, were estimated from the temperature dependence of the average lifetime, using the equation:

$$\tau_{life}(T) = \tau_0 \exp(E_{dis}/k_B T) \tag{10}$$

The results for all clusters studied are provided in Table 5. For visual illustration, the corresponding Arrhenius plots for six element Cu–vacancy clusters and for all vacancy clusters studied are given in Fig. 6 (all plots for all Cu–vacancy clusters can be found in [34]).

Table 5
Lifetime prefactor (in units of 10^{-14} s) and dissociation energy (eV) for all small clusters studied, specifying error bars, too.

N	Complex	τ_0			E_{dis}		
2	V_2	0.99	±0.12	0.75*	0.81	±0.007	0.81*
3	V_3	1.48	±0.19	0.98*	0.82	±0.006	0.91*
4	V_4	0.39	±0.07	0.3*	0.97	±0.008	1.08*
5	V_5	0.36	±0.08	0.9*	1.10	±0.010	1.2*
6	V_6	0.17	±0.07	0.04*	1.30	±0.019	1.45*
2	CuV	8.26	±0.12	8.3*	0.69	±0.007	0.71*
3	CuV_2	0.70	±0.23	2.45*	0.83	±0.015	0.81*
3	Cu_2V	6.53	±0.16	432*	0.72	±0.011	0.70*
4	CuV_3	5.18	±0.78	2.14*	0.80	±0.007	0.88*
4	Cu_2V_2	1.10	±0.16	0.65*	0.88	±0.006	0.96*
4	Cu_3V	3.16	±0.91	68.6*	0.83	±0.013	0.81*
5	CuV_4	2.65	±1.24		0.90	±0.021	
5	Cu_2V_3	4.37	±0.95		0.84	±0.009	
5	Cu_3V_2	1.76	±0.50		0.90	±0.012	
5	Cu_4V	3.45	±0.92		0.88	±0.011	
6	CuV_5	3.21	±1.43		0.96	±0.020	
6	Cu_2V_4	3.34	±0.97		0.92	±0.013	
6	Cu_3V_3	5.86	±2.02		0.86	±0.015	
6	Cu_4V_2	3.02	±0.57		0.90	±0.006	
6	Cu_5V	0.38	±0.11		0.98	±0.013	

Data denoted by * correspond to calculations from [24] and are added for comparison.

Fig. 7. Diffusion coefficient versus number of Cu atoms in precipitates at 400, 500 and 600 K. The solid lines correspond to the present work; dashed lines correspond to the results of [19]: full symbols are AKMC simulations, empty symbols calculations using the Binder–Stauffer model.
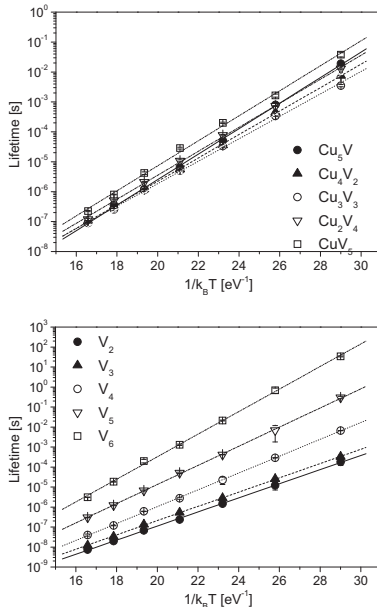

Fig. 6. Lifetime of Cu–vacancy clusters with six elements (above) and of vacancy clusters (below).

3.3. Mobility of large Cu clusters

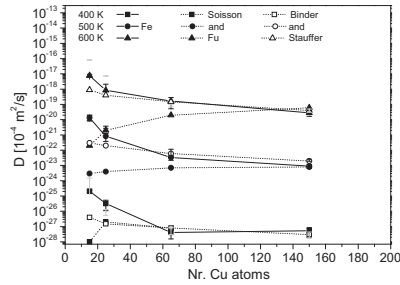Large clusters of 15, 25, 65 and 150 Cu atoms and only one vacancy were created in a $30 \times 30 \times 30$ cell simulation box, with a pure bcc Fe matrix. The FeCu-8nn ANNs were used for the calculation of the energy barriers. The mobility was studied by tracing the successive positions of the centre-of-mass of the cluster at three temperatures: 400 K, 500 K and 600 K. The diffusion coefficients were studied in the same way as for small clusters (Section 3.2.2). The results are given in Fig. 7, where the diffusion coefficient versus the number of Cu atoms in the cluster is shown for the three different temperatures studied. Results by Soisson and Fu obtained from similar studies conducted with a different AKMC model [19], and according to a mean-field Binder–Stauffer model [37–39], are shown for comparison on the same graph, after renormalisation. The latter was necessary because the raw diffusion coefficients obtained in the present study differed by several orders of magnitude from those obtained in [19]. One reason is that the diffusivity prefactor values bear a relationship with the chosen constant attempt frequency in Eq. (1): $\nu_{Fe} = 5 \times 10^{15}$ s^{-1} and $\nu_{Cu} = 2 \times 10^{15}$ s^{-1} in [19], while here $\nu_{Fe} = \nu_{Cu} = 6 \times 10^{12}$ s^{-1}. Moreover, in [19] time had been rescaled following the equation:

$$t = t_{MC} \frac{C_V^{MC}(Fe)}{C_V^{eq}(Fe)} \tag{11}$$

where $C_V^{MC}(Fe)$ is the vacancy concentration at the simulation box and $C_V^{eq}(Fe)$ is the vacancy concentration at equilibrium, exponentially proportional to the vacancy formation energy. We empirically determined that, by using 0.84 eV as "effective" vacancy formation energy, by applying Eq. (11) we could get values of the same order of magnitude as in [19], and therefore directly comparable, at least in terms of trends.

Qualitatively, the AKMC results from [19] exhibit the peculiarity that the mobility increases with size, instead of decreasing, as would seem more intuitive and logical, and as consistent with the mean-field Binder–Stauffer model [37]. According to our AKMC model the mobility of the Cu precipitates decreases with size, following closely the classical Binder–Stauffer trend. This different behaviour can most likely be rationalised in terms of difference between the vacancy formation energy in bcc Cu and bcc Fe. This difference is much more pronounced in the AKMC model from [19] than in our model ($E_V^f(Cu) = 0.82$ eV and $E_V^f(Fe) = 2.20$ eV in [19], while in our case $E_V^f(Cu) = 1.26$ eV and $E_V^f(Fe) = 1.71$ eV, [26]). This makes it possible that in Soisson and Fu's model the vacancy spends a significantly larger fraction of time inside the precipitate: the larger the precipitate, the longer. So, the possibility that the cluster migrates via surface or sub-surface hopping of the vacancy is higher than in our model.

113

Finally, the results shown in Fig. 7 were also obtained introducing the solubility limit of Cu atoms in the matrix. The diffusion coefficients obtained were essentially the same, except for the smaller clusters (15 and 25 atoms) at high temperature: in this case the introduction of Cu in the matrix reduced the mobility by about one order of magnitude.

4. Discussion

In this section, the reliability of the above results is discussed from two viewpoints: firstly, we analyse the capability of the artificial neural network to transfer the physical information that stems out of the interatomic potential to the AKMC model without significant loss; secondly, we shortly discuss the accuracy and the limitation of the interatomic potential, necessarily taking *ab initio* data as a reference.

4.1. Comparison between diffusivity results from ANN and tabulated barriers

Figs. 8 compare the dissociation energies obtained in the present work and in [24], for clusters of vacancies (upper panel) and mixed (lower panel). It can be seen that, despite the somewhat different methodology used in the two cases, the two sets of results are acceptably correlated. This suggests that, given an interatomic potential, the results are relatively robust versus the details of the way in which the energy barriers are actually estimated in the AKMC simulation.
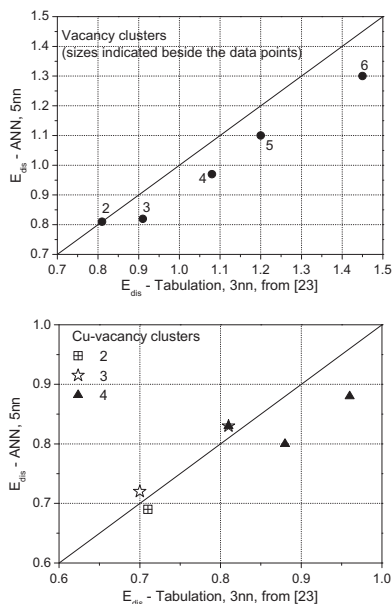

Fig. 8. Correlation between dissociation energies for vacancy clusters (above) and Cu-vacancy clusters (below) obtained with similar, yet different, methods, here (*Em* – ANN results) and in [24] (*Em* – Tabulation).

Next, the capability of the ANN (FeCuVac-5nn ANN) to reproduce the migration energies corresponding to the configurations encountered by the different clusters during the simulation was further tested, in the following way. A number of LAEs among those encountered were randomly extracted and the corresponding energy barrier rigorously calculated, using the NEB method. The values thereby obtained were then compared, for each cluster, with the prediction made by the ANN. The results of this comparison for all clusters are reported in [34]. Here a couple of examples of comparisons are given in Fig. 9 (upper panel), while the worst case of all is illustrated by Fig. 10. Clearly, the performance of the ANN is not the same for all clusters. This study suggests that the origin of the possible unreliability of a given ANN does not stem from the presence of a large number of defects in the LAC. Instead, it must be considered strictly as a solvable mathematical problem, related with the kind of examples (initially randomly chosen) on which the ANN is trained. Most likely, for the clusters whose energy barriers are less accurately predicted, the training set contained very few (too few) examples. The reliability of the ANN could therefore be no doubt further improved by retraining with a better selection of examples.

In order to quantify even better up to what extent errors committed by the ANN affect the prediction of the diffusion coefficient, the latter was calculated both using the ANN and the complete tabulation of NEB energy barriers, in the case of CuV_2 and CuV_3 (these clusters are small enough to allow the energy barriers for all possible 5nn LAEs to be calculated by NEB and tabulated; they represent average cases in terms of ANN accuracy). The results are shown in the lower panel of Fig. 9. It can be seen that the ANN error has only a negligible influence on the final result. It is thus concluded that, with the possible exception of extreme cases (such as Cu$_3V_3$, Fig. 10), the use of the ANN is essentially equivalent to the use of tabulations of NEB calculated barriers.

4.2. Limits of the obtained values

The previous section proves that the 'filter' introduced by the ANN between the properties of the interatomic potential used for energy barrier calculations and the AKMC model is generally 'permeable' and that using the ANN is equivalent to calculating by NEB all migration energies on-the-fly, though at a greatly reduced computational cost. There are, however, two *caveats*.

The first one is that in our AKMC model a constant attempt frequency is assumed ($\nu = 6 \times 10^{12}$ s^{-1}). In principle, this quantity is LAE-dependent, too. Given an interatomic potential the attempt frequency can be rigorously calculated for a given LAE, applying Vineyard's equation [40] and examples of calculations could be also used to train a separate ANN, specialised on attempt frequencies. The main reason to avoid this additional procedure is that small changes in the values of the migration energies, by entering an exponential function, will certainly have a much stronger impact than equally small changes in the attempt frequency (so long as these are not systematic). On the other hand, clearly the values for the jump frequencies and diffusion coefficients obtained in this work are scaled by the choice of the value of ν.

The second *caveat* concerns the reliability of the interatomic potential versus the only set of data we can use for validation, i.e. *ab initio* data, given that experimental data on binding and migration energies of Cu–vacancy clusters cannot exist. The migration energies are well reproduced by the potential when compared to the *ab initio* data available at the time of the fitting [26]. In particular, the potential we used has been demonstrated to be suitable to describe the dragging of Cu atoms by migrating vacancies, contrary to other potentials [26]. However, also *ab initio* data are affected by uncertainties. Most notably, the reference *ab initio* values of binding energies used to fit the potential were different from the most
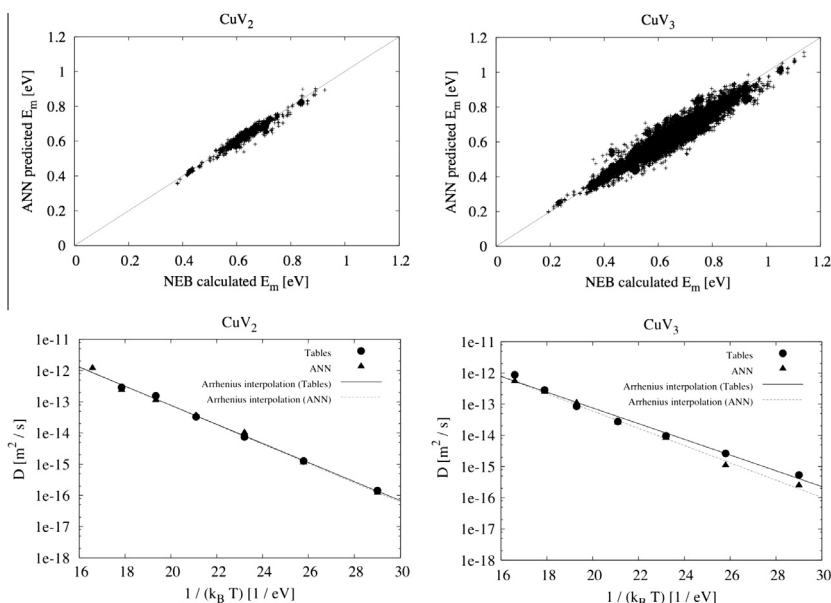
Fig. 9. Upper panel: Correlation between the migration energies predicted by the ANN and those calculated by NEB, for the two clusters CuV_2 and CuV_3 (mean errors: 1.77% for CuV_2, 2.81% for CuV_3). Lower panel: corresponding diffusion coefficient calculated from the complete tabulation of all possible migration energies and from the ANN.
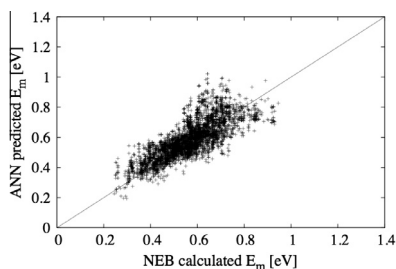

Fig. 10. Correlation between the migration energies predicted by the ANN and those calculated by NEB, in the case of the Cu_3V_3 cluster (worst case): mean error = 11.79%, Pearson's correlation index = 0.78.

recent ones reported here. As anticipated, the comparison made in Tables 1 and 2 suggests that the Cu–vacancy binding energies are systematically underestimated by the potential. (At the same time, it should be remembered that lower binding energy values than DFT ones were needed to properly model Cu precipitation in Fe with an AKMC model, as shown in [17,20,29].) If the potential underestimates the Cu–vacancy binding energies, of all quantities calculated here, the dissociation energy and the lifetime of the clusters are the most affected and are, therefore, probably underestimated by our results. Thus, in reality these clusters might be

more strongly bound and survive for exponentially larger times before dissolving. We are any way confident that at least the trends should remain largely acceptable.

5. Concluding remarks

We have shown here how it is possible to rely on atomistic kinetic Monte Carlo tools, exploiting advanced computational techniques such as artificial neural networks for the prediction of vacancy migration energies as functions of the local atomic environment, to provide an assessment of the stability and mobility of mixed copper–vacancy clusters in iron. These quantities are in practice inaccessible to experiments or to molecular dynamics simulations and can alternatively be obtained only by heavy *ab initio* calculations, for a limited number of cases. The reason for studying these clusters is that they are expected to play a key role in the process of copper precipitation in iron alloys under irradiation. The mobility and stability parameters deduced in this work can so now be used to parameterize models, such as object kinetic Monte Carlo, describing the nanostructure evolution of these alloys under irradiation.

The reliability of the artificial neural network to predict the migration energy of a vacancy as a function of the local atomic environment was generally good and in some cases excellent. It has been seen that relatively poor performances are not due to inherent problems of the method, but only to the lack of proper examples on which the neural network is trained. The main limitation of the method is that, clearly, it cannot do better than the potential used to produce the examples that are provided and the

number of examples that are required is fairly large. So, at the moment, it does not seem possible to train the neural network directly on *ab initio* results, although in principle this is a possible route to follow.

References

[1] G.R. Odette, G.E. Lucas, JOM 53 (7) (2001) 18–22.
[2] Y. Nagai, Z. Tang, M. Hassegawa, T. Kanai, M. Saneyasu, Phys. Rev. B 63 (2001) 134110.
[3] Y. Nagai, K. Takadate, Z. Tang, H. Ohkubo, H. Sunaga, H. Takizawa, M. Hasegawa, Phys. Rev. B 67 (2003) 224202.
[4] M. Lambrecht, L. Malerba, A. Almazouzi, J. Nucl. Mater. 378 (2008) 282.
[5] C.S. Becquart, C. Domain, Nucl. Instrum. Methods Phys. Res. B 202 (2003) 44.
[6] J. Marian, B.D. Wirth, G.R. Odette, J.M. Perlado, Comp. Mater. Sci. 31 (2004) 347–367.
[7] W.M. Young, E.W. Elcock, Proc. Phys. Soc. 89 (1966) 735.
[8] A.B. Bortz, M.H. Kalos, J.L. Lebowitz, J. Comp. Phys. 17 (1975) 10.
[9] F. Soisson, Philos. Mag. 85 (2005) 489.
[10] F. Soisson, J. Nucl. Mater. 349 (2006) 235.
[11] A. Chatterjee, D.G. Vlachos, J. Computer-aided Mater. Des. 14 (2007) 253.
[12] B.D. Wirth, G.R. Odette, Mater. Res. Soc. Symp. Proc. 540 (1999) 637.
[13] C. Domain, C.S. Becquart, J.-C. van Duysen, Mater. Res. Soc. Symp. Proc. 540 (1999) 643.
[14] C. Domain, C.S. Becquart, J.-C. van Duysen, Mater. Res. Soc. Symp. Proc. 650 (2001) R3.2.1–R3.2.6.
[15] S. Schmauder, P. Binkele, Comp. Mater. Sci. 25 (2002) 174.
[16] Y. LeBouar, F. Soisson, Phys. Rev. B 65 (2002) 094103.
[17] E. Vincent, C.S. Becquart, C. Domain, J. Nucl. Mater. 351 (2006) 88.
[18] E. Vincent, C.S. Becquart, C. Domain, Nucl. Instrum. Methods Phys. Res. B 255 (2007) 78.
[19] F. Soisson, C.C. Fu, Phys. Rev. B 76 (2007) 214102.
[20] E. Vincent, C.S. Becquart, C. Pareige, P. Pareige, C. Domain, J. Nucl. Mater. 373 (2008) 387.
[21] C. Domain, C.S. Becquart, L. Malerba, J. Nucl. Mater. 335 (2004) 121.
[22] Yu.N. Osetsky, Def. Diff. Forum 188–190 (2001) 71.
[23] F. Djurabekova, R. Domingos, G. Cerchiara, N. Castin, E. Vincent, L. Malerba, Nucl. Instrum. Method Phys. Res. B 255 (2007) 8.
[24] F.G. Djurabekova, L. Malerba, C. Domain, C.S. Becquart, Nucl. Instrum. Methods Phys. Res. B 255 (2007) 47.
[25] L. Malerba, R.C. Pasianot, SCK•CEN External Report ER-6, February 2006.
[26] R.C. Pasianot, L. Malerba, J. Nucl. Mater. 360 (2007) 118.
[27] N. Castin, L. Malerba, J. Chem. Phys. 132 (2010) 074507.
[28] H. Jonsson, G. Mills, K.W. Jacobsen, in: B.J. Berne, G. Ciccotti, D.F. Coker (Eds.), Classical and quantum dynamics in condensed phase simulations, World Scientific, Singapore, 1998.
[29] E. Vincent, PhD dissertation, University of Lille, 2006.
[30] G. Kresse, J. Hafner, Phys. Rev. B 47 (1993) 558;
ibid. 49 (1994) 14251.
[31] J.P. Perdew, Y. Wang, Phys. Rev. B 45 (1991) 13244.
[32] D. Vanderbilt, Phys. Rev. B 41 (1990) 7892;
G. Kresse, J. Hafner, J. Phys.: Condens. Matter 6 (1996) 8245.
[33] H.J. Monkhorst, J.D. Pack, Phys. Rev. B 13 (1976) 5188.
[34] M.I. Pascuet, N. Castin, L. Malerba, SCK•CEN External Report ER-93, May 2009.
[35] M.W. Guinan, R.N. Stuart, R.J. Borg, Phys. Rev. B 15 (1977) 699.
[36] P.G. Shewmon, Diffusion in Solids, McGraw Hill, 1963 (Chapter 2).
[37] K. Binder, D. Stauffer, J. Stat. Phys. 6 (1972) 49.
[38] K. Binder, M.H. Kalos, J. Stat. Phys. 22 (1972) 363.
[39] K. Binder, D. Stauffer, Phys. Rev. Lett. 33 (1974) 1006.
[40] G.H. Vineyard, J. Chem. Phys. Sol. 3 (1957) 121.

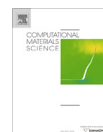# J Paper IX

Contents lists available at ScienceDirect

Computational Materials Science

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

ELSEVIER

Mobility of small clusters of self-interstitial atoms in dilute Fe–Cr alloy studied by means of atomistic calculations

D. Terentyev *, N. Castin

SCK-CEN, Nuclear Materials Science Institute, Boeretang 200, B-2400 Mol, Belgium

A B S T R A C T

Atomistic simulations have been used to characterize the interaction and mobility of small clusters of self-interstitial atoms (SIAs) in dilute Fe–Cr alloys. The variety of migration mechanisms for Di- and Tri-SIA clusters in the bcc Fe matrix were studied using the nudged elastic band method. The corresponding binding and migration energies for the SIA clusters interacting with isolated Cr atoms and Cr–Cr close pairs were calculated using the two-band model interatomic potential. The obtained results are discussed in the light of available experimental data for dilute Fe–Cr alloys and are compared with results obtained using *ab initio* calculations.

1. Introduction

The mobility and stability of point defect clusters that are formed under irradiation essentially determine the evolution of the microstructure. This work is devoted to the study of the mobility of self-interstitial atom (SIA) defects in Fe in the presence of Cr solutes. Here, we apply atomistic calculations to consider properties of small SIA clusters in dilute Fe–Cr alloys.

Resistivity recovery experiments suggest that in electron-irradiated Fe–Cr alloys containing less than 1%Cr, self-interstitial atoms in the form of stable mixed Fe–Cr dumbbells are created and migrate slightly faster than Fe–Fe dumbbells in pure Fe [1,2]. This was confirmed by recent *ab initio* calculations, which have shown that an Fe–Cr $\langle 1\,1\,0 \rangle$ dumbbell is stable and its migration energy (via movement of Cr) is lower than that of the Fe–Fe dumbbell by ~0.1 eV [3]. In dilute alloys no significant shift of the position of the peak corresponding to stage II (attributed to the long-range migration of small interstitial clusters) was observed in Fe–Cr alloys containing up to ~0.1%Cr [1,2], even though the total resistivity recovered during stage II was slightly higher in Fe [2]. By increasing the Cr concentration up to 3%, the damage retained

above stage I_E (attributed to the onset of long-range migration of self-interstitial atoms) up to the beginning of stage III (associated with free migration of vacancies) is observed to be higher than in pure Fe and dilute alloys [1]. This effect was ascribed to the trapping of self-interstitials in specific atomic configurations involving more than one Cr atom, which prevents recombination with vacancies [1]. In the concentrated alloys, the features of the stage II (i.e. number of peaks, their amplitudes and positions) strongly depend on the Cr content [4]. Whereas in pure Fe, the onset of Di-SIA migration is believed to determine the stage II [5]. The strength and concentration of traps for single SIAs is believed to be different in concentrated Fe–Cr alloys containing at least up to 16%Cr [4], which is in line with the atomistic calculations presently available [6]. It is therefore clear that the mobility and clustering of small self-interstitial defects differ significantly in dilute and concentrated alloys. In our previous work [7] we have carried out a set of atomistic calculations to validate the ability of the existing two-band model Fe–Cr empirical potential (EP) [8] to predict some important features of self-interstitial – Cr interaction. EP results were compared to data obtained by density functional theory (DFT) and were discussed in accordance with experimental indications. The applied EP was found to provide reasonable and sometimes unexpectedly good agreement with DFT data, we therefore use it in the present work to study the stability and mobility of

* Corresponding author. Tel.: +32 14 333197; fax: +32 14 321216.
E-mail address: dterenty@sckcen.be (D. Terentyev).

ARTICLE IN PRESS

small interstitial clusters interacting with Cr atoms. The main goal of this work is to see up to what extent the presence of Cr in dilute solution may affect the mobility of Di- and Tri-interstitial clusters that are believed to be responsible for the appearance of stage II, at least in Fe.

According to the atomic-level studies of self–interstitial defects in α-Fe performed up today [9–13], it is recognized that: (i) the migration of a single SIA in its ground state (i.e. ⟨1 1 0⟩ dumbbell) occurs via translation–rotation mechanism and the corresponding migration energy (E_m) is 0.34–0.37 eV; (ii) Di-SIA (I2), which is a set of two parallel ⟨1 1 0⟩ dumbbells situated as first nearest neighbours, migrates via the same mechanism with $E_m = 0.42$ eV [10]. In a recent work, focused on the simulation of resistivity recovery during isochronal annealing of electron-irradiated Fe, a combination of event kinetic Monte Carlo and DFT calculations was applied [13]. According to that work, the migration energy and migration mechanisms of the Tri-SIA cluster (I3) are essentially the same as for the I2. Larger clusters were assumed to be immobile in this work, which was essential for the reproduction of the stages III and V. Later on it has been shown that this assumption is not unreasonable because I3, I4, I5 and probably larger clusters may occupy practically immobile configurations [14]. With the help of molecular dynamics simulations, later on confirmed by more accurate DFT calculations, it was found that except for the high symmetry configurations of SIA clusters, some low symmetry (so-called 'non-parallel') configurations for I2, I3, I4 and probably larger clusters may also exist. The formation energy of the low symmetry configurations can be even lower than that of the canonical configurations [14]. In addition, the 'non-parallel' configurations are furthermore stabilized at finite temperature due to the excess of the vibration entropy [14].

Given that the occurrence of stage II in Fe can be explained by the onset of the long-range migration of the I2 clusters, we do not consider the above-mentioned unusual SIA clusters for the moment. Here, we study the effect Cr atoms on the migration mechanism and corresponding migration energy of single SIA and small 'canonical' SIA clusters (shown in the upper row of Fig. 1 in [14]) containing up to three self-interstitial defects. The main focus of this work is put on dilute Fe–Cr alloys, therefore we do not consider the interaction of SIA clusters with Cr clusters and include only pairs of Cr that might be formed while the Fe–Cr dumbbell, capable of dragging Cr [3], is migrating in the lattice. To do so we carry out a set of molecular static (MS) calculations to identify ground states and to estimate their migration energy and path applying the nudged elastic band method. In this work, calculations were performed using the two-band model EP from [8], as the number of configurations to be explored is significant and therefore the performance of this parametric study with DFT would be a heavily time consuming task. Some important configurations were however cross-checked against DFT calculations. The EP applied here was already extensively tested in terms of the description of different SIA-Cr and SIA-Cr–Cr configurations [7] and fitted to a number of key properties of the Fe–Cr system obtained from DFT calculations [8].

2. Simulation technique

The present MS calculations were carried out using the EP of embedded atom method type from [15] and [8] for Fe–Fe and Fe–Cr, Cr–Cr interactions, respectively. The calculation of migration paths was performed at constant volume (the lattice unit of Fe crystal was set to 2.8553 A) in supercells originally containing 1024 Fe atoms. The size of the supercells was varied to establish the convergence of the obtained migration energies, that in turn were calculated using the nudged elastic band (NEB) method [16].

The initial and final states of the system are first energy optimized with a conjugate gradient (CG) method [17], using an unconstrained static relaxation of all atomic positions. An SIA or SIA cluster initially introduced in the box may therefore change both its position and orientation during the relaxation. As we intend to calculate the transition barrier between two given configurations, described in a rigid lattice way, a subsequent analysis of the two relaxed states is necessary to check if they can still be considered as identical to those that were introduced in the unrelaxed box. The SIA central position and orientation are determined using the Wigner Seitz (WS) cell method and then compared with the desired ones. The relaxed SIA is considered to be identical to the desired one if: (i) its central position belongs to the same WS cell; (ii) the angle between its new orientation (obtained after relaxation) and desired one is smaller than between any other possible ⟨1 1 1⟩, ⟨1 1 0⟩ or ⟨1 0 0⟩ orientations. When finding the migration path for an SIA cluster, the same procedure is applied to each SIA in it.

Once the initial and final states (in terms of exact atomic positions) are determined, the NEB method finds the Minimum Energy Path (MEP) between them, starting from an initial guess calculated with a classical drag method [18]. Experience shows that the MEP found is rather sensitive to the following parameters:

(i) The choice of the atoms involved in the drag and NEB routines. In the present calculations, the best strategy was found to involve the most migrating (i.e. ones that show maximum displacement) atoms only, all the others were free of any constraints during the system relaxations. This way, we avoid freezing too many degrees of freedom in the system, allowing for an efficient energy optimization when calculating the transition images.

(ii) The exact definition of the drag algorithm applied to find the initial MEP. The migrating atoms, once dragged, can either be frozen or constrained in a plane perpendicular to the drag direction. The latter choice seems to be inappropriate in general because the 'system' may escape from the unstable states, created while atoms are dragged. Hence, the corresponding MEP may exhibit an unsmooth kinky shape, which is not a physical result but an artefact of the calculations. The best choice seems therefore to fix the migrating atoms and unconstraint all the others while applying the CG method.

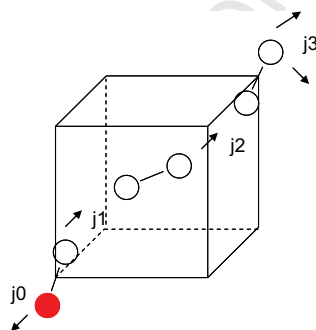**Fig. 1.** Schematic representation of the detrapping of a ⟨1 1 0⟩ SIA from an isolated Cr atom. The following migration barriers are to be overcome: $E_m(j1) = 0.22$ eV, $E_m(j2) = 0.32$ eV, $E_m(j3) = 0.36$ eV. The migration energy of a mixed dumbbell moving via displacement of a Cr atom is $E_m(j0) = 0.21$ eV. A Cr atom is shown by the filled circle.

(iii) The convergence criterion deciding when to stop the NEB iterations. The energy integral under the MEP very quickly reduces within first few iterations but the rate of decrease becomes very slow later on. The decision to stop the calculation is taken if only a small change in the saddle point energy over a significant number of iterations occurs. In this work the change in the saddle point energy should not exceed 0.01 eV over 500 iterations.

The following nomenclature will be used throughout the paper. E_t – absolute value of the total energy of the relaxed atomic configuration. E_m – migration energy of a defect. E_r – energy barrier for an SIA or SIA cluster to perform an on-site rotation from $\langle 1\,1\,0 \rangle$ into $\langle 1\,0\,1 \rangle$ or $\langle 0\,1\,1 \rangle$ configuration. E_u – energy barrier for an SIA or SIA cluster to perform an on-site rotation from $\langle 1\,1\,0 \rangle$ into $\langle 1\,1\,1 \rangle$ configuration. E_b – binding energy of an SIA or SIA cluster to another defect that could be a Cr solute atom, Cr–Cr pair and another SIA defect. E_b was calculated as a difference between the total energy of the crystals containing the two defects located apart from and next to each other. A positive value of the binding energy points at an attractive interaction between two considered defects.

3. Results

3.1. Single SIA

3.1.1. Properties in pure Fe

Following the investigation carried out in [9], where a similar potential [19] to the one used here was applied, the migration mode found for a single SIA in Fe is a translation–rotation jump proposed by Johnson [20]. The corresponding size-independent migration energy calculated here using constant volume relaxation conditions is 0.36 eV, in a close agreement with the DFT value (see introduction). The on-site rotation energy from the $\langle 1\,1\,0 \rangle$ state into $\langle 1\,0\,1 \rangle / \langle 0\,1\,1 \rangle$ or $\langle 1\,1\,1 \rangle$ orientation is $E_r = 0.41$ eV and $E_u = 0.48$ eV, respectively. Since the stability of the $\langle 1\,1\,0 \rangle$ versus $\langle 1\,1\,1 \rangle$ configuration is underestimated with the potential [9], E_u is underestimated as well.

3.1.2. Interaction with an isolated Cr atom

A single SIA is attracted to an isolated Cr atom and forms a mixed Fe–Cr dumbbell with a binding energy of 0.11 eV as already reported in [7]. The optimum migration path for the mixed Fe–Cr dumbbell (ICr) also corresponds to Johnson's mechanism [20] i.e. translation–rotation jump to the first nearest neighbour with $E_m = 0.21$ eV. The latter is only 0.02 eV lower than the DFT calculated value [3]. The second easiest migration mode for the ICr is the Johnson's jump of an Fe atom with $E_m = 0.22$ eV, the migration energy obtained with DFT is 0.07 eV higher [3]. The on-site rotation of the ICr complex into an equivalent $\langle 1\,1\,0 \rangle$ state has a barrier of $E_r = 0.41$ eV, which is higher than the DFT result by 0.05 eV [3]. The 1st nearest neighbour translation jump of Cr (without change of orientation) occurs via passing the $\langle 1\,1\,1 \rangle$ configuration with an energy barrier of 0.4 eV, while the DFT value is 0.42 eV [3]. Considering the Fe–Cr $\langle 1\,1\,0 \rangle$ dumbbell, we see that the energy barriers corresponding to different migration modes are reproduced with the applied potential in a good agreement with available DFT data. The average discrepancy for the migration barrier is about 0.05 eV. The most energetically favourable migration event for the ICr is Johnson's jump via movement of a Cr atom, exactly as suggested by a previous DFT study [3].

The attractive interaction between an isolated solute Cr atom and a $\langle 1\,1\,0 \rangle$ Fe–Fe dumbbell extends up to the 5th nearest neighbour distance (0.495 nm). The dissociation of the SIA from a Cr atom requires at least two successive migration jumps, shown in Fig. 1 as j1 and j2. The corresponding migrations energies for these two jumps are 0.22 eV and 0.32 eV. 3D migration of the ICr via movement of a Cr atom occurs with $E_m = 0.21$ eV. Hence, the ICr defect is expected to diffusive much faster than a Fe–Fe dumbbell and to cover substantial distance before the mixed dumbbell decays.

3.1.3. Interaction with Cr–Cr pairs

A moving Fe–Cr dumbbell may meet another Cr atom or Cr–Cr pair that *a priori* could act as a trap for an SIA defect as suggested in [1,2]. In our previous work [7] a number of SIA-Cr and SIA-Cr–Cr configurations were considered to identify the low energy states that could be associated with traps. The binding energies reported in [7] correspond to the interaction between an Fe–Fe $\langle 1\,1\,0 \rangle$ dumbbell and a Cr or Cr–Cr pair. When recalculating the corresponding binding energy for an Fe–Cr $\langle 1\,1\,0 \rangle$ dumbbell interacting with another Cr atom, we came to the conclusion that the maximum E_b is 0.08 and 0.05 eV according to the DFT and EP, respectively. Using the EP, we searched for other possible Cr arrangements (including configurations containing up to three Cr atoms) surrounding the ICr defect but no configuration with E_b larger than 0.1 eV was found. Thus, we estimate the dissociation energy of an Fe–Cr mixed dumbbell from another Cr atom to be 0.31 eV at most. Hence, a moving Fe–Cr dumbbell cannot be strongly trapped by isolated Cr atoms to form a complex involving a pair of Cr atoms and an SIA (ICr2). The latter will, however, be formed if a freely migrating Fe–Fe dumbbell meets a Cr–Cr close pair, as the biding energy can be as high as 0.19 and 0.15 eV, according to the data obtained with the potential, which is also well agrees with DFT [7]. The most probable pathway for dissociation of the ICr2 complex should occur via the escape of the ICr from the second Cr atom forming a pair. Thus, Cr–Cr close pairs (up to 3rd nn distance) present in the system can be broken apart after interaction with Fe–Fe dumbbells that will drag one Cr atom away. Whereas an isolated Cr atom is a weak trap ($E_b = 0.1$ eV) for an Fe–Cr mixed dumbbell.

3.2. Di-SIA

3.2.1. Properties in pure Fe

The lowest energy configuration for a Di-SIA cluster in alpha Fe is a pair of $\langle 1\,1\,0 \rangle$ dumbbells situated as 1st nearest neighbours, being parallel to each other, in such a way that the vector connecting the center mass of the dumbbells ($\mathbf{R}_{CM}$) is perpendicular to the cluster orientation, as shown in Fig. 2. DFT studies suggest that the lowest MEP is the successive translation–rotation jump of the two dumbbells changing their orientations from [1 1 0] into [1 0 1] configuration [10]. The cluster experiences an intermediate (metastable) state after dumbbell A (see Fig. 2, jump j1) has jumped. Another energetically equivalent jump that the cluster may perform is a [0 1 $\bar{1}$] configuration (the intermediate state for this jump is shown in Fig. 2 by j2). Careful inspection of all possible migration events corresponding to this migration mechanism allows us to formulate three rules determining the final configuration of the cluster after every possible jump and the displacement of the cluster center mass. These are: (i) the orientation of $\mathbf{R}_{CM}$ must be preserved for all ground states that the cluster undergoes while migrating; (ii) at the intermediate state, the two dumbbells must be separated by the distance equal to the third nearest neighbour distance; (iii) the vector corresponding to the orientation of the dumbbells must be orthogonal to $\mathbf{R}_{CM}$ for all stable states (i.e. the dumbbells are first nearest neighbours to each other) that the cluster undergoes while migrating.

For each particular configuration of a Di-SIA cluster only three energetically equivalent configurations (i.e. defining cluster orientation) may occur while the cluster migrates. For the Di-SIA cluster, shown in Fig. 2, they are: [1 1 0], [1 0 1] and [0 1 $\bar{1}$] (here and

4   *D. Terentyev, N. Castin / Computational Materials Science xxx (2009) xxx–xxx*
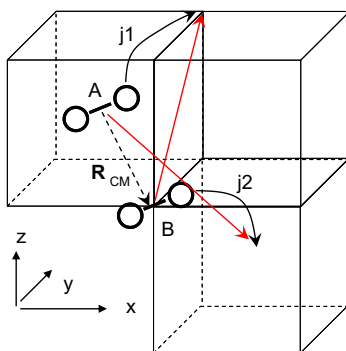

Fig. 2. 'Parallel' configuration of the I2 cluster that can migrate into two energetically equivalent states via successive movement of the two dumbbells following two different paths, namely, j1: dumbbell A migrates into the [1 0 1] orientation in the [1 1 1] direction ($E_m = 0.39$ eV), dumbbell B follows ($E_m = 0.14$ eV). j2: Dumbbell B migrates into the [0 1 1] orientation in the [1 1 $\bar{1}$] direction, dumbbell A follows. A migration jump into a non-equivalent metastable state occurs if dumbbell B makes a jump into the [0 1 1] orientation in the [1 1 1] direction ($E_m = 0.48$ eV), then dumbbell A follows it ($E_m = 0.23$ eV). Axes x, y and z are oriented along [1 0 0], [0 1 0] and [0 0 1] directions, respectively.

Table 1
Activation energy for different migration modes for the I2, I2Cr and I2Cr2 clusters. All values are given in eV. In the column 'successive jump', A and B indicate a particular dumbbell that performs a jump, as shown in Fig. 2.

Initial/final state	Simultaneous jump	Successive jump	On-site rotation
I2			
[1 1 0]/[1 0 1]	0.53	$A = 0.39$, $B = 0.14$	0.65
[1 1 0]/[0 1 1]	0.83	$B = 0.48$, $A = 0.23$	0.78
[1 1 0]/[1 1 1]			0.67
I2Cr			
[1 1 0]/[1 0 1]	0.47 (0.48)[a]	$A = 0.31$, $B = 0.12$	0.49
[1 1 0]/[0 1 1]	0.65 (0.7)[a]	$B = 0.38$, $A = 0.26$	0.68
[1 1 0]/[1 1 1]			Unstable
I2Cr2			
[1 1 0]/[1 0 1]	0.35	Unstable	0.37
[1 1 0]/[0 1 1]	0.47	Unstable	0.51
[1 1 0]/[1 1 1]			0.45

[a] The value in brackets is given for the case when the Cr atom is bound to the B dumbbell.

further we do not mention the mirror jumps resulting in the formation of [$\bar{1}$ $\bar{1}$ 0], [$\bar{1}$ 0 $\bar{1}$] and [0 $\bar{1}$ 1] configurations, since they are identical to the above-mentioned ones. Let us call these configurations A, B and C, respectively. Following the rules, defined above, it is easy to show that the vector corresponding to the displacement of the center mass of the cluster ($\mathbf{V}_{DCM}$) determined by a particular transition can be: [1 1 1] (from A to B and back), [1 1 $\bar{1}$] (from A to C and back) and [$\bar{1}$ 1 1] (from B to C and back). Given that in a perfect bcc lattice the Di-SIA cluster should move by performing uncorrelated jumps (i.e. all transition states realize with the same probability) $\mathbf{V}_{DCM}$ will undergo the three possible displacements: [1 1 1], [1 1 $\bar{1}$] and [$\bar{1}$ 1 1]. Since in the bcc lattice there is no common plain containing all of these vectors, the movement of the I2 (according to the considered mechanism) is not confined in a specific crystallographic plane and thus the I2 should move three-dimensionally via the mechanism explored in [10]. In the latter work, two possible migration paths were suggested, namely: (i) the simultaneous translation–rotation jump of the two dumbbells changing their orientation from [1 1 0] into the [1 0 1] configuration (henceforth this type of jump will be referred to as [1 1 0]/ [1 0 1]); (ii) a sequence of successive [1 1 0]/[1 0 1] jumps. Remarkably, both of these paths result in the same saddle point energy, estimated to be 0.42 eV [10]. The same calculations repeated with the potential from [18] applying constant pressure relaxation revealed that the path for the simultaneous migration event is unstable. The migration energy via two successive jumps was estimated to be 0.31 eV, according to the potential from [18].

Note that the I2 cluster can, in principle, perform the [1 1 0]/ [0 1 1] jump as well. This jump brings the I2 cluster in the non-equivalent metastable energy state (even though the two dumbbells are situated as first nearest neighbours). According to the calculations done here with the EP (at fixed volume) both successive and simultaneous pathways exist for the translation–rotation [1 1 0]/[1 0 1] and [1 1 0]/[0 1 1] jumps. The calculated saddle point energies together with the on-site rotation energies are summarized in Table 1. As one can see, the lowest migration path corresponds to the mechanism suggested in [10] for the sequence of

successive jumps. The estimated energy barrier is, however, lower than the one reported in [10] by 0.03 eV. The above-mentioned second migration mode ([1 1 0]/[0 1 1] jump) occurs with the energy barrier of 0.48 eV. The energy barriers for the simultaneous jumps and on-site rotation movements are distinctively higher. Thus, the migration via the rotation of the I2 into the ⟨1 1 1⟩ configuration, its glide and rotation back into the ⟨1 1 0⟩ configuration is unlikely to occur.

3.2.2. Interaction with an isolated Cr atom

An isolated Cr atom attractively interacts with the I2 cluster and occupies an interstitial position forming a mixed dumbbell attached to the Fe–Fe dumbbell, forming a configuration that we shall call the I2Cr complex. The binding energy is 0.11 eV and 0.2 eV in the ⟨1 1 0⟩ ground and ⟨0 1 1⟩ excited states, respectively. The Cr–I2 interaction range is 0.75 nm. The easiest migration mode for the I2Cr is the [1 1 0]/[1 0 1] jump ($E_m = 0.31$ eV), when the Cr atom is moving. The second lowest MEP (with $E_m = 0.38$ eV) corresponds to the [1 0 1]/[1 1 1] migration jump that again requires movement of the Cr atom. The ⟨1 1 1⟩ configuration of the I2Cr was found to be unstable (the cluster rotates back in the ⟨1 1 0⟩ configuration during CG relaxation) and therefore E_u could not be estimated. Migration barriers for other possible movements such as a simultaneous jump and on-site rotation movement are given in Table 1. The translation–rotation jumps corresponding to the displacement of the Fe–Fe dumbbell at the first place resulted in the unstable configurations.

Using the rules formulated above for the movement of the I2 cluster via [1 1 0]/[1 0 1] type of jumps, it can be shown that the I2Cr will be trapped by a single Cr atom if it would move only via [1 1 0]/[1 0 1] jumps accompanied by the displacement of the Cr atom. A sort of cage effect (well known in the case of face centered cubic lattice) would occur, so that the I2 cluster will perform A–B–A and A–C–A jumps only. Thus, the movement of the I2Cr cluster may occur via a sequence of [1 1 0]/[1 0 1] and [1 1 0]/ [0 1 1] jumps (via displacement of the Cr atom), and the highest migration energy barrier is 0.38 eV (see Table 1). Detrapping of the I2 from a single Cr atom requires a dissociation energy of 0.5 eV, calculated as a sum of the Cr–I2 binding energy and E_m for the I2. We see that the activation energy for the movement of the I2Cr complex is lower than its dissociation energy and hence the I2Cr cluster is expected to be mobile.

3.2.3. Interaction with Cr–Cr pairs

As for a single SIA, the presence of the I2 cluster in the vicinity of a Cr–Cr close pair compensates the Cr–Cr repulsion. Some

D. Terentyev, N. Castin / Computational Materials Science xxx (2009) xxx–xxx

5

important configurations involving the I2 cluster and Cr–Cr pairs are shown in Fig. 3. The configurations 1a–c occupy lower energy states than when a Cr–Cr pair is located far away from the I2 cluster (i.e. configuration 2a). The binding energy between the Cr–Cr pair and the I2 varies from 0.1 up to 0.15 eV, depending on the particular arrangement of Cr atoms in the cluster, see Fig. 3. The most energetically favourable configuration of the I2Cr2 complex is shown in Fig. 3 1c, for which the migration barriers are calculated and reported in Table 1. We see that the I2Cr2 can migrate faster than the I2Cr cluster. Note however, that the breakup of the I2Cr2 complex into I2Cr and an isolated Cr atom is an energetically favourable reaction, as can be seen comparing the total energy of the configurations 2c and 1c shown in Fig. 3. Thus, the freely migrating I2 cluster should break up Cr–Cr close pairs by dragging away one Cr atom. The I2Cr on the other hand should avoid the interaction with other Cr or Cr–Cr pairs, thus their presence is not expected to affect the mobility of the I2Cr complex. However, given that the I2Cr cluster will try to avoid interacting with another Cr atom(s) (to prevent the formation of the I2Cr2), its diffusion path towards a sink or vacancy available for recombination may be longer than in pure Fe.

To sum up the properties of the I2 cluster we may draw the following: (i) I2 exhibits attractive interaction with an isolated Cr and Cr–Cr close pairs ($E_b = 0.1$–0.2 eV); (ii) I2 can break up Cr–Cr close pairs and drag a single Cr forming the stable mobile I2Cr complex (with $E_b = 0.1$ eV); (iii) the migration energy of the latter is $E_m = 0.38$ eV i.e similar to that of the I2 in Fe ($E_m = 0.39$ eV); (iv) the freely migrating I2Cr should not be trapped by other isolated Cr atoms or by Cr–Cr close pairs. In the case of relatively high Cr concentration the movement of the I2Cr, avoiding interaction with Cr, can be significantly reduced; (v) the presence of Cr atom(s) in the I2 strongly reduces both E_r and E_u values.

3.3. Tri-SIA

3.3.1. Mobility in pure Fe

According to the EP, the lowest energy configuration of a tri-SIA cluster (I3) is a triplet of $\langle 1\,1\,0 \rangle$ dumbbells, located in the same

(1 1 0) plane, as shown in Fig. 4a. The dumbbells in this configuration form a triangle with sides oriented along $[\bar{1}\,1\,0]$, $[\bar{1}\,1\,1]$ and $[\bar{1}\,1\,\bar{1}]$ directions. The central dumbbell (top of the triangle in Fig. 4a) connects two dumbbells (side dumbbells) separated by the third nearest neighbour distance. Exactly the same ground state for the parallel I3 clusters was identified by the DFT calculations using both S'STA and VASP codes in [14]. In contrast to this, a linear configuration (all three [1 1 0] dumbbells aligned along $[\bar{1}\,1\,1]$ line) was claimed to be the ground state according to earlier DFT calculations done with the SIESTA code [10,13]. The latter result eventually proved a mistake. The difference in the energy between these two states is about 0.2 eV according to the EP and DFT.

The I3 in the 'linear' configuration can be translated in the equivalent position via the simultaneous or successive (if the intermediate states are metastable) translation–ration jump of all the dumbbells. The I3 in the 'triangle' configuration, on the other hand, does not have an equivalent state if each SIA forming the cluster would perform the [1 1 0]/[1 0 1] or [1 1 0]/[0 1 1] translation–rotation jump simultaneously. The configuration resulting from such a macro jump is not energetically equivalent and not even a metastable one, according to both EP and DFT calculations. Therefore, the I3 in its ground state cannot migrate by performing simultaneous translation–rotation jumps, neither can the barrier for the on-site rotation be calculated. We have performed a set of NEB calculations to estimate the path and corresponding energy barrier for the transition of the I3 cluster from a 'triangle' into a 'linear' configuration. That transition may have an activation energy comparable to the activation energy of the I2 cluster. Among a sequence of jumps (not presented here) the highest saddle point energy was determined to be 0.64 eV. The unfaulting energy was calculated to be 0.67 eV which is comparable with the vacancy migration energy. If one neglects the possible effects of vibration entropy, the movement of the I3 cluster via unfaulting into the $\langle 1\,1\,1 \rangle$ configuration and its glide along $\langle 1\,1\,1 \rangle$ direction should occur at a temperature when vacancies are already mobile. The contribution of the I3 to the stage II recovery therefore needs to be clarified.

Exploring different pathways that would include successive jumps of the dumbbells, we revealed that the movement of the
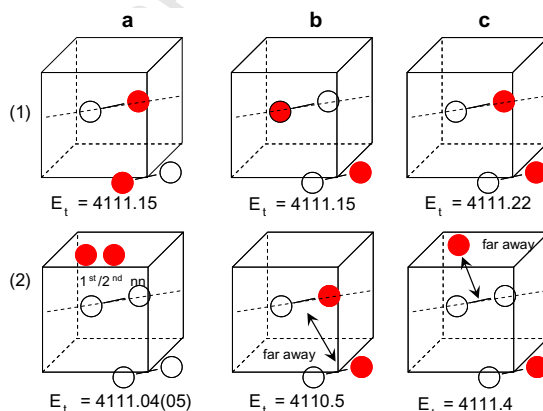

Fig. 3. Different arrangements of the I2 cluster interacting with a Cr–Cr close pair (1a–c), situated far away from the pair (2a), two mixed dumbbells situated far from each other (2b), I2Cr complex situated far from an isolated Cr (2c). The corresponding total energy of the fully relaxed crystal containing 1026 atoms is shown below each configuration. The filled circles represent Cr atoms.
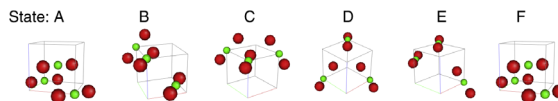
Fig. 4. Sequence of the metastable states that the I3 cluster undergoes while performing the migration jump between the two ground states. Red balls show the relaxed positions of self-interstitial atoms. Green balls show the lattice site positions which the SIAs belong to. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

I3 between the two ground states occurs via a set of independent jumps of the dumbbells. The cluster correspondingly undergoes a transition over a number of metastable states that are shown in Fig. 4. Details of the migration mechanism are described in the following. First, one of the side dumbbells performs the translation–rotation jump in the [1 1 1] direction to occupy a [1 0 1] configuration (see Fig. 4 b), overcoming the barrier of 0.44 eV. Then the other side dumbbell jumps in the [1 1 1] direction and ends up in the [0 1 1] configuration (see Fig. 4c), the corresponding barrier is 0.18 eV. At this point, all dumbbells have different orientations but their central positions lie in the same (0 0 1) plane. The migration of the central dumbbell in the [1 1 1] direction results in the unstable state that decays into the configuration shown in Fig. 4d. There the central positions of the three dumbbells (with different orientations) belong to the same $(1\,1\,\bar{1})$ plane. The migration energy between states C and D, shown in Fig. 4, is extremely small (~0.05 eV) and close to that of the ⟨1 1 1⟩ crowdion in Fe. The next metastable state, shown in Fig. 4e, is obtained by the migration of the side dumbbell. This state can be viewed as the complex made by the I2 and a single [0 1 1] dumbbell located in the tensile region of the cluster. The translation–rotation jump in the [1 1 1] direction completes the migration process and brings the I3 cluster in the ground state. The described 'macro jump' (since it is composed of a set of elementary jumps) between the two energetically equivalent states has a length of 5th nearest neighbour distance.

The migration barriers and excess formation energies for the above described metastable states are shown in Fig. 5. The maximum migration barrier among those corresponding to the elementary jumps is 0.44 eV, while the highest energy state with respect to the ground state is 0.53 eV. None of these values
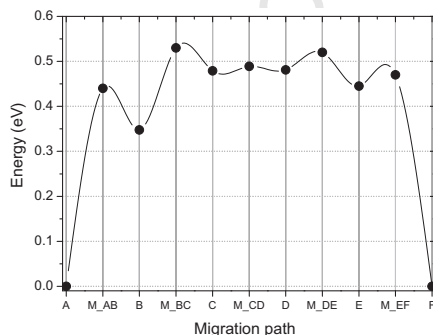
can be attributed to the migration energy of the I3 performing the macrojump. However, the lowest bound for the effective migration energy of the I3 can probably be characterized by the configuration corresponding to the highest energy excess (i.e. saddle point). In general, the assignment of the average time needed to accomplish the macrojump is more appropriate than the concept of the effective migration energy. The average time can be analytically or numerically estimated for a given temperature if the set of migration barriers for all elementary jumps is known. Therefore, one needs to consider the possibility for both forward and backward transitions between the metastable states and correctly account for the effect of vibration entropy contributing to the free energy of the metastable states shown in Fig. 5. Here we do not perform these calculations. We restrict ourselves to stating that in the low temperature range where the vibration entropy contribution is expected to be small, the effective migration energy for the I3 cluster is higher than the migration energy of the I2 cluster (0.39 eV according to the EP, 0.42 according to DFT). Yet, according to the EP, the I3 clusters might contribute to stage II, but the onset of its migration should occur only after the I2 cluster moves, with a shift of ~20–40 K on the temperature scale.

3.3.2. Interaction with an isolated Cr atom and Cr–Cr pairs

The I3 was also found to exhibit attractive interaction with a single Cr atom that enters the cluster by forming a mixed dumbbell. The maximum binding energy is 0.11 eV and it does not depend on the position of Cr in the cluster. Given that the macro jump of the I3 requires two successive translation–rotation jumps of each dumbbell forming the cluster, a Cr atom can not be dragged by the cluster. Therefore, we did not consider the migration energy landscape in the presence of Cr and simply state that Cr should act as a trap for the I3. Correspondingly, the I3Cr complex is stable but immobile. However, we point out that the stability of the I3Cr against unfaulting is enhanced in comparison with the I3. The corresponding rotation energy is 0.75 eV. We also note that the presence of Cr in the I3 cluster does not change the fact that the 'triangle' configuration is more stable than the 'linear' one. The excess formation energy is 0.25 eV, which is even slightly higher than in pure Fe.

As in the case of a single SIA and I2 cluster, the I3 can reduce Cr–Cr repulsion (see Fig. 6 configurations 1a–c vs. 2d). Thus, the I3 can be bound to a Cr–Cr pair with E_b up to 0.3 eV, depending on the particular arrangement of the Cr atoms in the cluster. Note that the gathering of the I2Cr and ICr mobile complexes also results in the formation of the I3Cr2 immobile cluster. The unfaulting energy of the latter is only 0.46 eV. Therefore, a competition between the two following detrapping mechanisms is expected. On the one hand, the cluster can jump away from the Cr–Cr pair via unfaulting and consequent ⟨1 1 1⟩ glide. On the other hand, it can also perform the macro jump already described. Given that the I3Cr complex is immobile, the freely moving I3 cluster can not break up the Cr–Cr close pairs, moving via translation–rotation jumps.

To sum up the most essential features of the I3 cluster we can outline the following: (i) the I3 cluster has distinctively higher
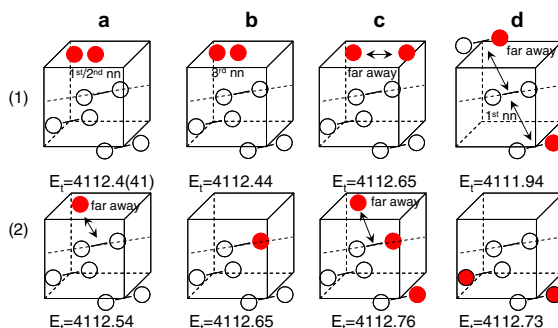

Fig. 5. The energy landscape for the I3 performing the macro jump migrating between the two ground states. M_XY refers to the saddle point energy from state X to Y.

ARTICLE IN PRESS

7

Fig. 6. Different arrangements of the I3 cluster interacting with an single Cr atom and Cr–Cr close pairs. Arrangements shown in figures correspond to the case when: (1a–b) the I3 cluster is situated far away from Cr–Cr close pairs; (1c) the I3 is situated far away from two isolated Cr atoms; (1d) the I2Cr and ICr are situated far away from each other; (2a) the 3I is situated far away from an isolated Cr atom; the I3Cr cluster is formed; (2b) the I3Cr cluster is situated far away from an isolated Cr atom; (2d) the I3Cr2 cluster is formed. The filled circles represent Cr atoms.

migration energy than the I2; (ii) the I3 exhibits attractive interaction with Cr–Cr close pairs and isolated Cr atoms; (iii) the I3 can not drag a single Cr atom or Cr–Cr pair.

4. Summary and conclusive remarks

The stability of the Fe–Cr mixed dumbbell and saddle point energy for different migration modes predicted by the used EP are clearly in line with DFT calculations [3] and with the interpretation given for resistivity recovery experiments in dilute Fe–Cr alloys [1,2], where the shift of stage I_E to a lower temperature has been observed. It is important to note that according to the obtained results, the effective energy characterizing the stage I_E is the migration energy of an Fe–Fe dumbbell, but the position of stage is determined by Cr content and dissociation energy of the ICr–Cr complex ($E_d = 0.31$ eV). The dissociation of the ICr–Cr complexes in Fe–2Cr alloy at ~110K was discussed in [4] and this temperature agrees well with the obtained dissociation energy. The obtained results also point out that the configurations involving Cr–Cr pairs can not essentially hinder fast migration of the Fe–Cr dumbbell and therefore cannot be considered as strong traps, which disagrees with suggestions from [1], implied to explain a global reduction of stage I in Fe–1Cr and Fe–3Cr alloys. The difference between the resistivity retained during stage II at 150–180 K in Fe–1,3Cr alloys and in pure Fe, can not be explained by the weak Cr–ICr interaction reported here. In principle, few recovery stages observed in [1] may occur due to the rearrangement of complexes of self-interstitials interacting with Cr atoms, resulting in the resistivity change rather than SIA-vacancy annihilation. Anyway, prior to draw any conclusion, the possible effect the Cr background on the Cr-SIA binding energy, long-range Cr–Cr interaction, and Cr–ICr detrapping mechanism should be validated by performing accurate DFT calculations.

According to the EP, the I2 cluster also exhibits attractive interaction with a single Cr atom and $E_b = 0.11$ eV is the same as for the I-Cr complex. Whereas the binding energy of the I2Cr2 complex depends essentially on the arrangement of Cr atoms in the cluster. The most striking observation we made is that according to the EP, the I2Cr2 cluster should break up into I2Cr and an isolated Cr atom spontaneously. Based on this result, we have made an essential conclusion stating that the mobility of the I2Cr complex should not be affected by the presence of other Cr atoms (i.e. the I2Cr will

not be trapped by another Cr). To validate the predictions of the EP concerning I2-Cr interaction we will compare them with some available (though not published yet) DFT results taken from [24]. All details on the DFT parameterization and settings used in [24] are exactly the same as in [14], where study of SIA clusters in pure Fe was performed. Briefly, DFT calculations were performed with the with VASP [21], a plane-wave code employing the Projector Augmented Wave (PAW) method [22]. Standard VASP potentials were used, with the Perdew–Wang parameterisation in the GGA. The plane-wave energy cut-off was set to 300 eV and the calculations were performed with spin polarization, the Brillouin zone being sampled by meshes of $3 \times 3 \times 3$ k - points, using the Monkhorst-Pack scheme. The Di-SIA clusters interacting with Cr atoms were relaxed in 252 atom supercells at constant volume.

The obtained DFT results from [24] can be summarized as follows: (i) there is the weak positive binding energy of 0.02 eV for the I2-Cr complex; (ii) there is relatively strong binding energy between the I2 and Cr–Cr close pair, which varies from 0.23 (for configuration shown in Fig. 3 1c) up to 0.32 eV (for configuration shown in Fig. 3 1b); the interaction energy of the ICr and isolated Cr atom is negative (i.e. it is repulsive) for all configurations shown in Fig. 3 row 1, except for the configuration 1b ($E_b = 0.062$ eV). The latter is the only case where the EP incorrectly predicts the sign of the binding energy. However, given that the DFT obtained positive binding energy is extremely small, one may consider the interaction to be negligible. Overall, we see that the EP correctly reproduces the trends in the Cr–Cr-I2 binding energy depending on the arrangement of Cr atoms in the cluster, such as: positive interaction between the I2 and isolated Cr, positive interaction of the I2 with a Cr–Cr close pair and very weak or even repulsive interaction of the I2Cr with an isolated Cr atom. We are therefore confident that the calculations done for the I2 with the EP are qualitatively consistent with the DFT data in most cases.

The reassessment of the migration properties of the I3 cluster in pure Fe, performed here, has revealed new features. The I3 cluster cannot perform a simultaneous jump when it is in its 'triangular' ground state. The obtained migration path suggests that the cluster can move via a series of dumbbell jumps and its lower bound migration energy is higher than that of the I2 cluster by at least 0.1 eV. According to the results obtained with the potential, the I2 and I3 clusters do not start to move simultaneously during annealing, so the recovery curve at the stage II peak should be

8   *D. Terentyev, N. Castin / Computational Materials Science xxx (2009) xxx–xxx*

somewhat more complex than just a single peak. The shape of the recovery peak obtained in pure Fe in [1,2] in the temperature range of 160–200 K is in favour of the above argument. The occurrence of the two distinctive peaks or their partial/complete overlap should depend on the concentration of the I2 and I3 complexes formed up to the beginning of stage II and the concentration of vacant recombination sites and traps for the I2 and I3 clusters (possibly carbon). Clearly, deeper investigations applying DFT and lattice dynamics calculations to estimate the 0 K activation energy for the I3 and the effect of vibration entropy respectively are needed.

In this respect we note that up to this moment two types of DFT calculations on SIA properties in iron exist. These were performed using the SIESTA [10,13] and VASP [11,12] codes. The VASP results are obtained using constant volume calculations whereas the other ones are obtained by relaxing the atomic positions and optimizing the box shape simultaneously, following the so-called 'constant volume per atom' method [23]. Since the strain field of the SIA and SIA clusters is strongly anisotropic, constant pressure calculations should be taken with great care especially when they are performed in small supercells (such as 128 atoms) [9,10,13] for relatively large SIA clusters (size 3 and higher). It is reasonable to expect that the migration energy could be underestimated due to the flexibility of the crystal shape, which compensates high atomic strain localized near the migrating atom(s). The underestimation should be directly related to the crystal size and number of simultaneously migrating defects. At the same time, the 'constant volume' calculations are expected to provide the convergence in a smaller supercell than one needs in fixed volume calculations. Because of the box shape optimization, the effect of self interaction (due to the imposed periodicity of a supercell) can be cancelled out by the release of the lattice strain. However, there is no guarantee that the seemingly converged value may essentially differ if recalculated but without crystal shape optimization. According to the potential, the migration energy for the I2 cluster, calculated using constant volume relaxation, converges when the MD box reaches size of 432 atoms. The higher energy migration path (i.e. the case of the [1 1 0]/[0 1 1] jump) can not be established in a box containing less than 250 atoms in the constant pressure simulations. The migration energy path for the I3 cluster, shown in Fig. 5, becomes independent on crystal size if the number of atoms exceeds 686. Some of the metastable states for the I3 cluster become unstable when calculated in the smaller crystals and/or applying volume optimization.

The migration energy of the I2 cluster according the present calculations is 0.39 eV (for the successive jumps) i.e. it is underestimated by 0.03 eV compared to the value obtained using DFT calculations (0.42 eV [10]). The migration energy of a single SIA is overestimated by 0.02 eV, on the other hand. This interplay cannot be explained directly by the box-size or volume optimization effects mentioned previously. The origin of the discrepancy between DFT and EP calculations is therefore most likely related to the explicit treatment of magnetism, naturally present in DFT and absent in the central force model calculations. For example, an atom migrating according to Johnson's mechanism exhibits a transformation from a ferro- to antiferromagnetic state near the saddle point configuration [10]. This magnetic transition of course affects the total configuration energy of the system and its existence can be put forward to explain why the E_m for the I2 cluster depends essentially on the migration path if calculations are done with the EP and does not depend on the migration path in DFT. Indeed, if the magnetic transition of the atom located in the compressed region results in the decrease of the configuration energy of the system, one understands why there is an essential overestimation of the E_m with the potential when the two dumbbells are displaced simultaneously ($E_m = 0.53$ eV). The clarification of the role of magnetism is highly desirable and in principle can be real-

ized in the framework of DFT by applying certain constraints on the atomic spins while dragging atoms. Even though the discrepancy between the EP and DFT calculated migration energies is very small (0.03 eV) and can be compared with the uncertainty of DFT calculations, the absolute difference between the E_m for a single SIA and I2 cluster is 0.03 eV and 0.09 eV according to the EP and DFT, respectively. The latter discrepancy is essential, since accurate reproduction of the temperature interval between the positions of peaks corresponding to the recovery stages I_E and II in Fe requires a difference in the activation energy of about 0.1 eV [5]. Hence, if one would redo calculations such as done in [13] using the migration barriers from the EP, the position of the stage II would be found at lower temperature and would not coincide with experimental data. Keeping in mind this essential remark, we provide a summary in the following paragraph and then give some implications of the obtained results.

Based on the results presented in Sections 3.1, 3.2 and 3.3 one can summarize the properties of SIAs and their small clusters (I2 and I3) with respect to their mobility and trapping by isolated Cr atoms and Cr–Cr close pairs. The compilation of the obtained numerical data is presented in Table 2. All SIA defects considered here were found to exhibit attractive interaction with isolated Cr atoms and Cr–Cr close pairs. The binding energy of a single Cr atom with either of the defects is 0.11 eV. Single SIA, I2 and I3 clusters can also be bound to a Cr–Cr close pairs with $E_b = 0.2$–0.3 eV, depending on the particular arrangement of solutes in the cluster. However, the I2Cr2 complex is unstable and should break up into the mobile I2Cr complex and isolated Cr atom which is left behind. Therefore, the presence of Cr–Cr close pairs should not affect the mobility of stable I2Cr complexes. On the other hand, Cr–Cr close pairs can be destroyed due to the interaction with freely moving single SIAs and Di-SIA clusters. The migration energy of the ICr and I2Cr complexes, capable of dragging one Cr atom, is 0.21 and 0.38 eV, respectively. Thus, the mobility of the I2Cr is comparable to that of the I2 ($E_m = 0.39$ eV). The mobility of the I3 cluster, not capable of dragging Cr atoms, could be reduced due to the interaction with both isolated Cr atoms and Cr–Cr close pairs. The most important is that the I3Cr and I3Cr2 clusters can be formed in a number of reactions involving I, ICr, I2 and I2Cr mobile complexes. Hence, the concentration of the I3Cr and I3Cr2 clusters should increase drastically during stage II with increasing Cr content in the alloy. Finally, we note that the presence of Cr atom(s) in the I2 and I3 clusters strongly reduces the energy barrier for the on-site rotation into the ⟨1 1 1⟩ configuration. Thus, the presence of Cr may activate another migration mechanism associated with glide along a ⟨1 1 1⟩ direction.

Based on the above-presented summary we can propose possible scenarios for the defect evolution during isochronal annealing performed up to the temperature corresponding to stage III in Fe. For the sake of simplicity and in the spirit of work by Fu et al. [13] and following the conclusions drawn in [14], we assume that SIA clusters containing more than three self-interstitials are immobile at least up to the temperature corresponding to the onset of vacancy migration. Trying to elucidate possible effects of Cr on the deviation of the damage recovery during annealing, we consider the following situations determined by the ratio of the concentration of Cr atoms and Frenkel pairs generated by irradiation, which are:

1. The concentration of radiation-induced Frenkel pairs is much lower than the Cr concentration, so that most of the interstitials are ⟨1 1 0⟩ Fe–Fe dumbbells. Stage I_E will be determined by the migration of Fe–Fe interstitials and should be the same as in pure Fe or slightly shifted towards lower temperature. The main reactions occurring at stage I_E will be: I + Cr → ICr; ICr + I-Cr → I2Cr2 → I2Cr + Cr; ICr + V = Cr. The intensity of stage I_E

ARTICLE IN PRESS

D. Terentyev, N. Castin / Computational Materials Science xxx (2009) xxx–xxx   9

Table 2
Properties of small interstitial clusters with respect to their mobility and trapping by isolated Cr atoms and Cr–Cr close pairs calculated in this work using the EP. Available DFT data is given in brackets.

Specie(s)	Name	Migration (E_m), binding (E_b) and dissociation (E_d) energy
Fe–Fe dumbbell	I	E_m = 0.36 (0.34–0.37 [10,11]) eV
Clusters of parallel Fe–Fe dumbbells	I2, I3	E_m(I2) = 0.39 (0.43 [10]) eV; E_m(I3) $\geqslant$ 0.53 eV
Mixed FeCr dumbbell	ICr	E_m = 0.21 (0.23 [3]) eV; E_b = 0.11 (0.08 [7]) eV; E_d = 0.47 (0.42, given that E_b = 0.08 and E_m = 0.34) eV
Mixed Fe–Cr dumbbell trapped at another neighbouring Cr atom (not entering the dumbbell)	ICr2	Immobile, E_b(ICr–Cr) = 0.1 (0.08 eV [7]) eV; E_b(I-Cr2) = 0.19 (0.15 [7]) eV
		Dissociation reaction occurs via the release of the ICr complex, E_d = 0.31 (0.31) eV
Di-SIA formed by Fe–Fe and Fe–Cr dumbbells	I2Cr	E_m = 0.38 eV; E_b(I2-Cr) = 0.11 (0.02 [24]) eV; E_d = 0.5 (0.45, given that E_b = 0.02 [24] + E_m = 0.43 [13]) eV.
Di-SIA formed by two Fe–Cr dumbbells	I2Cr2	Immobile, E_b(I2-Cr2) = 0.1–0.2 (0.23–0.32 [24]) eV; E_b(I2 Cr–Cr) = from −0.43 to −0.1 (−0.495–+0.062 [24]) eV
		Dissociation results in the formation of the I2Cr and isolated Cr atom, E_d = 0.21 eV
Tri-SIA formed by two Fe–Fe and one Fe–Cr dumbbells	I3Cr	Immobile, E_b(I3-Cr) = 0.11 eV, E_d = 0.63 eV
Tri-SIA formed by two Fe–Cr and one Fe–Fe dumbbells	I3Cr2	Immobile, E_b(I3-Cr2) = 0.2–0.3 eV (depending on the position of Cr in the cluster), E_d = 0.73–0.83 eV; Dissociation occurs via the release of the I3.

should be governed by the dissociation of ICr–Cr complexes but redistribution of Cr atoms may also occur during this stage. The onset of stage II will be mainly determined by the movement of the I2Cr complexes and therefore there should not differ from that in Fe, since the migration energy for the I2 and I2Cr are comparable, according to the results obtained here. The intensity of stage II and the total defect recovery up to stage III will depend on the ratio of the concentrations of the I3 and I3Cr (I3Cr2) clusters formed during stage II. If the fraction of the I3Cr complexes is substantial, a visible decrease of the intensity of the stage should be observed.

2. The fraction of Fe–Cr mixed dumbbells that were formed after irradiation and that have survived the spontaneous and correlated recombination stages is significant or even overwhelming compared to a number of Fe–Fe dumbbells. Since the migration energy for the ICr complex to moves is much lower than for a $\langle 1\,1\,0 \rangle$ Fe–Fe dumbbell the position of stage I_E should shift significantly towards lower temperature or even merge with the correlated recombination stage I_D. Again, the main complex surviving up to the onset of stage II will be the I2Cr cluster and hence the onset of stage II should not be affected. This time, however, most of the I3 and larger SIA clusters will contain Cr atoms. So that, the intensity of stage II should be strongly reduced.

Regarding the available experimental studies of dilute FeCr alloys, it seems that the alloys studied in [2] suit the first scenario suggested above, since both a progressive shift of stage I_E and a decrease of the recovery (by ~30%) at stage II were observed in the alloys (with $\langle 0.1\%$Cr) as compared to pure Fe. Whereas the recovery curves obtained in irradiated Fe-1%Cr and -3%Cr in [1] can be rationalized within by the second case scenario. In a forthcoming work, we will use kinetic Monte Carlo and rate equation modeling tools to study the defect recovery corresponding to irradiation and annealing conditions from [1,2], and try to reveal the range of Frenkel pair and Cr concentration wherein both suggested scenarios can be distinctively realized.

Acknowledgements

References

[1] F. Maury, P. Lucasson, A. Lucasson, F. Faudot, J. Bigot, J. Phys. F: Met. Phys. 17 (1987) 1143.
[2] H. Abe, E. Kuramoto, J. Nucl. Mater. 271&272 (1999) 209–213.
[3] P. Olsson, J. Nucl. Mater. 386–388 (2009) 86–89.
[4] A.L. Nikolaev, V.L. Arbuzov, A.E. Davletshin, J. Phys.: Condens. Matter 11 (1999) 8633–8644;
A.L. Nikolaev, Phil. Mag. 87 (2007) 4847–4874.
[5] S. Takaki, J. Fuss, H. Kugler, U. Dedek, H. Schultz, Radiat. Eff. 79 (1983) 87.
[6] D. Terentyev, P. Olsson, L. Malerba, J. Nucl. Mater. 386–388 (2009) 140–142.
[7] D. Terentyev, P. Olsson, T.P.C. Klaver, L. Malerba, Comput. Mater. Sci. 43 (2008) 1183–1192.
[8] P. Olsson, J. Wallenius, C. Domain, K. Nordlund, L. Malerba, Phys. Rev. B 72 (2005) 214119.
[9] F. Willaime, C.-C. Fu, M.C. Marinica, J. Della Torre, Nucl. Instr. Meth. Phys. Res. B 228 (2005) 92–99.
[10] C.C. Fu, F. Willaime, P. Ordejon, Phys. Rev. Lett. 92 (2004) 175–503.
[11] E. Vincent, C.S. Becquart, C. Domain, J. Nucl. Mater. 359 (2006) 227.
[12] P. Olsson, C. Domain, J. Wallenius, Phys. Rev. B 75 (2007) 014110.
[13] C.C. Fu, J. Dalla Torre, F. Willaime, J.L. Bocquet, A. Barbu, Nat. Mater. 4 (2005) 68.
[14] D.A. Terentyev, T.P.C. Klaver, P. Olsson, et al., Phys. Rev. Lett. 100 (2008) 145503.
[15] G.J. Ackland, M.I. Mendelev, D.J. Srolovitz, et al., J. Phys.: Condens. Matter 16 (2004) S2629.
[16] G. Henkelman, H. Jonsson, J. Chem. Phys. 113 (2000) 9978.
[17] M.C. Payne, M.P. Teter, D.C. Allan, T.A. Arias, J.D. Joannopoulos, Rev. Mod. Phys. 64 (1992) 1045.
[18] G. Henkelman, G. Johannensson, in: S.D. Swartz (Ed.), Progress on Theoretical Chemistry and Physics, Kluwer Academic Publishers, 2000, pp. 269–300.
[19] M.I. Mendelev, S.W. Han, D.J. Srolovitz, G.J. Ackland, D.Y. Sun, M. Asta, Phil. Mag. A 83 (2003) 3977.
[20] R.A. Johnson, Phys. Rev. 134 (1964) A1329.
[21] G. Kresse, J. Hafner, Phys. Rev. B 47 (1993) 558;
G. Kresse, J. Furthmuller, Phys. Rev. B 54 (1996) 11169.
[22] G. Kresse, D. Joubert, Phys. Rev. B 59 (1999) 1758.
[23] F. Willaime, J. Nucl. Mater. 323 (2003) 205.
[24] P. Klaver, P. Olsson, in preparation.   Q1