Ranking the Stars: A Refined Pareto Approach to Computational Materials Design

Kurt Lejaeghere,[1] Stefaan Cottenier,[1,2] and Veronique Van Speybroeck[1,*]

[1]Center for Molecular Modeling, Ghent University, Technologiepark 903, BE-9052 Zwijnaarde, Belgium

[2]Department of Materials Science and Engineering, Ghent University, Technologiepark 903, BE-9052 Zwijnaarde, Belgium

(Received 2 April 2013; published 14 August 2013)

We propose a procedure to rank the most interesting solutions from high-throughput materials design studies. Such a tool is becoming indispensable due to the growing size of computational screening studies and the large number of criteria involved in realistic materials design. As a proof of principle, the binary tungsten alloys are screened for both large-weight and high-impact materials, as well as for fusion reactor applications. Moreover, the concept is generally applicable to any design problem where multiple competing criteria have to be optimized.

DOI: 10.1103/PhysRevLett.111.075501 PACS numbers: 81.05.Zx, 07.05.Kf, 28.52.Fa, 62.20.−x

In recent years, materials design has benefitted considerably from computational searches [1]. The increase in computing power has allowed the screening of large numbers of hypothetical compounds for several well-defined design criteria. However, interpreting the output from such investigations is not a trivial task, especially when multiple objectives are optimized simultaneously. The field of multicriteria decision making [2,3] proposes the use of the so-called Pareto set to drastically reduce the number of materials under consideration. Multicriteria decision making notes that in the case of competing requirements, a single "best" solution does not exist. Instead, a number of solutions can be shown to outperform the rest, forming the Pareto set $\mathcal{P}$. Compared to such a Pareto solution $\tilde{\boldsymbol{x}}$, none of the alternative $\boldsymbol{x}$ improve all of the decision criteria $f^{(k)}$ simultaneously:

$$\tilde{\boldsymbol{x}} \in \mathcal{P} \Leftrightarrow \nexists\, \boldsymbol{x}\colon \begin{cases} f^{(k)} \geq \tilde{f}^{(k)} & \forall\ k = 1, \ldots, N \\ \exists\ k_0\colon f^{(k_0)} > \tilde{f}^{(k_0)}. \end{cases} \tag{1}$$

Here, we assume the design problem to be described by N normalized, maximizable objective functions $f^{(k)}$ with $k = 1, \ldots, N$, relating to each Pareto solution a set of coordinates. The assembled Pareto points are therefore also called the Pareto front or skyline, as they outline a hypersurface in N-dimensional space (full black line in Fig. 1).

Pareto optimality has already been successfully applied to computational materials design [4,5]. Unfortunately, as the dimensionality of the problem increases, so does the size of the Pareto set $\mathcal{P}$. It then becomes prohibitively time consuming to study every single Pareto compound in more detail. Several post-Pareto analysis methods therefore try to reduce the number of candidates even further [6–8], but none of them offer a quantitative ordering. This Letter, on the other hand, proposes a mathematically founded procedure that allows the *ranking* of the Pareto compounds, thus identifying the most optimal compromises with respect to the design requirements. It is, however, also more generally applicable to any multicriterion design problem, ranging from space technology [9] over urban studies [10] to linguistics [11].

Figure 1 illustrates how a ranking can be based on the tradeoff between two Pareto solutions, using a hypothetical 2D data set. If we consider an arbitrary Pareto point, such as b, its bottom left quadrant (hatched area) contains only suboptimal (dominated) points, while solutions outside this region offer a tradeoff with the properties of b, either because they improve X (e.g., c) or because they improve Y (e.g., a). The beneficial part of this tradeoff increases as we move away from the dominated quadrant, corresponding to a larger angle θ. When $\theta = 45°$, none of the points is particularly preferred: every loss ΔY is now compensated by an equivalent gain in ΔX, or vice versa. For angles

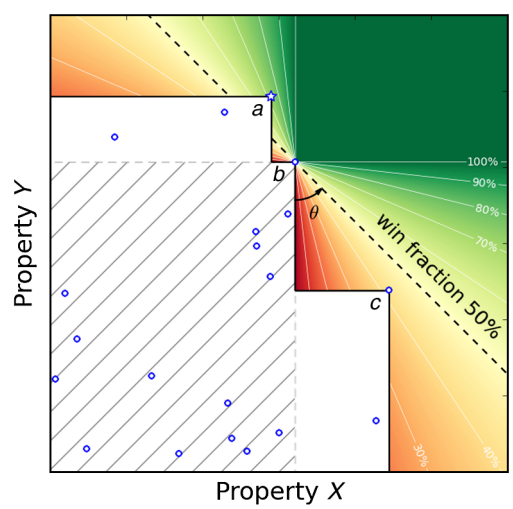

FIG. 1 (color online). Win fraction [Eq. (2)] with respect to Pareto solution b (wf_b), ranging from 0% (red, worst tradeoff) to 100% (green, best tradeoff). The quadrant dominated by point b has been hatched. At a deviation $\theta = 45°$ from the skyline, all solutions are equally valuable (50% win fraction). For the given two-dimensional Pareto front, however, a outranks b and b outranks c.

 © 2013 American Physical Society

larger than 45°, on the other hand, the tradeoff favors the new data point. When applied to Fig. 1, this means that a outranks b, while b outranks c.

To distinguish between two Pareto solutions, we therefore define the win fraction (wf) of $\tilde{\boldsymbol{x}}_j$ with respect to $\tilde{\boldsymbol{x}}_i$:

$$\mathrm{wf}_i(j) = \frac{\sum_k |\tilde{f}_j^{(k)} - \tilde{f}_i^{(k)}| \times \mathcal{H}(\tilde{f}_j^{(k)} - \tilde{f}_i^{(k)})}{\sum_k |\tilde{f}_j^{(k)} - \tilde{f}_i^{(k)}|}, \qquad (2)$$

where $\mathcal{H}(t)$ stands for the Heaviside step function (1 for $t > 0$ and 0 for $t < 0$). The denominator of Eq. (2) represents the total tradeoff between the two Pareto solutions, while in the nominator, only the positive terms remain. The win fraction $\mathrm{wf}_i(j)$ hence expresses what part of the total tradeoff between $\tilde{\boldsymbol{x}}_i$ and $\tilde{\boldsymbol{x}}_j$ favors $\tilde{\boldsymbol{x}}_j$. Since each part of the tradeoff favors one of both solutions, $\mathrm{wf}_i(j)$ and $\mathrm{wf}_j(i)$ always sum to 100%. In 2D, the win fraction is directly related to the skyline angle θ, as can be seen from the win-fraction contour lines in Fig. 1.

A ranking factor for a given solution $\tilde{\boldsymbol{x}}_j$ now follows from the minimum of the win fraction (mwf) with respect to all other Pareto compounds $\tilde{\boldsymbol{x}}_i$: $\mathrm{mwf}(j) = \min_i[\mathrm{wf}_i(j)]$. A large mwf value indicates that $\tilde{\boldsymbol{x}}_j$ provides a considerably better tradeoff than any other Pareto solution. In this way, an mwf-based ordering identifies the most striking features of the skyline.

We first apply this ranking criterion to two classic engineering problems, as a proof of principle: the design of large-weight and of high-impact materials. In both examples, the same search space is investigated, containing 285 ordered body-centered cubic (bcc) binary tungsten alloys with doping levels of 50%, 25%, 12.5%, and 6.25%. This includes all possible alloying elements up to Rn (without the lanthanides) and pure W. For each of these compounds, a case-specific set of properties was calculated by means of density-functional theory [12], using the projector augmented-wave formalism [13,14] in VASP [15,16]. All calculations were performed with the Perdew-Burke-Ernzerhof exchange-correlation functional [17]. Before the Pareto analysis, the value ranges of all properties were normalized to the corresponding range of the elemental crystals [18] and scaled with a weight factor. This provides the mwf the flexibility to express the relative importance of each objective, if user-specific preferences require so. In the examples discussed here, all weights were set to unity (equally important objectives). A summary of further computational settings and the calculated properties is given in the Supplemental Material [19].

As a first example, consider the search for an inexpensive weight, useful as a counterbalance, a press, or a wrecking ball. Acceptable candidate compounds should then combine a large mass density with a good price. The mass density (as well as the bulk modulus in a later example) was extracted from a Birch-Murnaghan fit [20] to a 13-point equation of state (between $V_0 \pm 6\%$), while price data were obtained from the Chemicool Web site, based on prices from Goodfellow [21]. The corresponding skyline is displayed in Fig. 2 (top). In addition, a buffer zone of one error bar is indicated by a dashed line to show the influence of computational uncertainties [18] (see the Supplemental Material [19]). By considering this buffer zone, we avoid overlooking compounds that belong to the *experimental* Pareto set, yet—due to computational inaccuracies—not to the *computational* Pareto set.

In Fig. 2, pure W intuitively appears as the most prominent Pareto solution: increasing the mass density comes at a tremendous cost (a 35-fold price increase), while relaxing the constraints on the mass density only marginally improves the price. Indeed, this also appears from its mwf (73%), which overshadows all other Pareto compounds. When the price of tungsten is considered to be too high, however, alloying offers a solution. The minimum win fractions of WPb (27%), W_3Pb (25%), W_7Pb (24%), and $W_{15}Pb$ (21%) are ranked just below pure W. When a mass density increase is desired, on the other hand, WIr is the highest-ranked alloy in the list, with an mwf of 10%.

Even when the search space is extended with 71 elemental crystals [18], tungsten remains the most interesting Pareto solution, with an mwf of 54%. It is now directly

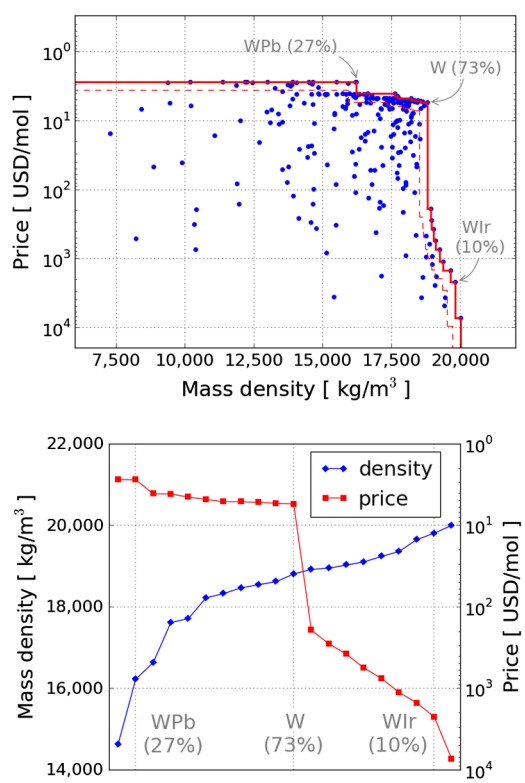

FIG. 2 (color online). Pareto analysis of the bcc binary tungsten alloys with respect to the mass density and the molar price. Top: skyline plot with a one-error-bar uncertainty zone (dashed line). Bottom: evolution of the objectives within the Pareto set. Three high-ranked solutions are indicated with their minimum win fractions.

followed by pure lead (46%) and pure iron (43%). Both are commonly used materials for large-weight applications, favored over tungsten because of their considerably lower price. They illustrate that an mwf-based ranking indeed identifies the most promising options.

In a second test case, we focus on the hardness and temperature resistance of tungsten. Although W itself already performs very well for these properties, using alloying elements might nevertheless yield an overall improvement. We therefore screen our search space for a good bulk modulus (hardness), cohesive energy (thermal resistance), and price. Materials that meet all of these objectives can be used in drill bits or armor-piercing projectiles, for example.

A 3D representation of the skyline (left-hand panel of Fig. 3) allows us to discern the two most prominent Pareto solutions, WC (in green) and pure W (in red). The minimum win fractions correctly suggest them as highest-ranked materials, with mwf values of 56% and 44%, respectively. The third Pareto compound WRe already performs much worse, with an mwf of only 13%. WC and W are indeed used commercially. Tungsten carbide is the harder material, both in its metastable cubic structure and in the ground-state hexagonal phase [22], while tungsten itself is more temperature resistant. Another commonly employed material, iron, enters the Pareto set at an mwf of 32% when we extend the search space with 71 elemental crystals.

These two test cases demonstrate that our procedure yields meaningful results for well-known examples. A last case study, in the domain of nuclear fusion research, shows its strength in tackling cutting-edge design problems. Because of the harsh operating conditions inside fusion reactors, materials development is a critical aspect there [23,24]. This particularly holds for the in-vessel components, which directly come into contact with the plasma. The divertor, for example, filters ions from the plasma but only by withstanding huge heat and particle fluxes [25]. Most designs for it are based on tungsten because of the associated high-temperature resistance and low activation [26]. Unfortunately, the ductile-to-brittle transition temperature of tungsten is too high for structural applications: in the low-temperature range of operating conditions, tungsten is in the brittle regime. This has initiated an ongoing search for ductilizing alloying elements [25].

Our test set of binary W alloys can contribute to this screening of potential divertor compounds. A suitable divertor material needs to improve the ductility of tungsten while maintaining its favorable high-temperature resistance and acceptable price. Evaluating these three objectives for each of the tungsten alloys may therefore provide the fusion community with some attractive candidate reactor materials.

Since at a defect-free atomic scale, ductility is an intrinsic property of the material, we use the Cauchy pressure P_C to estimate it. Large values of P_C indicate large nonpairwise contributions to the interatomic forces, suggesting an isotropic material [27]. Indeed, directional bonds hinder shear and result in brittle fracture more easily, while isotropic bonds allow for a certain plasticity. For cubic crystals, P_C equals the elastic constant difference $C_{12} - C_{44}$. In our W alloy test set, these elastic constants were extracted from the relation between a set of 2% strains and the corresponding stresses.

Contrary to the two previous test cases, we only consider the mechanically stable structures here. Mechanical stability, as prescribed by the Born criteria [28], is governed by the elastic parameters, and crystals that do not meet these requirements produce highly unrealistic Cauchy pressures.

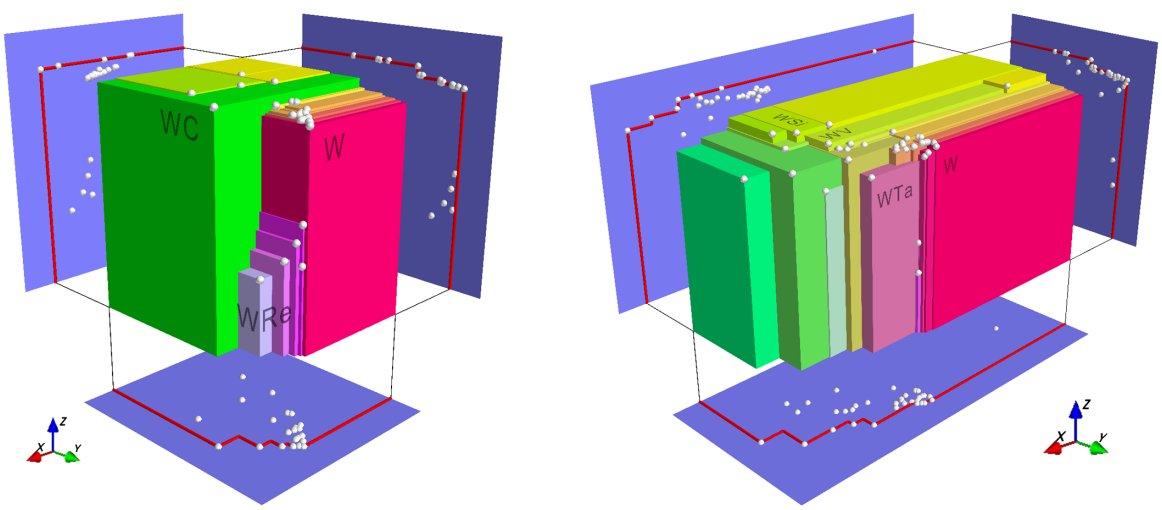

FIG. 3 (color online). Skyline plot and its projection onto the three coordinate planes [33] for the bcc binary tungsten alloys. Left: with respect to the bulk modulus (x), cohesive energy (y), and logarithmic price ($-z$). Right: with respect to the Cauchy pressure (x), cohesive energy (y), and logarithmic price ($-z$).

As this affects the Pareto front substantially, it is best to exclude them. An overview of the stability of the tungsten alloys is given in the Supplemental Material [19].

A Pareto analysis of the three objectives—Cauchy pressure, cohesive energy, and price—retains 36 possible solutions (Table I). They are depicted in Fig. 3 (right). If we moreover require the Pareto solutions to be energetically stable (with respect to phase segregation), this number can be reduced even further. Energetic stability mainly refers to the formation energy. When the formation energy exceeds several times the estimated error on the computed value (0.15 eV/at, Ref. [18]), it can safely be concluded that this compound is unlikely to be (meta)stable. This is the case for W_7C, W_3P, and $W_{15}C$, for example, which can therefore be neglected in the Pareto set of Table I. *A posteriori* analyses of the Pareto set are not necessarily limited to density-functional theory quantities. Other design requirements, such as avoiding radioactive activation, could be taken into account in a similar fashion.

The mwf procedure does not rank pure tungsten highest. At 22%, WTa provides a better tradeoff, yielding an improved Cauchy pressure (128 GPa versus 64 GPa), while the cohesive energy is almost the same (differing by only 30 kJ/mol). In addition, other doping levels have similar effects. For three of the four studied concentrations, the W-Ta system is Pareto efficient, and $W_{15}Ta$ lies only 6% of an error bar from the Pareto front. All W-Ta alloys are mechanically and energetically stable, even at 0 K [19]. Tantalum is three times more expensive than tungsten, however.

Table I also suggests vanadium as an interesting possibility. Although the cohesive energy deteriorates by up to 140 kJ/mol, combining tungsten with high concentrations of vanadium improves its intrinsic ductility even further than with tantalum ($P_C = 138$ GPa for WV). In addition, using vanadium lowers the price significantly, a 75% cut compared to W. All four W-V crystals belong to the Pareto set and are energetically and mechanically stable at 0 K [19].

Contrary to Ta and V, other high-mwf alloying elements yield ambiguous results. Silicon-doped alloys, for example, all belong to the Pareto set and alter the behavior of pure tungsten in the same way as vanadium. Unfortunately, they have quite high formation energies. The magnetic 3*d* transition metals rank well, too. Except for chromium, they all give rise to intrinsically less brittle metals. Mechanical and energetic stability is sometimes an issue for these materials, however. Rhenium, which is experimentally found to lower the ductile-to-brittle transition temperature of tungsten [29], does not appear from this Pareto analysis. Our calculations indeed find an improved Cauchy pressure but suggest that the same improvement can be achieved at better tradeoff conditions.

W-Ta and W-V compounds were investigated theoretically before, but based on the intrinsic ductility alone, their promising performance was not observable [30]. Other studies found Ta and V to embrittle W [25,31]. This discrepancy between the intrinsic and macroscopic ductility of tungsten alloys remains to be elucidated, especially in contrast to observations for other materials systems (e.g., for stainless steel [32]).

TABLE I. Pareto set for the bcc binary tungsten alloys with respect to the Cauchy pressure, cohesive energy, and price. The Pareto compounds are ranked according to the mwf.

mwf		mwf		mwf	
58.1%	W_7C	10.2%	$W_{15}V$	2.5%	W_7Ni
41.9%	W_3P	9.0%	W_7V	2.4%	W_7Ge
22.1%	WTa	8.6%	WV	2.4%	W_3Fe
19.0%	W	8.0%	WFe	2.1%	W_3V
18.9%	W_7Ta	7.1%	$W_{15}Si$	2.0%	W_7Os
16.4%	W_3Ta	3.6%	W_3Si	1.8%	$W_{15}Cr$
14.8%	$W_{15}Co$	3.5%	W_3Sb	0.8%	W_7Cr
14.6%	$W_{15}Ni$	3.4%	$W_{15}Fe$	0.7%	$W_{15}O$
14.2%	$W_{15}C$	3.4%	W_7Fe	0.4%	$W_{15}Rh$
13.9%	W_7Si	3.1%	WC	0.4%	W_3Co
10.7%	WSi	2.8%	$W_{15}P$	0.3%	$W_{15}N$
10.6%	WMn	2.5%	W_3Ni	0.0%	W_7O

All three test cases show that the minimum win fraction orders the Pareto set in an intuitive way. Contrary to conventional Pareto approaches, optimal designs can be compared to each other and ranked, drastically reducing the number of candidates that require closer inspection. Using materials design as an example, the procedure has been shown to yield meaningful results for both classic engineering problems and state-of-the-art applications. It can hence serve as an enabler for discoveries and developments in many research areas.

This work is supported by the Fund for Scientific Research—Flanders (FWO) and by the Research Board of Ghent University. S. C. acknowledges financial support from OCAS NV by an OCAS-endowed chair at Ghent University. Calculations were performed using the Stevin Supercomputer Infrastructure at Ghent University, funded by Ghent University, the Hercules Foundation, and the Flemish Government (EWI Department).

*Veronique.VanSpeybroeck@UGent.be

[1] S. Curtarolo, G. L. W. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, and O. Levy, Nat. Mater. **12**, 191 (2013).
[2] R. L. Keeney and H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Trade-Offs* (Cambridge University Press, Cambridge, England, 1993).
[3] K. M. Miettinen, *Nonlinear Multiobjective Optimization*, International Series in Operations Research and Management Science (Kluwer Academic, Boston, 1999).
[4] T. Bligaard, G. H. Jóhannesson, A. V. Ruban, H. L. Skriver, K. W. Jacobsen, and J. K. Nørskov, Appl. Phys. Lett. **83**, 4527 (2003).

[5] J. Greeley, T. F. Jaramillo, J. Bonde, I. Chorkendorff, and J. K. Nørskov, Nat. Mater. **5**, 909 (2006).
[6] I. Das, Struct. Optim. **18**, 30 (1999).
[7] G. K. Kao and S. H. Jacobson, Comput. Optim. Applic. **40**, 73 (2008).
[8] N. Lopez, O. Aguirre, J. F. Espiritu, and H. A. Taboada, in *Proceedings of the 41st International Conference on Computers & Industrial Engineering, Los Angeles, California, 2011*, edited by M. I. Dessouky, http://www.usc.edu/dept/ise/caie/34th/conference/website/documents/cover.html.
[9] N. Assadian and S. H. Pourtakdoust, Adv. Space Res. **45**, 398 (2010).
[10] S. Liu, R. Tao, and C. M. Tam, Habitat Int. **37**, 155 (2013).
[11] M. Choudhury, V. Jalan, S. Sarkar, and A. Basu, *Proceedings of Ninth Meeting of the ACL Special Interest Group in Computational Morphology and Phonology*, edited by J. Nerbonne, M. T. Ellison, and G. Kondrak (Association for Computational Linguistics, Prague, Czech Republic, 2007), pp. 65–74.
[12] R. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, England, 2004).
[13] P. E. Blöchl, Phys. Rev. B **50**, 17 953 (1994).
[14] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).
[15] G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).
[16] J. Hafner, J. Comput. Chem. **29**, 2044 (2008).
[17] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[18] K. Lejaeghere, V. Van Speybroeck, G. Van Oost, and S. Cottenier, arXiv:1204.2733 [Crit. Rev. Solid State Mater. Sci. (to be published)].
[19] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.111.075501 for additional computational details and a full overview of the calculated values.
[20] F. Birch, Phys. Rev. **71**, 809 (1947).
[21] Chemicool Periodic Table (February 2012), http://www.chemicool.com; Goodfellow Group, http://www.goodfellow.com.
[22] D. V. Suetin, I. R. Shein, and A. L. Ivanovskii, Phys. Status Solidi B **245**, 1590 (2008).
[23] A.-A. F. Tavassoli, J. Nucl. Mater. **302**, 73 (2002).
[24] I. Cook, Nat. Mater. **5**, 77 (2006).
[25] M. Rieth *et al.*, J. Nucl. Mater. **432**, 482 (2013).
[26] G. A. Cottrell, R. Pampin, and N. P. Taylor, Fusion Sci. Technol. **50**, 89 (2006).
[27] D. G. Pettifor, Mater. Sci. Technol. **8**, 345 (1992).
[28] M. Born, J. Chem. Phys. **7**, 591 (1939).
[29] G. A. Geach and J. E. Hughes, in *Plansee Seminar Proceedings*, edited by F. Benesovsky (Pergamon, London, 1955).
[30] D. Nguyen-Manh, M. Muzyk, K. J. Kurzydłowski, N. L. Baluc, M. Rieth, and S. L. Dudarev, Key Eng. Mater. **465**, 15 (2011).
[31] H. Li, S. Wurster, C. Motz, L. Romaner, C. Ambrosch-Draxl, and R. Pippan, Acta Mater. **60**, 748 (2012).
[32] L. Vitos, P. A. Korzhavyi, and B. Johansson, Nat. Mater. **2**, 25 (2003).
[33] P. Ramachandran and G. Varoquaux, IEEE Comput. Sci. Eng. **13**, 40 (2011).