Entropy based fingerprint for local crystalline order

Pablo M. Piaggi\textsuperscript{1,2} and Michele Parrinello\textsuperscript{3,2,*}
\textsuperscript{1) Theory and Simulation of Materials (THEOS), Ecole Polytechnique Fédérale de Lausanne, c/o USI Campus, Via Giuseppe Buffi 13, CH-6900, Lugano, Switzerland
\textsuperscript{2) Facoltà di Informatica, Instituto di Scienze Computazionali, and National Center for Computational Design and Discovery of Novel Materials MARVEL, Università della Svizzera italiana (USI), Via Giuseppe Buffi 13, CH-6900, Lugano, Switzerland
\textsuperscript{3) Department of Chemistry and Applied Biosciences, ETH Zurich, c/o USI Campus, Via Giuseppe Buffi 13, CH-6900, Lugano, Switzerland

(Dated: 1 August 2017)

We introduce a new fingerprint that allows distinguishing between liquid-like and solid-like atomic environments. This fingerprint is based on an approximate expression for the entropy projected on individual atoms. When combined with a local enthalpy, this fingerprint acquires an even finer resolution and it is capable of discriminating between different crystal structures.

I. INTRODUCTION

Atomic computer simulation is an important technique used in the study of a broad range of phenomena in materials science, chemistry, and condensed matter physics. In these fields, very often one is faced with the problem of identifying different local arrangements. A paradigmatic case is that of the nucleation of a crystal from the liquid where one is required to distinguish between solid-like and liquid-like atomic environments. The situation is even more complicated in systems exhibiting polymorphism since in these cases it is desirable to classify the atoms as belonging to one of the different polymorphic structures. This is a common occurrence in nucleation studies where Ostwald’s step rule is observed or where clusters exhibit a core-shell structure. Another area where the ability to distinguish between different local arrangements plays a role is in the identification of crystallites in nanocrystalline material.

Several methods have been proposed to distinguish between liquid-like and solid-like atoms and to identify local crystalline structures. One such method is the common neighbor analysis (CNA) which is an efficient algorithm able to distinguish between liquid, bcc, fcc, and hcp phases. However, it lacks robustness with respect to particle displacements such as those arising from thermal motion or stresses. Another popular method is based on the local Steinhardt parameters which are local, averaged versions of the original Steinhardt parameter. However, they also come at a high computational cost and presume that the nature of the crystal structure is known beforehand.

This work is inspired by a recent progress in the study of nucleation using metadynamics to enhance the probability of inducing the crystal formation in an accessible computer time. Metadynamics relies on the identification of appropriate collective variables (CVs). In Ref. we found that enthalpy and an approximate expression for entropy based on the two body correlation function, were useful CVs in this context. One of the features of this work was that the CVs did not contain any information on the geometry of the crystal structure. This suggested that perhaps from these two quantities one could extract fingerprints able to distinguish between different local atomic arrangements.

Enthalpy and entropy are global properties and in order to be able to use them as local parameters we have to project them onto each atom. We propose a method that is able to do so. We find that the local entropy thus defined is able to distinguish extremely well between solid-like and liquid-like atoms. Furthermore, in conjunction with local enthalpy it can distinguish well between different polymorphs, even in the subtle case of the difference between fcc-like and hcp-like arrangements.

II. ENTROPY APPROXIMATION BASED ON THE TWO BODY CORRELATION FUNCTION

Ref. was based on the consideration that in the liquid to solid transition there is a trade-off between entropy and enthalpy. The role of metadynamics was there to enhance the fluctuations of these two quantities so as to accelerate crystallization. This required designing CVs able to describe these two quantities. Entropy is easy to compute but entropy is extremely costly to evaluate. However, an expression that gives an approximate evaluation of the entropy is sufficient for the purpose of driving crystallization. Such an expression was derived from an expansion of the configurational entropy in terms of multibody correlation functions. In simple liquids the second term of the expansion, often called two-body excess entropy, involves only the pair correlation function and accounts for about 90% of the configurational entropy. This term is given by,

\[
S_2 = -2\pi \rho k_B \int_0^\infty [g(r) \ln g(r) - g(r) + 1] r^2 dr,
\]

\(a)\) Electronic mail: parrinello@phys.chem.ethz.ch
where $\rho$ is the system’s density, and $g(r)$ is the radial distribution function. Extensions of the expansion to multicomponent\cite{19,20} and inhomogeneous\cite{21} systems are also available. We also recall that entropy series expansions have been used for order-disorder phenomena\cite{19,20},\cite{21} and inhomogeneous\cite{22} systems are available. We also recall that entropy series expansions have been used to study order-disorder phenomena\cite{19,20}.

In order to come to grasp with $S_2$ and understand better why it works, we first contrast in Fig. 1 the different behaviors of $g(r)$ and the integral in Eq. (1)

$$ I(r) = [g(r) \ln g(r) - g(r) + 1] r^2. $$

The data were taken from a system with Lennard-Jones interactions at temperature $T = 1.15$ and pressure $P = 5.68$, that correspond to the solid-liquid coexistence point.\cite{23} The Lennard-Jones potential was truncated at 2.5 and tail corrections were included. We refer the reader to Appendix A for further computational details. As usual we use Lennard-Jones units, i.e. $\sigma = 1$. We have chosen these thermodynamic conditions because at this temperature and pressure the fcc, hcp, bcc, and liquid phases are all metastable allowing a fair comparison. The first observation is that while $g(r)$ has some difficulty at distinguishing between solid and liquid, it strikes the eye that $I(r)$ in the liquid phase is much more short ranged than in the solid phases. Furthermore, the $g(r)$ for the solid phases can hardly distinguish between the different polymorphs. In contrast, the bcc $I(r)$ appears clearly different from that of the closed packed structures. More subtle is the difference between fcc and hcp, that is revealed only if one goes as far out as the third neighbor shell.

![FIG. 1. Analysis of functions related to the entropy approximation $S_2$. a) Radial distribution function $g(r)$, and b) integrand in Eq. (1) $I(r) = [g(r) \ln g(r) - g(r) + 1] r^2$. These functions are compared for the liquid, fcc, hcp, and bcc phases of a Lennard-Jones fluid at the melting temperature. We use Lennard-Jones units, i.e. $\sigma = 1$.](image)

III. ENTROPY FINGERPRINT FOR SOLID-LIKE AND LIQUID-LIKE ENVIRONMENTS

The analysis of $I(r)$ suggests that, if properly projected onto the different atoms, $S_2$ could be used as a fingerprint to identify local structures. The projection on atom $i$ can be achieved using the expression:

$$ s^i_S = -2\pi \rho k_B \int_0^{r_m} [g^i_m(r) \ln g^i_m(r) - g^i_m(r) + 1] r^2 \, dr, \quad (2) $$

where $r_m$ is an upper integration limit that in principle should be taken to infinity, and $g^i_m$ is the radial distribution function centered at the $i$-th atom. To obtain a continuous and differentiable order parameter, we define a mollified version of the radial distribution function\cite{12}

$$ g^i_m(r) = \frac{1}{4\pi \rho r^2} \sum_j \frac{1}{\sqrt{2\pi \sigma^2}} e^{-(r-r_{ij})^2/(2\sigma^2)}, \quad (3) $$

where $j$ are the neighbors of atom $i$, $r_{ij}$ is the distance between atoms $i$ and $j$, and $\sigma$ is a broadening parameter. We shall choose $\sigma$ so small that $g^i_m(r) \sim g(r)$ yet large enough for the derivatives relative to the atomic positions to be manageable.\cite{12} A similar projection of $S_2$ has been used in Ref.\cite{25}.

If we use $s^i_S$ as defined in Eq. (2) it can be seen in Fig. 2 that, in the cases of Na\cite{26} at 350 K and A\cite{26} at 900 K (see Appendix A for technical details), the distribution of $s^i_S$ in the liquid and solid phases are peaked at two different positions but exhibit a large overlap. In order to calculate local order parameters whose distributions are more clearly distinct, we take cue from Lechner and

![Image](image)
Dellago[8] and define an average local entropy:

$$\bar{s}_S = \frac{\sum_j s_j^S f(r_{ij}) + s_i^S}{\sum_j f(r_{ij}) + 1} \tag{4}$$

where $j$ runs over the neighbors of atom $i$ and $f(r_{ij})$ is a switching function with cutoff $r_a$. Switching functions have a value of 1 for $r_{ij} \ll r_a$, 0 for $r_{ij} \gg r_a$, and decay smoothly from 1 to 0 for $r_{ij} \approx r_a$. We have used a switching function with the functional form:

$$f(r_{ij}) = 1 - \left(\frac{r_{ij}}{r_a}\right)^N$$

with $N = 6$ and $M = 12$. Such a form has proven useful in many other contexts. At variance with $s_i^S$, the distributions of $\bar{s}_S$ of the liquid and solid phases now have a negligible overlap (see Fig. 2). Henceforth, we shall drop the index $i$ when referring to distributions and we shall refer to $\bar{s}_S$ as entropy fingerprint.

The ability to distinguish sharply between solid-like and liquid-like molecules depend on a wise choice of the parameters $r_m$ and $r_a$. As $r_m$ is increased, more of the long range part of the integrand is included making the difference between liquid and solid more and more evident. On the other hand by increasing $r_a$, more neighbors are included in the summation in Eq. (4) and eventually the locality of $\bar{s}_S$ is lost. In the practice we have chosen for $r_m$ and $r_a$ the smallest values that still ensure sharp distinction between solid-like and liquid-like atoms. The parameters $r_m$, $r_a$, and $\sigma$ that were used are summarized in Table I.

It is interesting to investigate whether the entropy fingerprint can identify ordered structures in a complex situation, in a context different from nucleation. To this effect we generated a nanocrystalline structure (see Fig. 3) using a procedure described in Appendix A. The system is Al, as described by the potential in Ref. 30. It can be seen that the entropy fingerprint clearly brings out the nanostructure of the system and the network of grain boundaries. This indicates that the entropy fingerprint can also work in inhomogeneous situations where different atomic environments coexist.

### IV. IDENTIFICATION OF CRYSTAL STRUCTURES

In the previous section we have shown that $\bar{s}_S$ is able to distinguish liquid-like from solid-like atomic environments. We will now explore the possibility of distinguishing between fcc, hcp, bcc and liquid-like atomic environments. As we shall see, this is best achieved if we accompany our definition of local entropy with a measure of local enthalpy.

| Structure | Model | T (K) | $r_m$ (a) | $r_a$ (a) |
|-----------|-------|-------|-----------|-----------|
| bcc       | Na    | 350   | 1.8 (5NS) | 1.2 (2NS) |
| fcc       | Al    | 900   | 1.4 (3NS) | 0.9 (1NS) |
interatomic potential $U(\mathbf{r})$ that can be decomposed into energies $U_i(\mathbf{r})$ associated to individual atoms. Here $\mathbf{r}$ denotes the atomic coordinates of an $N$ atom system. The expression that we shall use is then,

$$\bar{s}^i_H = U_i(\mathbf{r}) + PV/N$$

where $P$ and $V$ are the system’s pressure and volume, respectively and, for simplicity, we have partitioned the volume of the system into $N$ equal parts. A more complex partition criterion is also possible. As done for the local entropy, we define an average local enthalpy,

$$\bar{s}^j_H = \frac{\sum_i s^j_H f(r_{ij}) + s^j_H}{\sum_j f(r_{ij}) + 1}$$

where the symbols have the same meaning as in Eq. (4).

We calculated the joint probability distributions of $\bar{s}^i_H$ and $\bar{s}^j_S$ of the fcc, hcp, bcc, and liquid phases of the Lennard-Jones system described in Section II. For this purpose we simulated systems in each of those phases for 200 ps. The thermodynamic conditions were the same as described in Section II. We used the following parameters to define $\bar{s}^i_H$ and $\bar{s}^j_S$: $r_m = r_a = 2.5$, and $\sigma = 0.1$. The $P(\bar{s}^i_H, \bar{s}^j_S)$ of each phase are shown in Fig. 4. Each $P(\bar{s}^i_H, \bar{s}^j_S)$ was normalized to one.

We now discuss the results in Fig. 4. We first notice that the distributions of the different phases in Fig. 4 have minimal overlap and therefore $\bar{s}^i_H$ and $\bar{s}^j_S$ are useful fingerprints. As in the case of Na and Al, the distributions of liquid and solid phases are very far apart and therefore the fingerprints distinguish very well between liquid-like and solid-like environments. The distributions of the solid phases are clustered together in the region of low enthalpy and entropy, and it is easy to distinguish between the structures using $\bar{s}^i_H$ and $\bar{s}^j_S$. We analyze in detail the challenging case of fcc and hcp. Both fcc and hcp structures are formed by stacking of close-packed planes. However, they differ in the way the close-packed planes are stacked. For this reason, these structures are usually not easy to discriminate. As seen in Fig. 4, the fingerprints introduced in this work discriminate well between fcc and hcp configurations.

V. CONCLUSIONS

To conclude, the degree of success of the entropy based fingerprint is at first sight surprising. However, the root of this success must lie on the point of view taken here that does not directly focus on the local geometry but on properties of deeper thermodynamic significance, like local entropy and enthalpy. It also points to the usefulness of looking at old problems from a different standpoint.

Appendix A: Computational details

We performed molecular dynamics (MD) simulations using LAMMPS. We employed an anisotropic Parrinello-Rahman barostat and the stochastic velocity rescaling thermostat. The fingerprints were programmed in a development version of PLUMED.

The Lennard-Jones simulations were performed at temperature $T = 1.15$ and pressure $P = 5.68$ (solid-liquid coexistence). As usual, we use Lennard-Jones units, i.e. $\sigma = 1$ and $\epsilon = 1$. The Lennard-Jones potential was truncated at 2.5 and tail corrections were included. The time step for the integration of the equations of motion was 0.002. The relaxation times of the barostat and thermostat were 5 and 0.05, respectively.

Na and Al were simulated using embedded atom models (EAM). The time step for the integration of the equations of motion was 2 fs. For Na we set the temperature at 350 K, close to the melting temperature (366 K) of the model. For Al the temperature was set to 900 K, near the melting temperature 931 K. In both cases the pressure was set to its standard atmospheric value. The relaxation times of the barostat and thermostat were 10 ps and 0.1 ps, respectively. The results presented in Fig. 2 were obtained by performing independent simulations in the liquid and solid phases of Na and Al at the above cited temperatures. Each simulation had a length of 200 ps and the distributions of $\bar{s}^j_S$ and $\bar{s}^i_S$ were calculated taking samples every 1 ps.
The configuration of the nanocrystalline Al was constructed using Voronoi tessellation. The mean grain size was 5 nm and the system contained 255064 atoms. We performed an annealing at 600 K for 0.2 ns, then the temperature was ramped to 300 K in 0.2 ns, and finally the temperature was kept constant at 300 K for 0.2 ns. For these simulations we employed a different EAM potential. The configuration in Fig. 3 corresponds to the last in this trajectory. The simulation details were the same as those used for Al above.

EAM potentials have a natural way to partition the energy between the atoms as needed in Eq. (6), i.e.

\[ U_i(R) = \sum_{j \neq i} \phi(r_{ij}) + F \left( \sum_{j \neq i} \rho_{atom}(r_{ij}) \right) \]  

where \( \phi \) is a pairwise potential, \( F \) is the embedding energy function, and \( \rho_{atom} \) is the electron charge density function. We have used this partition criterion.

ACKNOWLEDGMENTS

This research was supported by the NCCR MARVEL funded by the Swiss National Science Foundation. The authors also acknowledge funding from the European Union Grant No. ERC-2014-AdG-670227 / VARMET. The computational time for this work was provided by the Swiss National Supercomputing Center (CSCS) under project ID mr3. Calculations were performed in CSCS cluster Piz Daint.

1. P. R. ten Wolde and D. Frenkel, Physical Chemistry Chemical Physics 1, 2191 (1999).
2. F. Giberti, M. Salvadagio, M. Mazzotti, and M. Parrinello, Chemical Engineering Science 121, 51 (2015).
3. P. R. Ten Wolde, M. J. Ruiz-Montero, and D. Frenkel, Physical Review Letters 75, 2714 (1995).
4. W. Lechner, C. Dellago, and P. G. Bolhuis, Physical Review Letters 106, 085701 (2011).
5. M. A. Meyers, A. Mishra, and D. J. Benson, Progress in Materials Science 51, 427 (2006).
6. J. D. Honeycutt and H. C. Andersen, Journal of Physical Chemistry 91, 4550 (1987).
7. A. Stukowski, Modelling and Simulation in Materials Science and Engineering 20, 045021 (2012).
8. W. Lechner and C. Dellago, The Journal of chemical physics 129, 114707 (2008).
9. P. Steinhardt, D. Nelson, and M. Ronchetti, Phys. Rev. B 28, 784 (1983).
10. A. Laio and M. Parrinello, Proceedings of the National Academy of Sciences 99, 12562 (2002).
11. A. Barducci, G. Bussi, and M. Parrinello, Physical Review Letters 100, 020603 (2008).
12. P. M. Piaggi, O. Valsson, and M. Parrinello, Physical Review Letters 119, 015701 (2017).
13. H. S. Green, The molecular theory of fluids (North-Holland Publishing Company Amsterdam, 1952).
14. R. Nettleton and M. Green, The Journal of Chemical Physics 29, 1365 (1958).
15. A. Baranyai and D. J. Evans, Physical Review A 40, 3817 (1989).
16. D. C. Wallace, The Journal of chemical physics 87, 2282 (1987).
17. D. C. Wallace, International journal of quantum chemistry 52, 425 (1994).
18. B. B. Laird and A. Haymet, Physical Review A 45, 5680 (1992).
19. J. Hernando, Molecular Physics 69, 319 (1990).
20. S. Prestipino and P. V. Giaquinta, Journal of Statistical Mechanics: Theory and Experiment 2004, P09008 (2004).
21. T. Morita and K. Hiroike, Progress of Theoretical Physics 25, 537 (1961).
22. R. Kikuchi, Physical review 81, 988 (1951).
23. J.-P. Hansen and L. Verlet, Physical Review 184, 151 (1969).
24. D. Frenkel and B. Smit, *Understanding molecular simulation: from algorithms to applications*, Vol. 1 (Academic press, 2001).
25. M. Leocmach, J. Russo, and H. Tanaka, The Journal of chemical physics 138, 12A530 (2013).
26. S. Wilson, K. Gunawardana, and M. Mendelev, The Journal of chemical physics 142, 134705 (2015).
27. J. B. Sturgeon and B. B. Laird, Physical Review B 62, 14720 (2000).
28. G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, and G. Bussi, Computer Physics Communications 185, 604 (2014).
29. A. Stukowski, Modelling and Simulation in Materials Science and Engineering 18, 015012 (2009).
30. M. Mendelev, M. Kramer, C. A. Becker, and M. Asta, Philosophical Magazine 88, 1721 (2008).
31. S. Plimpton, Journal of computational physics 117, 1 (1995).
32. M. Parrinello and A. Rahman, Journal of Applied physics 52, 7182 (1981).
33. G. Bussi, D. Donadio, and M. Parrinello, The Journal of chemical physics 126, 014101 (2007).
34. P. Piaggi, E. Bringa, R. Pasianot, N. Gordillo, M. Panizo-Laiz, J. del Rio, C. G. de Castro, and R. Gonzalez-Arrabal, Journal of Nuclear Materials 458, 233 (2015).
35. M. S. Daw and M. I. Baskes, Physical Review B 29, 6443 (1984).
36. M. Finnis and J. Sinclair, Philosophical Magazine A 50, 45 (1984).