

# Spatial uniformity as a principle for determination of atomistic structural models

Bin Cai<sup>1</sup>, Andrew L Goodwin<sup>2</sup> and David A Drabold<sup>1</sup>

<sup>1</sup> Department of Physics and Astronomy, Ohio University, Athens, OH 45701, USA

<sup>2</sup> Department of Chemistry, University of Oxford, Oxford OX1 3QR, UK

E-mail: [drabold@ohio.edu](mailto:drabold@ohio.edu)

Received 24 September 2010, in final form 3 March 2011

Published 24 March 2011

Online at [stacks.iop.org/MSMSE/19/035010](http://stacks.iop.org/MSMSE/19/035010)

## Abstract

We discuss the concept of structural uniformity as a tool for model construction, and employ the recently proposed reverse Monte Carlo + invariant environment refinement technique (RMC+INVERT) method to create computer models of amorphous materials. We further develop the RMC+INVERT method, and apply it to model glassy GeO<sub>2</sub>, by jointly fitting partial pair-correlation functions rather than the total correlation function as in previous calculations. Then the technique was applied to model water-ice (a-H<sub>2</sub>O), for which the uniformity is imposed on the medium range order. For these two cases, we found that imposing spatial uniformity may significantly improve the atomistic model, and the RMC+INVERT method is shown to offer significant advantage over the conventional RMC technique for particular systems.

(Some figures in this article are in colour only in the electronic version)

## 1. Introduction

The lack of translational periodicity is an essential feature of amorphous solids. Because of the absence of long-range order, it is always a challenge to create realistic atomistic models. Still, *local* (short-range) structural order does exist in all amorphous materials. Experimental information about this local ordering is provided by x-ray and neutron diffraction, and other probes. Information-based modeling, including the reverse Monte Carlo (RMC) method [1], may be used to generate models consistent with such diffraction measurements. Unfortunately, the models obtained from such methods are not unique. In other words, various, and indeed discrepant models can reproduce the same RDF data, and these models usually have an unphysical concentration of chemical and/or coordination defects. To remedy this, empirical assumptions are usually made and applied with RMC in the form of constraints on coordination numbers, bond angles, etc. For different atomic species, these assumptions vary, so it is hard to form a routine or universal RMC approach. This paper can be understood as suggesting

an implementation of RMC with a new kind of constraint [2]<sup>3</sup>: structural uniformity. In some ways, coordination number or angular constraints [2] are similar to our strategy, though different in detail.

Recently, a new method, the ‘invariant environment refinement technique’ (INVERT) has been shown to improve chemical and topological order for C<sub>60</sub>, a-Si and a-SiO<sub>2</sub> [3]. By insisting on maximum spatial uniformity, other empirical constraints become unnecessary. Thus, the single requirement of maximum uniformity may become a generic prescription for modeling many homogeneously disordered by materials. In this paper, we further develop the RMC+INVERT method and apply it to a-GeO<sub>2</sub> and a-H<sub>2</sub>O. For materials with a relatively uniform structure, such as GeO<sub>2</sub>, fitting partial RDFs, INVERT+RMC lead to much improved models in terms of both chemical order and local topological order (as gauged by the distribution of bond angles). In the case of a-H<sub>2</sub>O, we show that the new technique can impose suitable medium range order and easily handle the rigid molecule constraint. Moreover, for both cases, INVERT+RMC method is more efficient than traditional RMC method.

Of course, no method is a panacea, and INVERT+RMC is no exception. For systems that are not structurally uniform, the method is difficult to apply. We are currently trying to develop suitable generalizations to handle such cases, but the jury is still out. Nevertheless, many important amorphous systems are quite structurally homogeneous, and for these, it is clear that INVERT+RMC is helpful. Moreover, our method is very much based on real space where the whole ‘INVERT’ strategy is most naturally implemented. Employing information from structure factor could be also considered; the INVERT, radial distribution functions and structure factors could be combined together into the modeling process.

Philosophically, it is clear that an ideal structure inversion method should produce models consistent with our full knowledge base: an ideal model contradicts no experiments that we believe and must also be a suitable energy minimum according to accurate interatomic interactions. Thus, INVERT+RMC is useful both as a tool for inverting experimental data, and as a means of requiring a model to agree with experiments within an experimentally constrained molecular relaxation approach [4, 5].

## 2. Method and development

Compared with traditional RMC, the INVERT+RMC method requires a system to have maximum structural uniformity consistent with the experimental data: namely each distinct site is required to have an identical local environment to the maximum degree possible. By introducing atomic RDF,  $g_i(r)$ , the idea is to connect each individual site to experimental RDF data. To implement this strategy, two major terms are included in the cost function for each individual site: the RDF fitting term (1) and the variance term (2) which maximize uniformity.

$$\chi_{\text{RDF},i}^2 = \sum_r [g_i(r) - g(r)]^2 \quad (1)$$

$$\chi_{\text{VAR},i}^2 = \sum_j \frac{[d_i(j) - \langle d(j) \rangle]^2}{\langle d(j) \rangle^2}. \quad (2)$$

The total cost function would then have the form

$$\chi^2 = \sum_{i=\text{all,atoms}} (\chi_{\text{RDF},i}^2 + \chi_{\text{VAR},i}^2). \quad (3)$$

<sup>3</sup> For the utility of constraints in a-Si, see [2].

The RDF term quantifies the difference between the calculated atomic RDF data for each site in the model and the experimental data. The variance term measures the dispersion in the local environment for each site. The simplest types of materials, containing a single atomic local environment such as C<sub>60</sub> and a-Si, have been discussed in a previous paper [3]. For a system with two or more distinct sites, the cost functions will be modified to the form (4) and (5).

$$\chi_{\text{RDF},i\alpha}^2 = \sum_{\beta=\text{all species}} \sum_r [g_{i,\alpha\beta}(r) - g_{\alpha\beta}(r)]^2 \quad (4)$$

$$\chi_{\text{VAR},i\alpha}^2 = \sum_{\beta=\text{all species}} \sum_j \frac{[d_{i\alpha}(j, \beta) - \langle d_{\alpha}(j, \beta) \rangle]^2}{\langle d_{\alpha}(j, \beta) \rangle^2}. \quad (5)$$

Equations (2) and (5) can be interpreted as a mathematical constraint imposed upon the RMC process. In the RDF term, experimental data (a partial RDF) are used for each distinct site. For a site  $i$  with type  $\alpha$ , different partials  $g_{\alpha\beta}(r)$  will be fitted separately according to different type of  $\beta$ . In the variance term, the distance to different neighbors with respect to different types will be considered separately.  $d_{i\alpha}(j, \beta)$  means the distance from a site  $i$  with type  $\alpha$  to its  $j$ th neighbor with type  $\beta$ ;  $\langle d_{\alpha}(j, \beta) \rangle$  means the average distance from atoms with type  $\alpha$  to their  $j$ th with type  $\beta$ . The partial RDF can be obtained experimentally using the method of isotopic substitution in neutron diffraction [6, 7], or other methods.

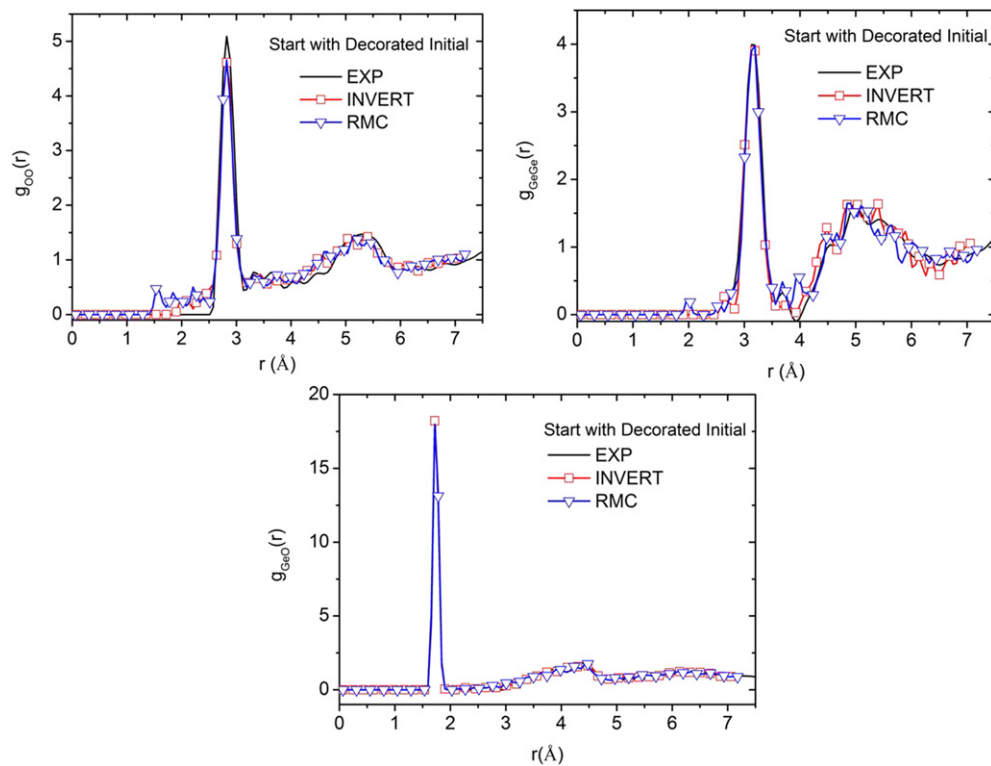
### 3. Results

#### 3.1. a-GeO<sub>2</sub>

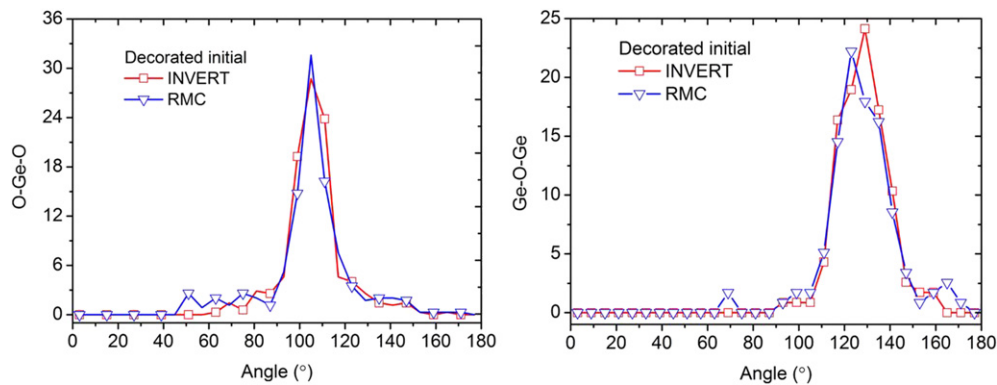
We first applied the INVERT+RMC method to model a binary material: a-GeO<sub>2</sub> and compared it with traditional unconstrained RMC method. The experimental partial RDF data are obtained from [7, 8]. There are two distinct sites in the a-GeO<sub>2</sub> network: Ge and O sites. In theory, Ge atoms with its four O atoms form tetrahedral structures with O–Ge–O bond angle around 109.47°; O atoms are two-fold and form an average bond angle around 132° [7–9]. However, we should emphasize that according to experimental results from Salmon and his colleagues, the mean coordination of Ge,  $n_{\text{Ge-O}}$ , is 3.8 (less than 4, we enforce) with  $r_{\text{Ge-O}} = 1.73\text{\AA}$ .

Two initial configurations are used: decorated initial model and random initial model. The decorated initial configuration is generated by substituting Ge for Si in an a-Si cell (made with WWW method [11]), decorating all the Ge–Ge bonds with a bond-center O, and rescaling the cell to satisfy experimental mass density [10]. The random initial configuration has totally randomized atomic coordinations with NO constraints. For comparison, we applied both INVERT and traditional RMC on 192-atom cells with the initial configurations mentioned above. We plot the results with decorated initial configuration in figures 1 and 2, and random initial configurations in figures 3 and 4.

We see that all models have good fit to the Ge–O partials with peak position around 1.73 Å, but a few homopolar bonds exist in traditional RMC models and lead to unrealistic peaks in Ge–Ge and O–O partials. These results indicate that for both initial models, INVERT converges to a minimum faster and manages to avoid some particular local minima. The non-smooth pattern, compared with experimental data, is mainly due to the size effect. For the angle distribution, all models predict a similar peak position around 109° for O–Ge–O and around 129° Ge–O–Ge, which are close to experimental values 109.47° and 132°. However, INVERT provides a better O–Ge–O distribution and much better Ge–O–Ge distribution. Again, the unrealistic peak in angle distribution of RMC model is mainly due to the mis-coordinated atoms in the network. Considering different initial configurations, the decorated model, as

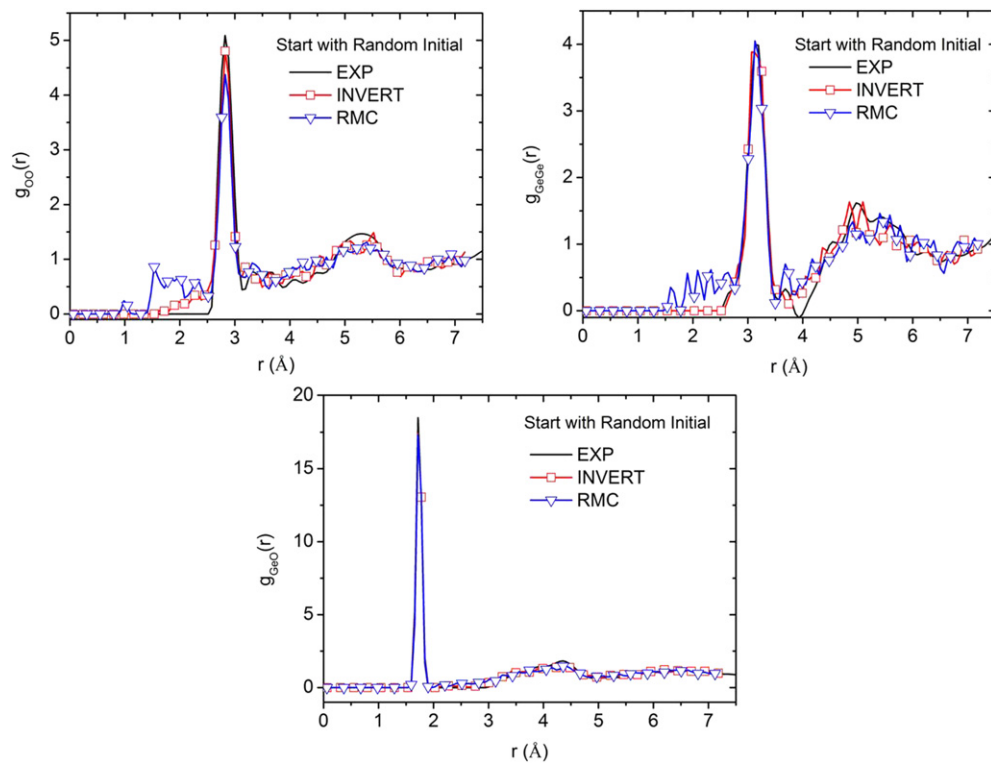


**Figure 1.** Comparison of partial radial distribution functions for 192-atom a-GeO<sub>2</sub> model obtained by different methods starting with decorated initial configuration. (See text.)

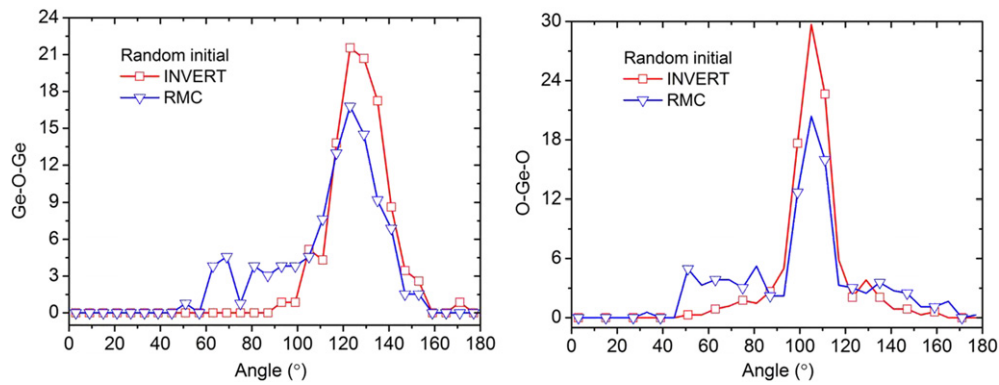


**Figure 2.** Comparison of angle distributions calculated for 192-atom a-GeO<sub>2</sub> model obtained by different methods starting with decorated initial configuration.

expected, gives a better result than random initial, since the decorated model starts with perfect chemical order which provides a better starting point than the random model. We report the structural statistics of these four models in table 1. The result confirms that INVERT+RMC with decorated initial configurations gives the best result with  $n_{\text{Ge}} = 3.8$ . Compared with the INVERT model, more homopolar bonds exist in the traditional RMC model. In figure 5, we show the topology of these four models. Atoms involved in homopolar bonds are marked as



**Figure 3.** Comparison of partial radial distribution functions for 192-atom a-GeO<sub>2</sub> model obtained by different methods starting with random initial configuration. (See text.)

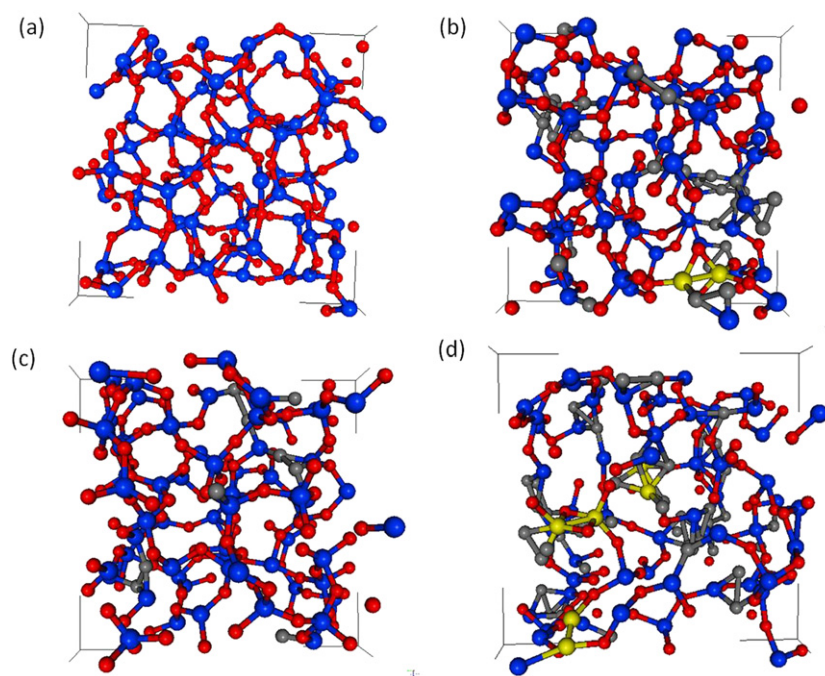


**Figure 4.** Comparison of angle distributions calculated for 192-atom a-GeO<sub>2</sub> model obtained by different methods starting with random initial configuration.

gold/light/big for Ge and gray/light/small for O. The illustration indicates that there are more defects in the RMC models, implying a poorer chemical order, compared with INVERT+RMC, consistent with our prior experience [3]. All these results reveal that by fitting the partial RDFs for each distinct site, the INVERT method achieves a credible network while the unconstrained RMC result is not very chemically realistic. We should also point out that to obtain these models, the RMC method is more computationally demanding than the INVERT method.

**Table 1.** Structural statistics of 192-atom models.

	INVERT (dec)	INVERT (ran)	RMC (dec)	RMC (ran)
$n_{\text{Ge}}$	3.81	3.75	3.8	3.6
$n_{\text{O}}$	1.91	1.99	2.1	1.85
$n_{\text{Ge-Ge}}(\%)$	0	0	1	5
$n_{\text{O-O}}(\%)$	0	6	10	29
$r_{\text{Ge-O}}(\text{\AA})$	1.73	1.73	1.73	1.73
$\text{O-Ge-O}(\text{\textcircled{^\circ}})$	109	109	109	109
$\text{Ge-O-Ge}(\text{\textcircled{^\circ}})$	129	127	129	124



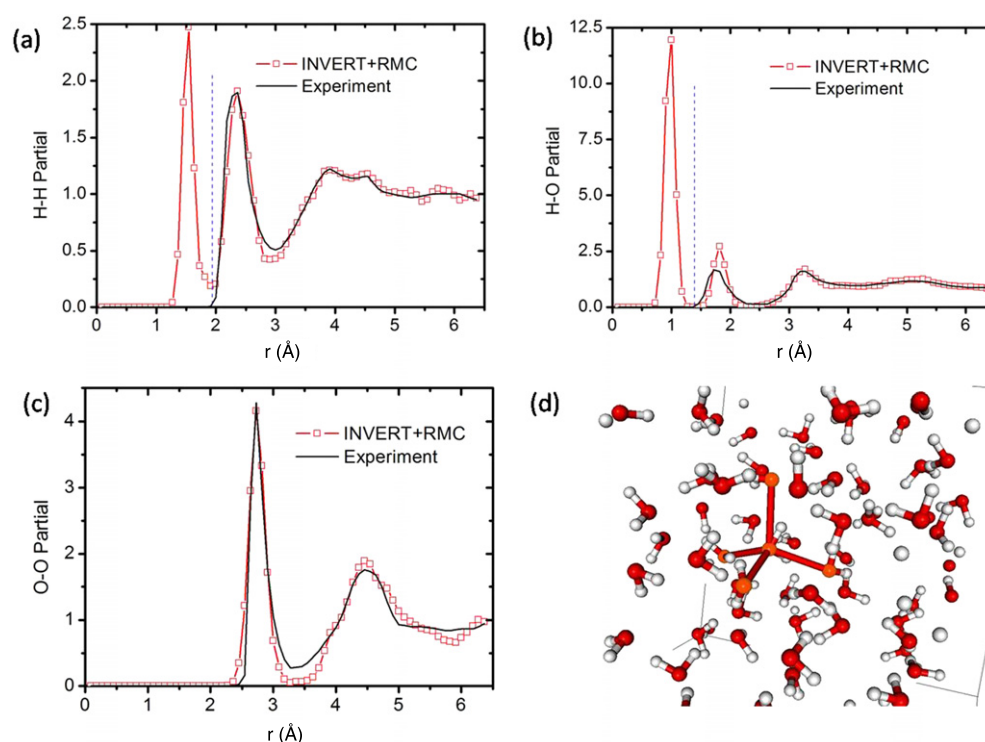
**Figure 5.** Comparison of topology of 192-atom a-GeO<sub>2</sub> obtained by (a) INVERT with decorated initial; (b) traditional RMC with decorated initial, (c) INVERT with random initial and (d) RMC with random initial. Blue (dark) and big atoms are Ge; gold (light) and big atoms are Ge atoms with homopolar bond (trouble Ge site); red (dark) and small atoms are O; gray (light) and small atoms are O atoms with homopolar bond (trouble O site).

(Since the INVERT+RMC model only considers the RDF for each individual site, it is an order  $N$  calculation; while it is order  $N^2$  for traditional RMC to calculate the average RDF.) The RMC results could be improved by adding coordination constraints, shortest distance constraints, which are not necessary for the INVERT+RMC method.

### 3.2. Low density amorphous ice: a-H<sub>2</sub>O

Another interesting case is low density amorphous ice (LDAI). In our approach, we use the simplest model for water molecules: O atoms connected rigidly to two H neighbors with bond length 1.0 Å and bond angle 106°. For a-H<sub>2</sub>O, one possible model is that water molecules are packed in a way such that O atom with its first four O neighbors forms a tetrahedral structure with medium range order [12]. Since intramolecular OH bonds are known, the peaks

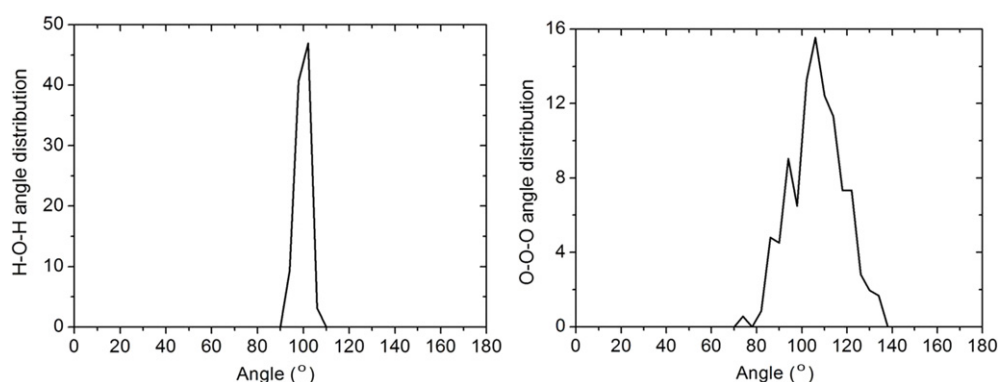




**Figure 6.** Three partial distribution functions and structure for 192-atom a-H<sub>2</sub>O. Red (bigger) atoms are O and gray (small) atoms are H.

arising from O–H bond within small  $r$  region are often suppressed in reported RDF data. Thus, the H<sub>2</sub>O molecules are often treated as rigid units and RMC is used only to fit the intermolecular peaks. Molecular dynamic simulations have also been used to model a-H<sub>2</sub>O, but density functional theory (DFT) does not correctly treat the weak interactions such as hydrogen bonding properly, and fails to reproduce experimental RDF data. In this section, we show that the RMC+INVERT method provides a convenient way to model LDAI, and it might improve DFT results on medium range order and generate an experiment credible model, perhaps using ECMR [5].

The rigid molecule constraint could be realized by setting the average distance to a constant value from atoms to their neighbors in the variance term. For example, for an O site, the distance to its first two H neighbors is fixed to 1.00 Å; for an H site, the distance to its first H neighbor is set to be 1.6 Å (this satisfies the H–O–H angle requirement). For the RDF part, we could still fit partial RDF but ignore the first peak in O–H and H–H partial. We start with previously relaxed models through DFT. 648- and 192-atom models are generated. The fitting result for 192-atom cell is plotted in figure 6 and compared with experiment [12]. In figures 6(a) and (b), we only fit the large- $r$  region ( $r$  to the right of the dashed line) for H–H and O–H partials. After fitting, the first peaks in O–H and H–H partials are centered at 1.00 Å and 1.6 Å which implies the intramolecular structure is maintained. Compared with experimental data, peaks almost perfectly fit the H–H, O–O partials, but are a little bit off for O–H partial. The off-fitting on second peak of O–H partial may be due to the fixed-distance constraint and the weighting factors in the variance term. Future work should better characterize the weight factor for the variance term relative to the experimental terms. Considering the first minimum after the



**Figure 7.** H–O–H intramolecular angle and O–O–O intermolecular angle distribution for 192-atom a-H<sub>2</sub>O.

primary peak, data show a higher value which RMC+INVERT cannot fit perfectly. The final configuration is also shown in figure 6(d). One sample is shown in this plot showing how O and its four neighbors may form a tetrahedral structure.

We plot the angle distribution in figure 7. Again, the plot reveals the short-range order and medium range order for water molecules. H–O–H distribution possess a peak at 103° which is close to 106° and represents the V-shaped water molecule. The major peak in O–O–O distribution is located at 109°, implying a tetrahedral structure.

Finally, by fixing the average distance in the variance term and assigning low tolerance, INVERT+RMC could easily handle the rigid unit constraint. Meanwhile, the RDF term could be applied to fit a second peak and force maximum uniformity in medium range order.

#### 4. Conclusion

We have developed the INVERT+RMC method to fit experimental data of partial radial distribution functions. The result on a-GeO<sub>2</sub> indicates that the coordination number and angle distribution both are improved using the INVERT technique. For the a-H<sub>2</sub>O case, fixing the average distance in the variance term and assigning these terms low tolerance allows the technique to handle the rigid unit constraint. For non-uniform system, our scheme is difficult to apply and additional work is needed to determine if it can be usefully employed in this case. Our method can be interpreted as a new and useful kind of constraint to employ in RMC structural modeling.

We have shown that the imposition of the uniformity constraint is a significant boon for homogeneous systems, and also has computational advantages over the usual implementation of RMC. We have shown how to apply the method with *partial* structure factors, and have opened a broader inquiry into how the method may be generalized to systems that are substantially but not ideally uniform. We have offered useful new models of a-GeO<sub>2</sub> and amorphous ice.

#### Acknowledgments

We gratefully acknowledge the support of the National Science Foundation under grants DMR-09-3225 and DMR-0844014. ALG thanks the EPSRC (UK) for support under grant



ep/g004528/2. We would like to thank Professor Philip S Salmon providing experimental data for GeO<sub>2</sub>; and strongly acknowledge Dr F Inam for providing ideas and models for LDAI. We are grateful to Dr Mingliang Zhang for helpful discussions.

## References

- [1] McGreevy R L and Pusztai L 1988 *Mol. Simul.* **1** 359–67
- [2] Biswas P, Atta-Fynn R and Drabold D A 2004 *Phys. Rev. B* **69** 195207
- [3] Cliffe M, Dove M T, Drabold D A and Goodwin A L 2010 *Phys. Rev. Lett.* **104** 125501
- [4] Biswas P, Tafen D N, Atta-Fynn R and Drabold D A 2004 *J. Phys.: Condens. Matter* **16** S5173  
Biswas P, Tafen D and Drabold D A 2005 *Phys. Rev. B* **71** 054204
- [5] Drabold D A 2009 *Eur. Phys. J. B* **68** 1
- [6] Petri I, Salmon P S and Fischer H E 2000 *Phys. Rev. Lett.* **84** 2413–6
- [7] Salmon P S, Barnes A C, Martin R A and Cuello G J 2006 *Phys. Rev. Lett.* **96** 235502
- [8] Salmon P S, Barnes A C, Martin R A and Cuello G J 2007 *J. Phys.: Condens. Matter* **19** 415110
- [9] Giacomazzi L, Umari P and Pasquarello A 2005 *Phys. Rev. Lett.* **95** 075505
- [10] Tafen D and Drabold D A 2003 *Phys. Rev. B* **68** 165208
- [11] Wooten F and Weaire D 1987 *Solid State Physics* ed H Ehrenreich *et al* (New York: Academic) p 40
- [12] Pusztai L 2000 *Phys. Rev. B* **61** 28