

Reverse Monte Carlo modeling of amorphous germanium

J. K. Walters and R. J. Newport

Physics Laboratory, The University, Canterbury, Kent CT2 7NR, United Kingdom

(Received 8 August 1995)

The reverse Monte Carlo (RMC) method has been used to generate a model for the atomic structure of amorphous germanium (*a*-Ge). Fitting to experimental neutron diffraction data and applying coordination number and "triplet" constraints, the positions of 3000 "atoms" in a box, with full periodicity, were altered until the associated model structure factor, $S(Q)$, and pair correlation function, $g(r)$, agreed with the analogous experimental data, within the errors. The model generated is then analyzed to obtain coordination number, bond angle, and ring size distributions. These, in turn, are compared to the results obtained from random network models and other RMC studies. The effects of increasing the number density from the experimentally determined bulk value are also investigated. The results are consistent with an atomic structure which has a characteristic disordered tetrahedral network. For the highest density RMC model, the mean bond angle is 109.4° and the average coordination number is 3.49.

I. INTRODUCTION

At present there is no model for the structure of amorphous germanium (*a*-Ge), which can reproduce all the features of the experimental diffraction data: it is thought that the structure resembles a closely tetrahedrally bonded random network, although there is still much discussion about the detailed structure. The most successful models for describing tetrahedrally bonded amorphous materials are random network models (see, for example, Refs. 1–6), although others have been proposed.^{7–10} The original model, as applied to amorphous semiconductors, was developed by Polk.¹ However, when compared with experimental data from diffraction measurements, none of these models for *a*-Ge reproduce the features of the data satisfactorily¹¹ over more than a short-distance scale.

Neutron diffraction is a well-established method of obtaining structural information on disordered systems, such as amorphous solids. From the experimentally measured pair correlation function, accurate peak positions and quantitative areas can be obtained; however, a three-dimensional model of the structure cannot be built solely on this intrinsically one-dimensional information. The reverse Monte Carlo (RMC) method¹² offers one possible route towards this goal.

RMC is a method for producing three-dimensional models of the structure of disordered materials that agree quantitatively with the available data, usually diffraction data. (In this limited sense the method is comparable to the use of Reitveld refinement in determining crystal structures.) Although the model produced must be consistent with the experimental data, it cannot be regarded as unique given the intrinsic limitation imposed by using a one-dimensional basis data set. Unlike molecular dynamics and other Monte Carlo-based simulation methods, RMC modeling requires no interatomic potential, and it also allows complete data sets from different sources, e.g., diffraction of neutrons and x rays, to be fitted simultaneously, together with other appropriate constraints derived from prior chemical knowledge. These factors make it appealing in the study of a system such as *a*-Ge, where an interatomic potential is difficult to define, but high

quality experimental data is available. Some work has already been done in trying to develop structural models for *a*-Si and *a*-Ge,^{13,14} using the RMC method, and the results presented here will be compared to those obtained in previous studies.

II. EXPERIMENTAL BACKGROUND

The neutron diffraction data was obtained by Wright *et al.* in an experiment carried out at the Institute Laue-Langevin (Grenoble, France), on the D4 twin-axis diffractometer. Further details of the experimental procedure can be found in Ref. 11.

In performing a diffraction experiment, the quantity we wish to obtain is the structure factor $S(Q)$, where, for an amorphous material (i.e., an isotropic scatterer)¹⁵

$$S(Q) = 1 + \frac{4\pi\rho}{Q} \int_0^\infty r[g(r) - 1] \sin(Qr) dr, \quad (1)$$

where ρ is the average number density of atoms in the material, $|Q| = |k_i - k_f|$ is the wave-vector transfer associated with the diffraction experiment, and $g(r)$ is the pair distribution function, which is a measure of the atomic density at a distance r from a given atom at the origin. The pair distribution function may be obtained by Fourier transformation of the structure factor, which is directly related to the measured neutron scattering intensity.

The experimental data are subject to several corrections before the structure factor is obtained. Following subtraction of the background counts and corrections for absorption, multiple scattering, and self-shielding,¹⁶ the data were normalized using the Krogh-Moe-Norman technique.^{17,18} After normalization the diffraction pattern did not oscillate correctly about the self-scattering (Placzek corrections were applied according to the formalism of Yarnell *et al.*¹⁹) but dropped below it at high Q . This was identified by Wright *et al.* as being due to hydrogen contamination in the form of residual absorbed water and was corrected for using the neutron diffraction data of Beyster²⁰—see Ref. 11 for details.

The basic RMC algorithm has been described in detail elsewhere.^{12,21} In essence, “atoms” in a box are moved until the derived pair correlation function $g(r)$ and/or the structure factor $S(Q)$ matches the experimentally measured data. The important steps in the method are given below.

(1) Define a box with edge dimensions at least twice the value of r at which statistically significant oscillations in $g(r)$ disappear and then fill the box with “atoms” (either at random, or using a simple lattice) so that the number density matches the measured bulk value. The number density is an important parameter in RMC model building, as in all other modeling and simulation work. The fact that the modeling process ought to be conducted using the microscopic density (which often cannot be measured accurately, or at all), rather than the measured bulk value is discussed more fully below.

(2) The atoms are moved at random, and at each stage a model $S(Q)$ is calculated. Note that an atom is moved subject to the constraint that it does not overlap with a neighboring atom [an “excluded volume” is therefore associated with each atom type; this may be determined from the experimental $g(r)$ and need not rely on prior other knowledge].

(3) The model $S_{\text{mod}}(Q)$ is compared with the experimental $S_{\text{expt}}(Q)$, and a new configuration is accepted if the associated χ^2 has been reduced:

$$\chi^2 = \sum_i [S_{\text{expt}}(Q_i) - S_{\text{mod}}(Q_i)]^2 / \sigma_i^2.$$

Rejection is subject to a probability function dependent on the experimental error σ_i . This ensures that the algorithm’s model is not “trapped” in a subsidiary minimum.

(4) The process is repeated until the model $S(Q)$ reproduces experiment to within the experimental errors. The process is further iterated until an “equilibrium” configuration is obtained. Where more than one data set is used, the χ^2 calculations and acceptance or rejection criteria are applied to each. This is also the case for any constraints imposed on the bonding of the atoms.

For the model presented here, a box edge of 42.3 Å was used, containing 3000 atoms placed initially at random sites, subject to the criterion that defined distances of closest approach were not violated. In other RMC studies^{13,14} a box size of 1728 particles was used; this gives a box length of only 32 Å, which implies that the significant oscillations in $g(r)$ should not appear beyond 16 Å. In the *a*-Ge system, however, where there is significant medium range order, this assumption is unlikely to be realized, and the larger box size used in this study should significantly reduce the truncation errors, which may result from using a smaller box length. This also illustrates one of the important benefits of RMC over the hand-built “Polk-type” models. The largest such model for *a*-Ge consists of 563 atoms,⁵ which is too small realistically to generate ensemble average figures for ring statistics, bond distances, etc. Neutron diffraction $S(Q)$ and $g(r)$ data were modeled simultaneously using a modified version of the original code supplied by McGreevy.²² This was run on a DEC Alpha 3000 processor, where on average $\sim 10^6$ moves may be attempted in a 24 h period. The total number of accepted moves for the models presented here was $\sim 10^6$, compared to $2\text{--}3 \times 10^5$ for the models by Gerben and Pusztai.¹⁴ Considerations for deciding whether or

not the model has reached equilibrium are generally based on the ratio of moves tried to moves accepted. The criterion we have adopted is that the algorithm be run until this ratio fell to $\sim 1000:1$, i.e., out of 1000 generated moves, only 1 move was found to be acceptable. During the fitting process, secondary minima were avoided by cycling the maximum move size through the range 1–0.001 Å (and via the move rejection probability function determined on the basis of the error associated with the experimental data).

In our work modeling amorphous hydrogenated carbon (*a*-C:H),^{23,24} we have come across several problems intrinsic to applying the RMC method to covalently bonded amorphous systems, where diffraction data is used. This has led to the introduction of constraints on the model to avoid chemically and physically unreasonable features.

The first of these constraints to be adopted in the present study is one that removes “triples,” that is, three atoms forming an equilateral triangle with side length equal to the near-neighbor distance Ge-Ge. The existence of these is typified by a sharp peak at 60° (Refs. 13, 14, and 25) in the bond-angle distribution. Since the formation of these results in three bonds at the required distance, and therefore a relatively large fall in the χ^2 , it is not surprising that they form so readily; however, such a conformation is highly strained and unlikely to occur in any substantial quantity in the real *a*-Ge network. A constraint was therefore introduced that prevents the formation of triplets and removes the majority of those already in existence after the initial box-filling operation. This is done by examining the box for 60° bond angles and then preferentially moving the atoms involved in forming that angle, until the triplet is removed; in the same way, any move that creates a triplet is rejected.

In other cases, this problem has been overcome by constraining the bond angle distribution¹³ and by requiring 100% fourfold coordination.¹⁴ However, constraining bond angles in this way is very expensive on computer time and is difficult to implement simply; also our results show that this is unnecessary. Constraining the coordination number of every atom to be equal to 4 can be problematic, since it excludes all dangling bonds, some of which are known to be present because the experimentally determined average coordination number is less than 4.

In the present study we assume that each Ge atom can form three or four bonds, i.e., can have three or four nearest neighbors, according to basic chemical information. It is therefore reasonable to constrain the coordination number of each Ge atom to be less than five, but greater than two. This can be done quite simply by discouraging configurations such that atoms have zero, one, two, or five nearest neighbors between 2.23 and 2.74 Å, where distances are taken from the lower and upper limits of the first peak in the $g(r)$, respectively.

Since both these constraints arise from chemical considerations, the initially random arrangement of atoms was made to satisfy these conditions as completely as possible, before fitting to the experimental data was begun.

From our previous work on *a*-C:H,²⁴ it is also apparent that the number density used in generating the model is a crucial parameter. The original value adopted for this *a*-Ge sample, determined empirically from the experimental neutron data¹¹ is 0.039 75 atoms Å⁻³, which is a bulk or mac-

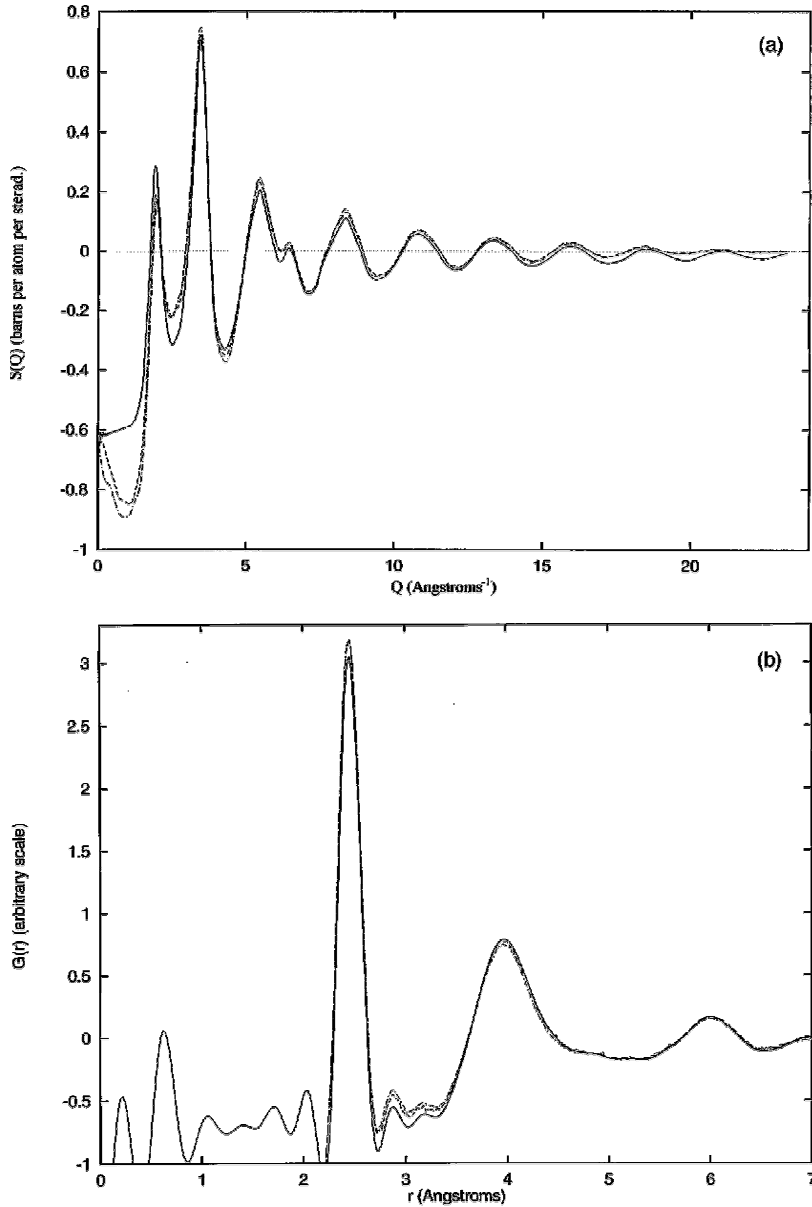


FIG. 1. (a) RMC-generated fits to the experimental structure factor: experimental data (—), low-density RMC model (---), and high-density RMC model (-·-·-). (b) RMC-generated fits to the experimental pair correlation functions: experimental data (—), low-density RMC model (---), and high-density RMC model (-·-·-).

roscopic number density. However, the value required for the RMC method is the microscopic density, which cannot be determined accurately by experimental methods. Therefore, using a value of $0.03975 \text{ atoms \AA}^{-3}$ is likely to be in error for a material such as *a*-Ge, where voids are known to be present. Random network models for this data (e.g., Refs. 2–6 and 8) give densities in the range 0.03952 – $0.04497 \text{ atoms \AA}^{-3}$. RMC calculations by Gereben and Pusztai¹⁴ give an optimum microscopic density for *a*-Ge of $0.043 \text{ atoms \AA}^{-3}$, somewhat higher than the experimentally determined value.

To explore the effect of density on the RMC work, two models have been generated using densities of $0.03975 \text{ atoms \AA}^{-3}$ (corresponding to the experimentally measured bulk density) and $0.04275 \text{ atoms \AA}^{-3}$. The same method of fitting was used for both, and the effects on the final configuration are presented.

III. RESULTS AND DISCUSSION

Figures 1(a) and 1(b) show the fits to the neutron diffraction $S(Q)$ and $g(r)$, respectively, for the two different densities used to produce the models. Both the RMC fits agree well with the experimental $S(Q)$ data, across the whole dynamic range. This matches the quality of fit achieved by Gereben and Pusztai,¹⁴ although we have fitted out to $\sim 24 \text{ \AA}^{-1}$, rather than $\sim 12 \text{ \AA}^{-1}$. It is noted that the quality of the fit to the experimental data in the low- Q region is not as good as that obtained by Gereben and Pusztai. This may result from our fitting over a wider Q range but is more probably caused by imposing coordination constraints that will preferentially fit over the short real-space distances and therefore will mean better fitting at intermediate range Q values. Figure 2 shows the $S(Q)$ differences between the data and the RMC, where it appears that above $\sim 15 \text{ \AA}^{-1}$, the oscillations in the RMC-generated plots are more heavily

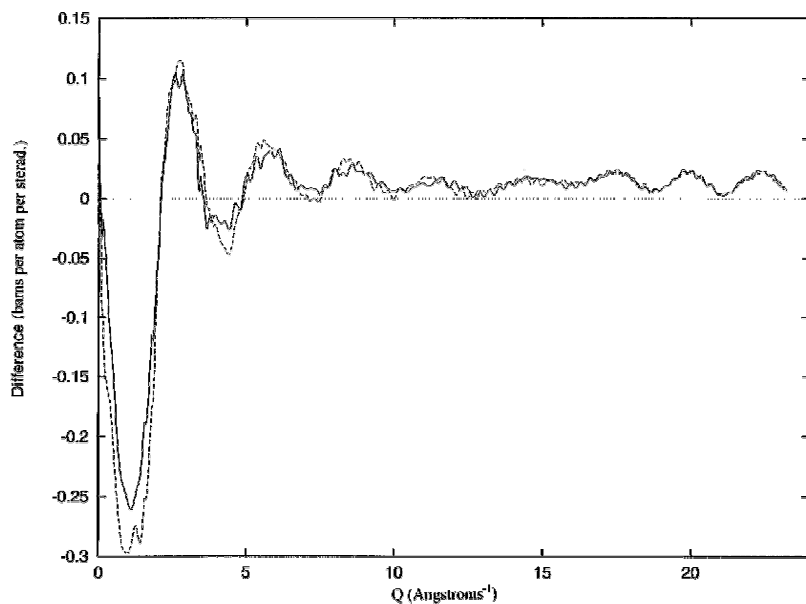


FIG. 2. Differences between the RMC fit and the experimental structure factor: low-density model (—) and high-density model (---).

damped. However, it should be noted that the experimental data drops off in this region, so that it no longer oscillates evenly about the axis, which the RMC $S(Q)$ is constrained to do. This could well be because of an error in correcting for hydrogen contamination and will not have any significant effect on the model. The overall task could be further improved if the Q range of the data were increased until the $S(Q)$ oscillations had more fully decayed towards their asymptotic limit; this would require additional experiments using a pulsed neutron source. The real-space $g(r)$ fits are also very good for both densities. This is in contrast to the random network models, which fail to reproduce the correct intensity of real-space features over more than a short range.

It is interesting to note that the quality of the fits is very similar for both densities. However, a good fit to the experimental data can be produced very easily using RMC, even though the resulting atomic configuration may show some peculiar, and physically and chemically implausible, arrangements when examined more closely at the atomistic level, so this agreement alone is not sufficient, and the models must be analyzed in further detail.

The nearest-neighbor distributions for the two models are given in Table I. The lower density gives an average first coordination number of 3.28, whereas for the higher density it has increased to 3.49 (c.f. the value determined by fitting

TABLE I. Coordination number distribution.

Number of neighbors	$n=0.03975 \text{ atoms } \text{\AA}^{-3}$	$n=0.04275 \text{ atoms } \text{\AA}^{-3}$
	Number of atoms	Number of atoms
0	0	0
1	2	2
2	95	73
3	1972	1370
4	931	1555
Average	3.28	3.49

Gaussian peaks to the experimental data of 3.68). So, a 7.5% increase in number density has resulted in a 6.4% increase in the average coordination number, but without producing any significant change in the fit to the experimental data. These coordination number distributions give 65% and 46% of Ge atoms with dangling bonds, respectively. The residual difference between the value obtained by peak fitting and those from the RMC probably arises from inaccuracies in correcting empirically for water contamination, and from the precise choice of maximum r values for the first-neighbor distance in the RMC model.

Figure 3 shows the Ge-Ge-Ge bond-angle distributions obtained from the two models. The main, well-defined peak in the distribution gives a mean bond angle at 109.4° (with a standard deviation of 8.5°) for both densities. This indicates that the network has a disordered tetrahedral character. The average bond angle determined from direct peak fitting to the data is $(108.5 \pm 1)^\circ$; the mean values from the random network models are in the range 109.2° – 109.5° . Both RMC models are both consistent with these previous results. Note

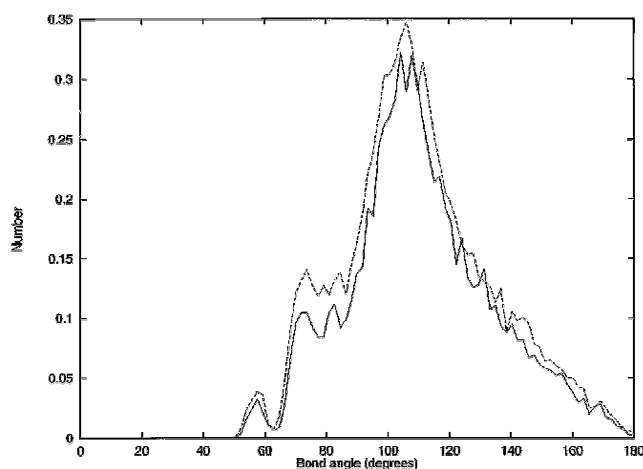


FIG. 3. Bond-angle distributions obtained from the RMC models: low-density model (—) and high-density model (---).

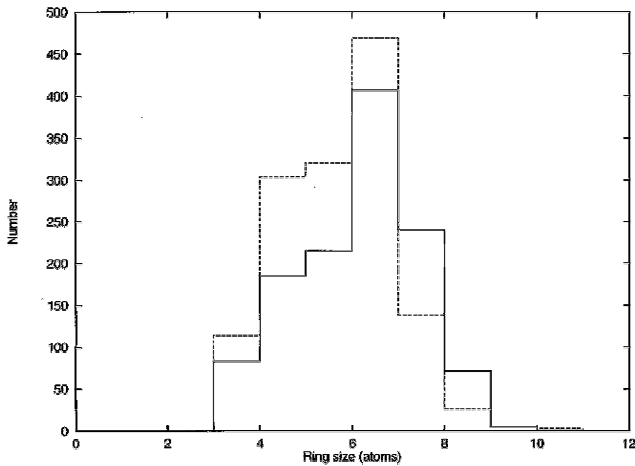


FIG. 4. Ring size distributions obtained from the RMC models: low-density model (—) and high-density model (---).

that the bond-angle distribution in the region $\sim 50\text{--}90^\circ$ is distorted by the triplet constraint, which discourages the formation of 60° bond angles but does nothing to reduce the probability of the formation of squares. The shape of the distribution is the same for both densities, but the “number” for the higher density is increased. This is consistent with more bonding in the network; i.e., a higher connectivity. The bond-angle distributions generated from the RMC models for *a*-Ge generated by Gereben and Pusztai¹⁴ show an unphysical sharp peak at $\sim 60^\circ$, except for their model constrained to yield 100% fourfold coordination starting from an initial box of atoms with the diamond structure. It is an important progression here that by using a larger box and by disallowing 60° angles and anything other than threefold and fourfold coordination we have been able to generate a model with a slightly narrower bond-angle distribution showing no serious unphysical characteristics and started from an initially random arrangement of atoms.

Finally, Figure 4 shows the ring statistics for the models. The most favored number of atoms in a ring is six, for both densities, which again indicates a model with disordered tetrahedral characteristics. There are also a few three-membered rings still present in the models. The order of probability for different ring sizes is six- \rightarrow seven- \rightarrow five- \rightarrow four-membered rings for RMC low density; six- \rightarrow five- \rightarrow four- \rightarrow seven-membered rings for RMC high density; and six- \rightarrow five- \rightarrow seven- \rightarrow eight-membered rings for random network models. It should be noted, however, that the number of four-membered rings in the RMC models is artificially high because of the effect of the triplet constraint to discourage three-membered rings. Taking this into account, the higher-density RMC model shows strong similarities to the random network models, although the fit to the full-range experimental $g(r)$ is much better for the RMC-generated model. The total number of rings is greater in the high-density model compared to the low-density one, which again shows an increased degree of bonding in the network and is, of course, consistent with the increased coordination number.

The main effects on the RMC model of increasing the number density are an increase in the average coordination number, the production of a more bonded network, and a change in the shape of the ring size distribution. These are

relatively subtle effects, which are not apparent simply by examining the fits to the experimental data. Both densities result in models that exhibit characteristics of a disordered tetrahedral network, which is also shown in earlier random network models, but, which of the two RMC models is “correct”? The model generated using the measured bulk density has an average coordination number lower than expected on the basis of direct analysis of the experimental data and gives a slightly different order of preference for ring sizes. Therefore, although the models are similar in all other respects, and the overall quality of their fits to the experimental data are the same, the higher density model is a better representation of the structure. This higher density also agrees with the value for the optimum density found by Gereben and Pusztai.¹⁴

One of Gereben’s and Pusztai’s conclusions is that the key characteristic of these structures is not the fraction of fourfold coordinated atoms but the fraction of tetrahedral angles. However, the two models presented here show almost identical bond-angle distributions and give the same fits to the experimental data, but have 31% and 52% fourfold coordination, respectively. This shows that the fraction of fourfold coordinated atoms remains one of the principal parameters in examining these structures; this is an especially important point, given the fact that no constraints were placed on this fraction in our models. Indeed, the work presented here shows that defects such as threefold coordinated atoms and four-membered rings, which are almost completely excluded by earlier random network models, are important features of the structure. Recent work by Gilkes, Gaskell, and Robertson²⁶ modeling the structure of tetrahedral amorphous carbon (*ta*-C) shows that the introduction of threefold coordinated sites into a tetrahedral network improved the agreement with the experimental $g(r)$ considerably. Wooten and Weaire²⁷ have also investigated the effects of introducing pair defects into their “bond switching” models; however, these also disallowed four-membered rings, so a direct comparison of the results is not strictly possible. They did, however, find that the introduction of such defects resulted in a large angular distortion within their models, but this problem was not evident in our RMC models. It is interesting to note in passing that the characteristics associated with a desirable kind of model outlined by Wooten and Weaire,²⁷ i.e., that the models should contain several hundred atoms at least, that the models should conform to periodic boundary conditions, and that the local tetrahedral bonding should not be too grossly distorted, are all satisfied by our RMC models. However, a model of 3000 atoms, i.e., ~ 42 Å, will not be able to reproduce microstructure (e.g., voids) over regions that are large in comparison with the RMC models. Indeed, such microstructural variation may prevent adequate fitting of the data in a single model, even of several thousand atoms, with periodic boundary conditions and other constraints. In the future, as computational power develops, models produced using hundreds of thousands of atoms will be a realistic possibility and may yield further insight into the structure of these materials.

IV. CONCLUSIONS

The RMC method has been used successfully to produce a model for the structure of amorphous germanium contain-

ing 3000 atoms. Chemically and physically unreasonable features have been discouraged by the application of constraints on coordination numbers and ring size. Models produced at two different densities fit the experimental data well, give a mean bond angle of 109.4° , and show six-membered rings to be the most probable. The main difference between these models is in the average coordination numbers (3.28 and 3.49 for the low and high densities, respectively), which reflects an increased amount of bonding in the model generated using a higher density, also apparent in the increase in the total number of rings. The RMC models show similarities to the random network models in that both exhibit characteristics of a disordered tetrahedral network, although the RMC models fit the experimental data more closely than any of the earlier random network models. This work also represents an improvement on previous RMC work.

Defects such as threefold coordinated atoms and four-membered rings have been shown to be important features in the structural model obtained for *a*-Ge, consistent with the experimental data. According to the criteria of Wooten and Weaire,²⁷ RMC provides a straightforward method of producing a realistic model structure.

ACKNOWLEDGMENTS

We are grateful to A. C. Wright and J. C. Dore for supplying the neutron diffraction data and to R. L. McGreevy for providing the original RMC code. Also, we would like to acknowledge D. W. Huxley and J. D. Wicks for developing and providing some of the additional code used for this work. Finally, we acknowledge the EPSRC for financial support.

-
- ¹D. E. Polk, *J. Non-Cryst. Solids*, **5**, 365 (1971).
²D. E. Polk and D. S. Boudreaux, *Phys. Rev. Lett.* **31**, 92 (1973).
³M. G. Duffy, D. S. Boudreaux, and D. E. Polk, *J. Non-Cryst. Solids* **15**, 435 (1974).
⁴G. A. N. Connell and R. J. Temkin, *Phys. Rev. B* **9**, 5323 (1974).
⁵D. L. Evans, M. P. Teter, and N. F. Borelli, *J. Non-Cryst. Solids* **17**, 245 (1975).
⁶D. Beeman and B. L. Bobbs, *Phys. Rev. B* **12**, 1399 (1975).
⁷F. C. Weinstein and E. A. Davis, *J. Non-Cryst. Solids* **13**, 153 (1973).
⁸D. Henderson, *J. Non-Cryst. Solids* **16**, 371 (1974).
⁹R. Grigorovici and R. Manaila, *J. Non-Cryst. Solids* **1**, 371 (1969).
¹⁰P. H. Gaskell, *Philos. Mag.* **32**, 211 (1975).
¹¹G. Etherington, A. C. Wright, J. T. Wenzel, J. C. Dore, J. H. Clarke, and R. N. Sinclair, *J. Non-Cryst. Solids* **48**, 265 (1982).
¹²R. L. McGreevy and L. Pusztai, *Mol. Sim.* **1**, 359 (1988).
¹³S. Kugler, L. Pusztai, L. Rosta, R. Bellisent, and P. Chieux, *Phys. Rev. B* **48**, 7685 (1993).
¹⁴O. Gereben and L. Pusztai, *Phys. Rev. B* **50**, 14 136 (1994).
¹⁵R. J. Newport, in *Neutron Scattering at a Pulsed Source*, edited by R. J. Newport, B. D. Rainford, and R. Cywinski (Hilger, London, 1988), Chap. 13, p. 233.
¹⁶P. A. V. Johnson, A. C. Wright, and R. N. Sinclair, *J. Non-Cryst. Solids* **58**, 109 (1983).
¹⁷J. Krogh-Moe, *Acta Crystallogr.* **9**, 951 (1956).
¹⁸N. Norman, *Acta Crystallogr.* **10**, 370 (1957).
¹⁹J. L. Yarnell, M. J. Katz, R. G. Wenzel, and S. H. Koenig, *Phys. Rev. A* **7**, 2130 (1973).
²⁰J. R. Beyster, *Nucl. Sci. Eng.* **31**, 254 (1968).
²¹R. L. McGreevy, M. A. Howe, D. A. Keen, and K. Clausen, *IOP Conference Proceeding No. 107* (Institute of Physics and Physical Society, London, 1990), p. 165.
²²M. A. Howe, R. L. McGreevy, and J. D. Wicks, *RMCA Version 3: A general purpose reverse Monte Carlo code*, October 1993 (unpublished).
²³J. K. Walters, J. S. Rigden, and R. J. Newport, *Phys. Scr.* **57**, 137 (1995).
²⁴J. K. Walters and R. J. Newport (unpublished).
²⁵D. W. Huxley, R. J. Newport, A. N. North, and J. K. Walters, in *Novel Forms of Carbon*, edited by C. N. Renschler, J. J. Pouch, and D. M. Cox, *MRS Symposia Proceedings No. 270* (Materials Research Society, Pittsburgh, 1992), p. 493.
²⁶K. W. R. Gilkes, P. H. Gaskell, and J. Robertson, *Phys. Rev. B* **51**, 12 303 (1995).
²⁷F. Wooten and D. L. Weaire, *J. Non-Cryst. Solids* **64**, 325 (1984).