INSTITUTE OF PHYSICS PUBLISHING MODELLING AND SIMULATION IN MATERIALS SCIENCE AND ENGINEERING

Modelling Simul. Mater. Sci. Eng. **12** (2004) 337–356 PII: S0965-0393(04)73231-3

Application of design of experiments methodology to optimization of classical molecular dynamics generation of amorphous SiO_2 structure

Evgueni Chagarov[1], James B Adams[1] and John Kieffer[2]

[1] Science and Engineering of Materials Program, Arizona State University, Tempe, AZ 85287-6006, USA

[2] Materials Science and Engineering Department, University of Michigan, Ann Arbor, MI 48109-2136, USA

Received 28 January 2003
Published 16 February 2004
Online at stacks.iop.org/MSMSE/12/337 (DOI: 10.1088/0965-0393/12/2/013)

Abstract

We introduce a design of experiments (DOE) methodology to the optimization of a-SiO_2 structures using classical molecular dynamics (MD) simulations. This approach produces defect-free systems in good agreement with experimental results of radial distribution functions and angular distribution functions. Application of DOE techniques allows systematic development of optimal MD simulation sequences of amorphous structures. A DOE approach allows control over the number of defects in the sample, as well as other properties, which may not be possible simply by maximizing annealing time and minimizing the cooling rate. In addition, DOE statistical analysis can give detailed insight into the dependences between preparation parameters and the properties of the samples so generated. For some response values, analytical forms have been fit, while for some other responses, accurate mapping on a space of affecting parameters has been performed. Although our investigation is restricted to the generation of a-SiO_2 structures, the approach is very general and can be effectively used for development and optimization of arbitrary amorphous systems.

1. Introduction

Amorphous materials play an important role in modern technology applications, and many researchers have carried out computer simulations of them using molecular dynamics (MD) [1–16], Monte Carlo (MC) [17, 18] and molecular modelling [19, 20] techniques. As a first step, any computer simulation must begin with preparation of a structural model, which should be as close in its physical and structural properties to the real system as possible. Whereas forming of ordered crystalline models is trivial, preparation of amorphous structures is challenging since the geometry of amorphous structures cannot be specified in advance.

0965-0393/04/020337+20$30.00 © 2004 IOP Publishing Ltd Printed in the UK

The structure of amorphous systems greatly affects their physical and chemical properties and therefore should be reproduced with high accuracy. The most important set of such attributes includes the number of dangling bonds, radial distribution functions (RDFs) and angular distributions functions (ADFs) for all types of pairs and angles, respectively.

We present an efficient and self-consistent application of design of experiments (DOE) techniques to optimization of a-SiO_2 structures using classical MD simulation, which finally produces a system in good agreement with experimental results for the above-mentioned attributes.

Although our investigation is restricted to generation of a-SiO_2 models, the proposed approach is very general and can be effectively used for development and optimization of arbitrary amorphous systems.

One of the most important applications of a-SiO_2 is the formation of inter-level dielectric layers in ULSI chips by controlled thermal oxidation of a Si surface. In addition, a-SiO_2 is important for chemical–mechanical polishing (CMP) of Si wafers since a thin layer of a-SiO_2 covers the Si wafer.

The properties of a-SiO_2 have been extensively investigated [1–22] using classical [1, 5, 8, 11, 12, 14, 21, 22] or *ab initio* MD simulations [15, 16] or a combination of classical and *ab initio* [13] MD simulations to generate a-SiO_2 model structures. Although the majority of the prepared models were satisfactory for the particular research, many of the preparation sequences suffered from a lack of a systematic approach and did not necessarily produce optimal structures. The application of DOE techniques enables one to systematically develop and optimize MD preparation sequences to produce structures in good agreement with experimental results. In addition, DOE statistical analysis can give detailed insight into dependences between preparation factors and the qualities of the generated sample as well as derive analytical forms of such dependences.

DOE is a widely and successfully used methodology in many fields of modern engineering. Originally, it was developed as a compilation of the theory of groups, linear algebra and statistics and remains a topic of intense research nowadays. A detailed description of it can be found elsewhere [23].

2. Underlying principles

SiO_2 exists in many different stable crystalline polymorphs such as α- and β-quartz, high- and low-cristobalite, coesite, stishovite, tridymite, keatite and others [24]. It is a strong glass former, and the silica glass obtained from the melt forms a random network of corner-sharing SiO_4 tetrahedra. In this structure, each Si atom is bound to four O neighbours in $Si(O_{1/2})_4$ coordination, while each O atom is connected to two Si atoms linking the tetrahedra in a continuous network. This arrangement was confirmed by diffraction experiments [25], reverse MC [26] and numerous MD simulations [1–22].

One of the standard ways of preparation of a-SiO_2 is MD quenching of the equilibrated SiO_2 melt. The general preparation sequence we investigated is shown in figure 1 and consists of four stages:

(a) heating with heating rate Rh up to a high temperature, Th, above the melting point of the system;
(b) equilibrating the system at Th for some time Lh;
(c) cooling with cooling rate Rc to room temperature;
(d) equilibrating at room temperature for some time Ll.

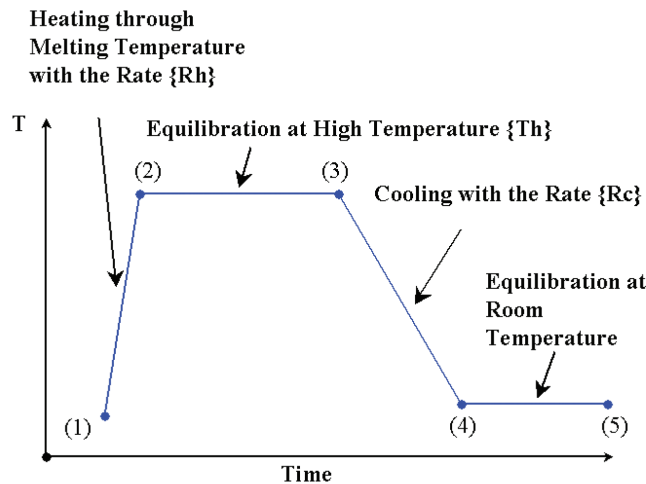

Figure 1. MD preparation sequence of a-SiO_2.

The set of five parameters {Rh, Th, Lh, Rc, Ll} and the initial structure entirely determine the preparation sequence and are targets of optimization.

The quality of prepared samples has been evaluated by the set of so-called 'response values', which are the following:

(1) number of defects, 'Ndef', which is the number of over-coordinated or under-coordinated Si or O atoms per sample;
(2) position of the main peak of the Si–O RDF—'$RDF_{Si–O}$';
(3) full-width half-maximum (FWHM) of the main peak of the Si–O RDF—'$FWHM_{Si–O}$';
(4) position of the main peak of the Si–Si RDF—'$RDF_{Si–Si}$';
(5) FWHM of the main peak of the Si–Si RDF—'$FWHM_{Si–Si}$';
(6) position of the main peak of the O–O RDF—'$RDF_{O–O}$';
(7) FWHM of the main peak of the O–O RDF—'$FWHM_{O–O}$';
(8) position of the main peak of the Si–O–Si ADF—'$ADF_{Si–O–Si}$';
(9) FWHM of the main peak of the Si–O–Si ADF—'$FWHM_{Si–O–Si}$';
(10) position of the main peak of the O–Si–O ADF—'$ADF_{O–Si–O}$';
(11) FWHM of the main peak of the O–Si–O ADF—'$FWHM_{O–Si–O}$'.

Each 'response value' is considered as a function of the parameters {Rh, Th, Lh, Rc, Ll} in parameter space. The general goal of our research is to optimize the set of parameters {Rh, Th, Lh, Rc, Ll} to generate structures that give the closest match of response values (2)–(11) to experimental results [29–31] and give Ndef $= 0$ since in real silica the concentration of defects is very low, around one per million. Optimization of all parameters is performed simultaneously by evaluating the approach of all response values to the reference ones and having different weights assigned to different response values to reflect their relative importance and typical variability.

The density of the prepared a-SiO_2 was not included in the set of response values since it was found that its value was quite close to experimental results for all runs. We suppose that density is mainly determined by inter-atomic potential, system temperature, external pressure and relaxation at the final temperature, which is long enough to allow the system to equilibrate its volume.

Table 1. Low- and high-level values of preparation parameters.

Parameter	Units	Low-level value	High-level value
Rh	$K\,fs^{-1}$	1.00	5.00
Th	K	5000.0	10 000.0
Lh	ps	40.0	80.0
Rc	$K\,fs^{-1}$	0.10	1.00
Ll	ps	20.0	60.0

The usual philosophy of glass preparation by MD simulations is to minimize the cooling rate, Rc, as much as possible and to maximize the annealing times, Ll and Lh. However, this approach cannot guarantee the absence of internal defects in the prepared sample, especially if annealing parameters are sub-optimal. It was found many times that runs with a low Rc and long Ll and Lh produced samples with internal defects. A DOE approach successfully solves this problem by searching for optimal annealing conditions. It generates accurate analytical forms for each response value and predicts the optimal parameters to obtain final properties in best agreement with experimental results. Although DOE techniques require multiple preliminary runs, the proper planning of these simulations can minimize this set, providing a higher or comparable total efficiency in comparison with the standard approach.

Each parameter is assigned a low- and a high-level value. These levels are chosen in a manner to avoid possible oscillations and bifurcation points in a region restricted by the low and high values of all parameters, so that the response value as a function is smooth, monotonous and single-valued inside that region. This requirement is not so critical and cannot violate the general DOE principles, though it provides higher reliability of prediction and subsequent optimization. Table 1 specifies low- and high-level values that were assigned to each parameter. The low- and high-level limits have been chosen according to general physical considerations and computational reasonability. There are different DOE schemes with different number of levels, but in our case a two-level scheme proved to have sufficient accuracy and reliability.

Five parameters assigned to five mutually orthogonal coordinate axes form a five-dimensional parameter space. In this space, points that correspond to the high- and low-level values of each parameter form the vertices of a five-dimensional parallelepiped. The general idea of the DOE methodology is to measure or simulate the magnitude of each response value at carefully chosen points in parameter space and to fit the response with an analytical form as a function of all parameters. Then, optimization methods can be used to search for the parameters that yield the closest simultaneous approach of all responses to the experimental values. Additional inclusion of some internal points like the centre point of the parallelepiped and the centres of its four-dimensional facets provides increasing accuracy and prediction reliability. A five-dimensional parallelepiped has $2^5 = 32$ vertices, and it can be time-consuming to perform simulations for each vertex in addition to possible central points. In this case, a so-called 'fractional factorial design' can be used, which requires simulations only on the fraction of all vertices, resulting in a 2^{5-n} ($n = 1, 2, 3, 4$) design. In this case, a special choice of vertices to simulate should be undertaken to minimize loss of accuracy and reliability in comparison with a full $2^5 = 32$ design. Details of fractional factorial design and choice of vertices to simulate can be found elsewhere [23].

The general methodology of DOE involves selecting a set of initial parameters (which probably can affect at least one of the response values) and then performing a so-called 'screening set of experiments', which in turn can identify which parameters are really important and which are not. Discarding unimportant parameters decreases the dimension of the

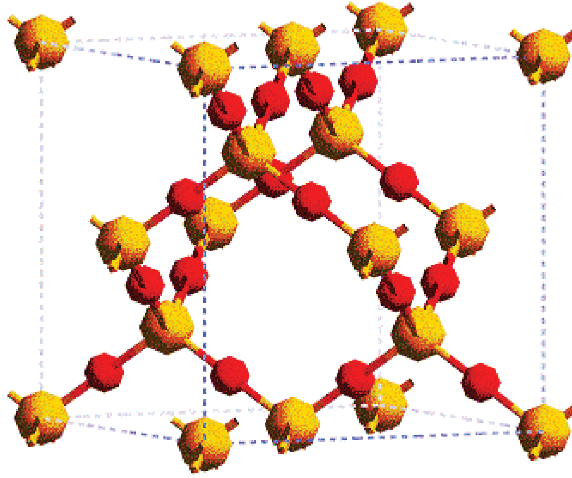

Figure 2. Unit cell of β-cristobalite.

parameter space and simplifies the optimization. Then a new set of simulations can be carried out, which will provide a maximum of information from a minimal number of simulations [23].

3. Computational procedure

Classical MD simulations have been performed according to the general preparation sequence of figure 1 with a specific set of parameters {Rh, Th, Lh, Rc, Ll} for each run. The initial system consisted of a 648-atom block of crystalline β-cristobalite (figure 2) relaxed for 20 ps prior to the simulations. Besides atom coordinates and velocities, the code utilized in its input file requires some data for calculation of charge–dipole and dipole–dipole interactions, which is not known prior to simulations but obtained through a self-consistent iterative procedure. During the first several hundred time steps, the code acquires self-consistency, which results in large forces on atoms and distortion of the initial β-cristobalite structure. Nevertheless, after relaxation, the initial sample did not have any defects. All simulations have been performed in three-dimensional periodic boundary conditions at a constant pressure of 1 atm controlled by the Andersen scheme [27]. The inter-atomic interactions were modelled by Kieffer's classical potential, which included a two-body part, a three-body angular term (equation (1)) and charge–dipole and dipole–dipole interaction terms [28].

$$\phi_i = q_i \sum_{j=1}^{N} \frac{q_j}{4\pi\varepsilon_0 r_{ij}} + \sum_{i=1}^{NC} A_{ij}\left(1 + \frac{z_i}{n_i} + \frac{z_j}{n_j}\right) \mathrm{e}^{(\sigma_i+\sigma_j-r_{ij})} + \sum_{j=1}^{NC-1}\sum_{k=j+1}^{NC} (\varphi_{ij} + \varphi_{ik})\mathrm{e}^{-\gamma_{ijk}(\bar{\theta}-\theta_{ijk})^2} \tag{1}$$

The temperature was controlled by rescaling the velocities of all atoms each time step. RDF and ADF curves for all types of pairs and angles have been calculated for the final system geometry, taking into account periodic boundary conditions. The correspondent RDF curves were calculated on a basis of ~42 000 Si–O pairs, ~21 000 Si–Si pairs and ~84 000 O–O ones. The Si–O–Si ADF curves were calculated for the ~860 Si–O–Si triples, while the O–Si–O ADF curves were calculated for the ~2500 O–Si–O triples. The main RDF and ADF peaks have been fit by Gaussian functions to eliminate some collateral noise and to determine the peak position and its FWHM as response values. To investigate the importance of time averaging,

Table 2. Chart of preparation simulation parameters. Plus (+) or minus (−) signs in parameter columns specify a high or low level of the parameter to be used. A sign (0) denotes an average of low- and high-level values of the parameter. For each run, a complete set of 11 response values has been obtained, but only 'Ndef' is presented here.

Run number	Rh	Th	Lh	Rc	Ll	Ndef
1	(+1)	(+1)	(+1)	(−1)	(−1)	2
2	(−1)	(−1)	(+1)	(−1)	(−1)	0
3	(−1)	(+1)	(−1)	(−1)	(−1)	2
4	(+1)	(−1)	(+1)	(+1)	(−1)	0
5	(−1)	(+1)	(−1)	(+1)	(+1)	12
6	(+1)	(+1)	(−1)	(+1)	(−1)	6
7	(+1)	(−1)	(−1)	(+1)	(+1)	0
8	(+1)	(−1)	(+1)	(−1)	(+1)	0
9	(−1)	(−1)	(−1)	(−1)	(+1)	0
10	(+1)	(+1)	(−1)	(−1)	(+1)	2
11	(−1)	(−1)	(−1)	(+1)	(−1)	0
12	(+1)	(+1)	(+1)	(+1)	(+1)	6
13	(−1)	(−1)	(+1)	(+1)	(+1)	0
14	(−1)	(+1)	(+1)	(+1)	(−1)	8
15	(+1)	(−1)	(−1)	(−1)	(−1)	0
16	(−1)	(+1)	(+1)	(−1)	(+1)	0
17	(−1)	(0)	(0)	(0)	(0)	12
18	(+1)	(0)	(0)	(0)	(0)	6
19	(0)	(−1)	(0)	(0)	(0)	0
20	(0)	(+1)	(0)	(0)	(0)	2
21	(0)	(0)	(−1)	(0)	(0)	2
22	(0)	(0)	(+1)	(0)	(0)	4
23	(0)	(0)	(0)	(−1)	(0)	2
24	(0)	(0)	(0)	(+1)	(0)	4
25	(0)	(0)	(0)	(0)	(−1)	2
26	(0)	(0)	(0)	(0)	(+1)	2
27	(0)	(0)	(0)	(0)	(0)	2

we compared the results obtained by averaging the response at five different times over the last 500 fs of the simulation. This did not significantly change the averaged response because the final structures contained a large number of pairs and triplets in the investigated ensemble.

In addition, the final geometry was investigated to determine the number of defects, Ndef. Each under-coordinated or over-coordinated Si and O atom was counted as an internal defect. The DOE software package Design Expert 6.0 and Interactive Data Language (IDL) were used for construction of the simulation set, statistical analysis, fitting and factor optimization.

4. Preparation of a-SiO_2 sample

4.1. Framework simulations

Initially a 'screening' set of 16 simulations according to a 2^{5-1} scheme of fractional factorial design [23] was performed to determine unimportant parameters that did not affect any response value and thus could be eliminated from further consideration. This set of 16 MD simulations in fact was determined by half the vertices of a five-dimensional parallelepiped in parameter space and corresponded to the first 16 runs of table 2. The symbols (+1) and (−1) in parameter

Table 3. Variability of all response values on the basis of the first 27 runs.

	Mean ($\bar{\nu}$)	Std. dev (σ)	$\sigma/\bar{\nu} \times 100\%$
Ndef	2.81	1.95	69.395
$RDF_{Si–O}$ (Å)	1.59	9.698×10^{-4}	0.061
$FWHM_{Si–O}$ (Å)	0.081	5.587×10^{-3}	6.898
$RDF_{Si–Si}$ (Å)	3.05	4.363×10^{-3}	0.143
$FWHM_{Si–Si}$ (Å)	0.16	0.013	8.125
$RDF_{O–O}$ (Å)	2.59	2.135×10^{-3}	0.082
$FWHM_{O–O}$ (Å)	0.29	0.014	4.828
$ADF_{Si–O–Si}$ (deg)	146.34	0.40	0.273
$FWHM_{Si–O–Si}$ (deg)	12.44	0.56	4.502
$ADF_{O–Si–O}$ (deg)	108.66	0.19	0.175
$FWHM_{O–Si–O}$ (deg)	15.65	0.77	4.920

columns of table 2 specify either a high or low level of the parameter (table 1). Half the vertices were chosen among all the ones given by the constructor $I = \mathrm{Rh} * \mathrm{Th} * \mathrm{Lh} * \mathrm{Rc} * \mathrm{Ll}$, which is explained elsewhere [23]. In the DOE approach, the fraction of vertices to be simulated is chosen in such a way as to decrease the number of runs to be performed while maximizing the model robustness. In many cases, when system properties do not exhibit fast oscillations on preparation parameters, the reduction of the number of runs leads to little loss of accuracy [23]. The $2^{5-1} = 16$ runs should not be considered as the smallest set of required simulations, and in many cases the investigation can be performed using $2^{5-2} = 8$ or even fewer runs.

For each performed run, all response values were obtained as described above and became the target of statistical analysis, which revealed that each parameter or its product with another parameter affected at least one response value and therefore had to be preserved. Then the initial design was augmented by an additional 11 runs (numbers 17–27 in table 2), which corresponded to ten centres of ten facets and to one central point of the five-dimensional parameter parallelepiped. The addition of the parallelepiped centre and the centres of its facets increases the accuracy and reliability of the model. The symbol (0) in the parameter columns of table 2 denotes an average of low- and high-level values of the parameter. Table 2 also presents one of the obtained response values, Ndef, as a most important and dynamic one. The average magnitudes of the response values and their absolute and relative standard deviations are summarized in table 3.

4.2. Optimization and validation run

The magnitudes of each response value obtained in the previous 27 framework runs were fit by a quadratic five-dimensional polynomial as a function of preparation parameters, producing a separate analytical form for each response. These analytical forms are used to obtain the set of optimal preparation parameters {Rh, Th, Lh, Rc, Ll}, that are expected to produce the sample in the best possible agreement with experimental results. The set of optimal parameters, {Rh, Th, Lh, Rc, Ll}, is obtained by simultaneous optimization along the fit response analytical forms evaluating the total deviation of the predicted properties from the experimental ones. The set of optimized preparation parameters is presented in table 4 together with low- and high-level parameter extremes, which restrict the possible argument values and make the prediction more reliable.

A final MD simulation with the optimized preparation parameters was performed to validate the prediction of the DOE analysis. The set of experimental reference values together

Table 4. Optimized parameter values. Parameter extreme points.

Parameter	Lower limit	Upper limit	Optimized parameter	Optimized coded parameter
Rh (K fs^{-1})	1.0	5.0	4.37	0.6850
Th (K)	5 000.0	10 000.0	5 292.43	−0.8830
Lh (fs)	40 000.0	80 000.0	58 049.70	−0.0975
Rc (K fs^{-1})	0.1	1.0	0.39	−0.3555
Ll (fs)	20 000.0	60 000.0	20 525.18	− 0.9737

Table 5. Experimental references for each response value [29–31]. Simulated response values correspond to the final simulation with optimized preparation parameters. Predicted response values are magnitudes of the fit analytical forms for each response with optimized preparation parameters as arguments. Parameter α defines deviation of the MD-prepared silica sample properties from the experimental references, estimating the general sample quality. Parameter β determines deviation of the analytically predicted properties from the simulated ones, estimating the general model accuracy.

Response value	Weight	Experimental reference	Simulated response	Predicted response	α (%)	β (%)
Ndef	5	0	0	2.222×10^{-11}		
$RDF_{Si–O}$ (Å)	3	1.610	1.5900	1.5920	1.24	0.13
$FWHM_{Si–O}$ (Å)	3	0.050	0.0687	0.0690	37.4	0.44
$RDF_{Si–Si}$ (Å)	3	3.080	3.0651	3.0624	0.48	0.09
$FWHM_{Si–Si}$ (Å)	3	0.100	0.1086	0.1180	8.60	8.66
$RDF_{O–O}$ (Å)	3	2.632	2.5991	2.5995	1.25	0.02
$FWHM_{O–O}$ (Å)	3	0.089	0.1466	0.1868	64.72	27.42
$ADF_{Si–O–Si}$ (deg)	3	142.00	147.8649	146.827	4.13	0.70
$FWHM_{Si–O–Si}$ (deg)	5	26.0	9.1438	9.4784	64.83	3.66
$ADF_{O–Si–O}$ (deg)	3	109.7	108.9844	109.108	0.65	0.11
$FWHM_{O–Si–O}$ (deg)	5	10.6	8.2012	10.5648	22.63	28.82

with simulated and predicted response values are presented in table 5. The simulated response values of table 5 correspond to the final simulation with the optimized preparation parameters presented in table 4. Predicted responses (table 5) are magnitudes of the functional forms fit for each response value with the optimized preparation parameters as arguments (table 4). The parameter α in table 5 is defined as $\alpha = |((R_{exp} - R_{sim})/R_{exp})| \times 100\%$ and determines the relative deviation of optimal MD-prepared silica properties from experimental ones. Parameter β is defined as $\beta = |((R_{sim} - R_{pred})/R_{sim})| \times 100\%$ and determines the relative deviation of the model-predicted silica properties from simulated ones estimating the general model accuracy. During optimization, each response value had an assigned weight (table 5) that reflected the importance and variability of the specific value in comparison with other ones. In the cases of $FWHM_{Si–O–Si}$ and $FWHM_{O–Si–O}$, the choice of such a high weight is dictated by the high relative variability of these values.

The set of 27 framework simulations is enough to perform reliable optimization and to obtain a satisfactory match with experimental results. A further increase in the number of performed simulations will likely provide only a minor improvement in the results.

4.3. Comparative analysis. Accuracy and restrictions of the method

DOE analysis provides an alternative approach to the a-SiO_2 MD preparation sequence and it is important to evaluate its advantages and disadvantages in comparison with the standard

approach. Under the 'standard approach', we imply the method of maximizing the annealing and relaxation times and minimizing the cooling rate.

In the standard approach, Th is usually chosen in an arbitrary manner. It does not always eliminate defects in the prepared sample, which can be present even after very long runs. However, for many cases the absence of internal defects in samples is the critical requirement. In many cases, the preparation of a silica sample by the standard approach can include weeks and even months of simulations. The DOE approach requires more simulations, but these runs can be relatively short.

The advantage of the DOE approach is that it provides more control over the final properties of the sample including the number of defects. The analytical models derived from the framework runs are quite accurate to predict the optimal preparation parameters, which will produce a defect-free sample in reasonable agreement with the other experimental properties like RDF and ADF peaks. In addition, the derived analytical dependences of the DOE model can provide deep insight into the physics of this process, which will be demonstrated in section 5.

Analysis of the values in table 5 shows that the MD-prepared silica reproduces experimental results with various degrees of accuracy.

Most importantly, the obtained sample has no internal defects, which is in complete accordance with the simulation goal and the analytical prediction.

The $RDF_{Si\text{–}O}$, $RDF_{Si\text{–}Si}$ and $RDF_{O\text{–}O}$ response values demonstrate very small variability from one run to another (table 3). They are generally very close to the experimental results, which primarily is a consequence of accurate fitting of the inter-atomic potential. In addition, analytic expressions fit to the first 27 runs turn out to be quite accurate in their predictions, which is reflected by the small magnitudes of parameter β in table 5.

The $FWHM_{Si\text{–}O}$, $FWHM_{O\text{–}O}$ and $FWHM_{Si\text{–}Si}$ demonstrate a larger variability from one run to another than the corresponding RDF peak positions (table 5). The simulated response values are greater than the experimental ones, especially $FWHM_{O\text{–}O}$. The discrepancy between experimental and simulated FWHM values is probably caused by the two-body part of the inter-atomic potential being too soft near the point of equilibrium.

The $ADF_{Si\text{–}O\text{–}Si}$ and $ADF_{O\text{–}Si\text{–}O}$ values are very close to experimental results, which is primarily a consequence of the accuracy of the potential's three-body part. The particular fit analytical forms demonstrate high prediction accuracy (table 5).

The simulated silica sample has somewhat lower $FWHM_{Si\text{–}O\text{–}Si}$ and $FWHM_{O\text{–}Si\text{–}O}$ values than experimental ones (table 5). However, it should be taken into account that the experimental values of FWHM include broadening due to measurements of the neutron or x-ray structure factor and its subsequent Fourier transformation. Thus, it is difficult to reliably compare the experimental and simulated FWHM for the angles, but we try to optimize it as best as we can.

The quadratic analytical model of $FWHM_{O\text{–}Si\text{–}O}$ demonstrates a higher deviation from the simulated value. This property will be investigated later in a more detailed manner, providing increased accuracy of analytic description. In addition, it will be shown that the deviation of the simulated $FWHM_{Si\text{–}O\text{–}Si}$ from the experimental result is caused by the trade-off with the Ndef response value, which as a rule has a high magnitude at the region where $FWHM_{Si\text{–}O\text{–}Si}$ would be closer to the experimental reference value.

Despite its efficiency and predictive power, application of the DOE methodology to the generation of a-SiO_2 sample has one natural limitation. This limitation is the inter-atomic potential, which predetermines many properties of the sample in its stable final state, decreasing the effect of the preparation sequence. Probably, the number of defects is the only response value that is mainly determined by the preparation sequence but not by the inter-atomic potential itself. The other response values are primary determined by the inter-atomic potential and to

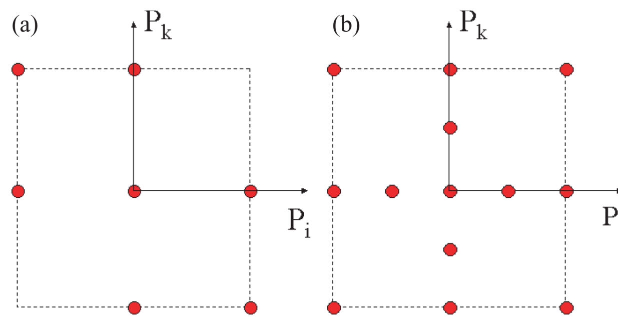

Figure 3. (*a*) Design graph before second augmentation. (*b*) Design after second augmentation. Previously omitted vertices of a five-dimensional cube and internal five-dimensional cross are added. The point of optimal parameters and its inversion have been included too, though they are not shown.

a lesser extent by the preparation sequence. Despite this limitation, we believe that the DOE approach allows a controlled search for the best possible correlation with the experimental requirements, within limitations imposed by the inter-atomic potential.

This paragraph completes the actual sample preparation section. The next sections will mainly concentrate on the more fundamental investigation of how the annealing sequence affects the structure and will require some additional simulations to be performed.

5. Analytical forms of response values

Analytical forms for the response values are able to provide insight into the amorphization process and a-SiO_2 network formation. Although 27 runs are enough to produce an a-SiO_2 system in good agreement with experiment, for more fundamental investigation it is desirable to expand the set of simulations to increase the accuracy of the functional forms of response values. Thus the initial design was augmented as shown in figure 3(*b*) by including 16 previously omitted vertices of the five-dimensional cube and internal five-dimensional cross of ten points to clarify trends close to the central point. In addition, the point obtained by the previous optimization with $\{\mathrm{Rh, Th, Lh, Rc, Ll}\} = \{4.37\,\mathrm{K\,fs^{-1}}, 5292.43\,\mathrm{K}, 58\,049.70\,\mathrm{fs}, 0.39\,\mathrm{K\,fs^{-1}}, 20\,525.18\,\mathrm{fs}\}$ as well as its inversion relative to the central point were included, giving a total of 55 runs. For each additional run, all 11 response values were obtained and fit together with the previous 27 runs. It is useful to represent the fit response value functions in terms of so-called 'coded parameters', which are defined as

$$P_{\mathrm{coded}} = -1 + 2\frac{P - P_{\mathrm{min}}}{P_{\mathrm{max}} - P_{\mathrm{min}}} \tag{2}$$

where P_{max} and P_{min} are the high- and low-level values of the parameter as stated in table 1.

Coding the parameters normalizes all of them with respect to each other by projecting segment $[P_{\mathrm{min}}; P_{\mathrm{max}}]$ to segment $[-1.0; 1.0]$. The representation of the response value analytical expression in terms of coded parameters normalizes the coefficients of each term with respect to the other ones, which in turn clarifies the contribution of each term to the selected response value.

The accuracy of the fit analytical model can be limited if the selected response value has a low variability. For example, if the RDF or ADF main peaks have a very narrow distribution, and that distribution is discretized on a fine but finite grid, then it is hard to fit the amplitudes exactly. If those RDFs and ADFs also depend weakly on the parameters, then it is difficult to

Table 6. Variability of all response values on the basis of 55 runs.

	Mean ($\bar{\nu}$)	Std. dev (σ)	$\sigma/\bar{\nu} \times 100\%$
Ndef	2.64	1.52	57.57
$RDF_{Si–O}$ (Å)	1.60	1.495×10^{-3}	0.09
$FWHM_{Si–O}$ (Å)	0.083	4.946×10^{-3}	5.96
$RDF_{Si–Si}$ (Å)	3.06	3.794×10^{-3}	0.12
$FWHM_{Si–Si}$ (Å)	0.16	0.015	9.38
$RDF_{O–O}$ (Å)	2.60	3.145×10^{-3}	0.12
$FWHM_{O–O}$ (Å)	0.29	0.030	10.34
$ADF_{Si–O–Si}$ (deg)	146.85	0.39	0.27
$FWHM_{Si–O–Si}$ (deg)	12.51	1.32	10.55
$ADF_{O–Si–O}$ (deg)	109.11	0.20	0.18
$FWHM_{O–Si–O}$ (deg)	16.01	1.73	10.81

optimize those response values. On the other hand, response values that manifest enhanced relative variability can be satisfactorily fit by quadratic multidimensional polynomials. The variability of all 11 response values based on 55 runs is summarized in table 6.

The following analytical expressions for several response values in terms of coded factors were obtained with satisfactory accuracy by least-squares fitting of a quadratic five-dimensional polynomial:

$$FWHM_{Si–Si} = 0.19 + 0.035 * Th - 0.014 * Rh * Rh - 0.041 * Th * Th + 0.011 * Lh * Lh + \varepsilon \quad (3)$$

$$FWHM_{O–O} = 0.36 + 0.092 * Th - 0.017 * Rh * Rh - 0.1 * Th * Th + 0.016 * Lh * Lh + 0.02 * Th * Rc + \varepsilon \quad (4)$$

$$ADF_{Si–O–Si} = 146.32 - 0.71 * Th - 0.1 * Rc + 0.75 * Th * Th + 0.15 * Lh * Lh + 0.16 * Rc * Rc - 0.25 * Ll * Ll + 0.17 * Rh * Rc - 0.28 * Th * Rc + 0.12 * Lh * Rc + \varepsilon \quad (5)$$

For equations (3)–(5), the FWHMs are measured in angstroms, while the main ADF peak position is measured in degrees. The parameters being 'coded' belong to segment [−1.0; 1.0]. Although $FWHM_{Si–Si}$, $FWHM_{O–O}$ and $ADF_{Si–O–Si}$ were fit initially to the same functional form, some coefficients were so much less than the other ones that they were discarded, resulting in somewhat simpler final functional forms in equations (3)–(5).

The fit of $RDF_{Si–O}$, $FWHM_{Si–O}$, $RDF_{Si–Si}$ and $RDF_{O–O}$ by quadratic five-dimensional polynomials does not yield a physically meaningful analytical form because of the very low variability of those responses. Those properties were not significantly affected by variations in the simulation parameters and were primarily controlled by the inter-atomic potential. As seen in table 5, the obtained $RDF_{Si–O}$, $FWHM_{Si–O}$, $RDF_{Si–Si}$ and $RDF_{O–O}$ were in good agreement with the experimental properties as well as highly accurate in their analytical predictions, which emphasizes the robustness of the DOE approach.

The $ADF_{O–Si–O}$ has a larger standard deviation (table 6), but it would require a more complex analytical form than a five-dimensional quadratic polynomial to obtain a more physically adequate fit. However, table 5 demonstrates a high degree of $ADF_{O–Si–O}$ quality ($\alpha = 0.65\%$) as well as accurate analytical prediction of the simulated response value ($\beta = 0.11\%$) derived from the 27 framework runs.

The obtained expressions in equations (3) and (4) have a clear physical sense. It is remarkable that both $FWHM_{Si–Si}$ and $FWHM_{O–O}$ have identical functional forms including

identical signs for the respective terms. The only exception is a presence of an additional component in equation (4). The negative quadratic term, $Rh * Rh$, in equations (3) and (4) denotes that a low heating rate slightly increases smearing of Si–Si and O–O bond length distributions, while a fast heating narrows the final bond length distribution. This demonstrates that at the end of the 5th stage of thermal treatment, the structure is only slightly affected by its first stage. The normalized coefficients of $Rh * Rh$ terms in equations (3) and (4) are very small, which indicates a rather weak effect and a very limited degree of influence of the initial stage. The detection of this result is due to the high sensitivity of DOE statistical analysis.

The contribution of Th in equations (3) and (4) is represented by a quadratic expression that has a maximum at $Th_{(max)} = 0.44$ (coded units) = 8600 (K) for Si–Si and at $Th_{(max)} = 4.6$ (coded units) for O–O bonds. The fact that the Th contribution to $FWHM_{O–O}$ reaches its maximum at $Th_{(max)} = 4.6$ denotes steady smearing of O–O bond length distribution as Th grows from the low to the high level. The evolution of $FWHM_{Si–Si}$ is more unusual. Initially it grows up to $Th_{(max)} = 0.44$ (coded units) = 8600 (K), but after that it begins to decrease slowly. A physical interpretation of this phenomenon is not so clear, though the statistical accuracy of the fit justifies the reliability of this result. A similar trend was found for $FWHM_{Si–O–Si}$ and $FWHM_{O–Si–O}$ response values and will be discussed later.

The presence of the positive term $Lh*Lh$ in both equations (3) and (4) denotes that keeping the system at Th increases the bond length smearing.

Interpretation of equation (5) is more complex because of the large number of mixed terms. The positive quadratic term, $+0.15 * Lh * Lh$, denotes expansion of the Si–O–Si angle after keeping the system at Th. The negative quadratic term, $-0.25*Ll*Ll$, in turn reflects the shrinking of the angle being equilibrated at room temperature. The component $+0.16*Rc*Rc$ denotes the fact that in the case of fast cooling the system does not have enough time to relax completely and as a consequence the Si–O–Si angle remains slightly expanded, which is transferred into the solid a-SiO_2 network from the high temperature melt. The contribution of Th is less trivial. The quadratic component $(-0.71 * Th + 0.75 * Th * Th)$ has a minimum at $Th_{(min)} = 0.47$ (coded units) = 8675 (K), which is very close to the $Th_{(max)} = 0.44$ (coded units) = 8600 (K) for $FWHM_{Si–Si}$. In general, this term describes initial decreasing of the Si–O–Si angle, which later begins to grow. This transition temperature may denote some structural transformation occurring in the system.

To investigate the robustness of the optimal preparation parameters, the new set of parameters was optimized on the basis of 55 runs and compared with the optimal parameters obtained from the first 27 runs (table 4). The new set of optimal parameters demonstrated very limited deviation from the first set presented in table 4. The Rh, Rc and Th deviated within 4%, while Lh and Ll deviated within 7%. This small deviation validates the smaller model. This demonstrates that the correct application of the fractional factorial design provides minimal loss of quality and reliability with a decreased number of performed runs [23].

The other response values, Ndef, $FWHM_{Si–O–Si}$ and $FWHM_{O–Si–O}$, exhibit quite significant absolute and relative variability, and taking into account their importance will be investigated in a more detailed manner below.

6. Dependences of FWHM of ADF

One of the most important and powerful capabilities of DOE analysis is the ability to clearly identify which parameters or their interactions affect the selected response values and which do not. This is accomplished by evaluating the ratio of the mean squares of the response value and the statistical noise [23]. Such information can be obtained either by drawing normal probability plots, which provides clear visual though slightly subjective interpretation,

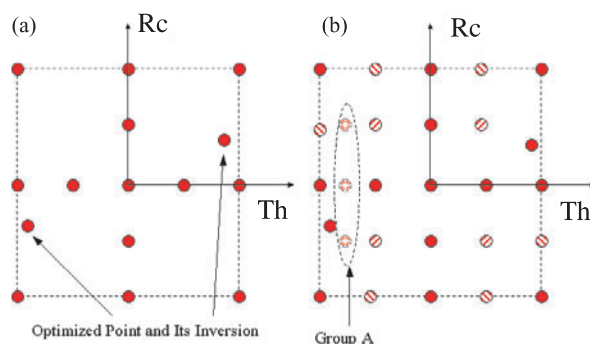

Figure 4. (*a*) Initial pattern, which is a projection of five-dimensional cube and its internal points to a two-dimensional plane. (*b*) Pattern after augmentation. Added points are pattern-filled ones. Group A is a special set of points.

or by a purely statistical computational procedure, which is more accurate [23]. Statistical analysis of $FWHM_{Si–O–Si}$ and $FWHM_{O–Si–O}$ unambiguously shows that these two response values depend only on Th and Rc and do not depend on the other parameters. This conclusion is rather powerful since it reduces a five-dimensional model to a two-dimensional one, greatly decreasing the number of required runs. Projection of the five-dimensional cube and internal points to the {Th, Rc} surface reveals that variability between points with the same {Th, Rc} is much less than between points with different {Th, Rc} pairs, which confirms the above-mentioned conclusion. After projection of the current design to the {Rh = 0, Th, Lh = 0, Rc, Ll = 0} plane, where parameters are 'coded' ones, the obtained two-dimensional design has been augmented by an additional 13 runs as is shown in figure 4. The initial pattern in figure 4(*a*) has 15 points in the {Th, Rc} plane, while the augmented pattern in figure 4(*b*) has 28 points. This expansion increases the predictive capability and accuracy of the analytical fitting. The reason for including the so-called Group A points in figure 4(*b*) will be clarified later. For an additional 13 runs corresponding to the 13 new points on the pattern, $FWHM_{Si–O–Si}$, $FWHM_{O–Si–O}$ and Ndef were calculated and analysed together with previous data. The resulting sets of 28 distributed points for $FWHM_{Si–O–Si}$ and $FWHM_{O–Si–O}$ were initially interpolated by a minimum curvature surface algorithm to the regular mesh with step of 0.05 and after that were fit to two-dimensional polynomial of sixth degree. Polynomials of lower degrees have been found to give a less accurate description of the simulated data. The interpolated surface of the $FWHM_{O–Si–O}$ response value is presented in figure 5.

The most important observation from figure 5 is that $FWHM_{O–Si–O}$ has a sharp increase in magnitude along some threshold temperature range. Figure 6 shows a set of curves of $FWHM_{O–Si–O}$ that correspond to several fixed Rc values. The temperature range of such transitions is almost independent of Rc and is limited by Th $\sim$ [−0.9; −0.3] (coded units) or Th $\sim$ [5250.0; 6750.0] (K). The presence of that sharp barrier explains the inclusion of Group A points shown in figure 4(*b*), which allows a rather accurate representation of that sharp transition.

A similar procedure was used for the set of $FWHM_{Si–O–Si}$ values. The interpolated response surface is shown in figure 7. In general, the surface of $FWHM_{Si–O–Si}$ is similar to the surface of $FWHM_{O–Si–O}$, though it has one peculiarity. $FWHM_{Si–O–Si}$ has an anomalous increase at {Th, Rc} = {0.5, 0.0} (coded units) or {Th, Rc} = {8750 (K), 0.55 (K fs^{-1})}, which is absent on the $FWHM_{O–Si–O}$ surface. That point has been double-checked and confirmed. Appearance of that hill can imply formation of different specific structural phases with enlarged smearing of the Si–O–Si angle and increasing of the total entropy of the system.

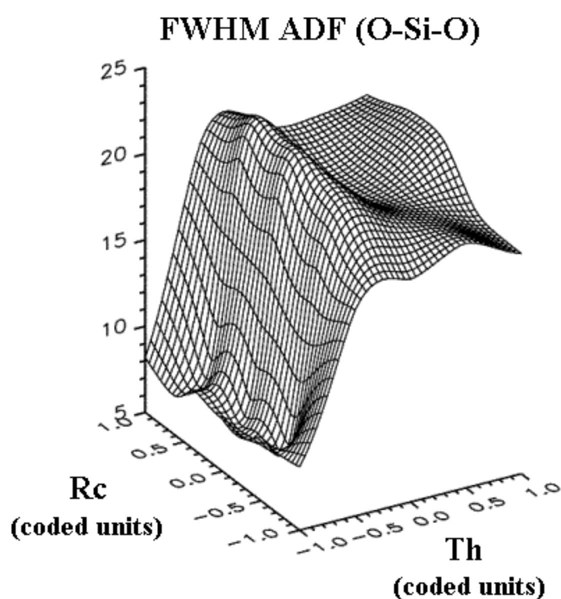

Figure 5. Surface of $FWHM_{O-Si-O}$ as a function of {Th, Rc}.

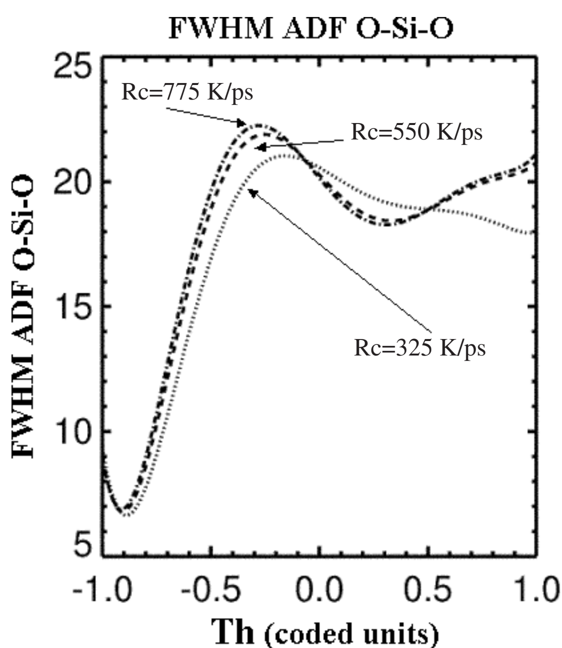

Figure 6. $FWHM_{O-Si-O}$ as a function of Th at several different Rc values.

The temperature range of the $FWHM_{Si-O-Si}$ threshold increasing was evaluated as before by plotting $FWHM_{Si-O-Si}$ as a function of Th at several fixed Rc values, which gave the same range of transition temperatures, Th $\sim$ [5250.0; 6750.0] (K), as before, which can be explained by the fact that both Si–O–Si and O–Si–O angles belong to the same geometrical system and condensed from the same liquid phase.

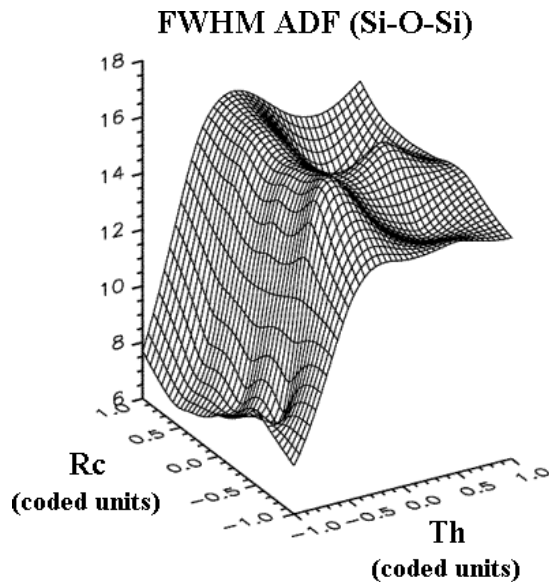

Figure 7. Surface of $FWHM_{Si-O-Si}$ response value as a function of {Th, Rc}.

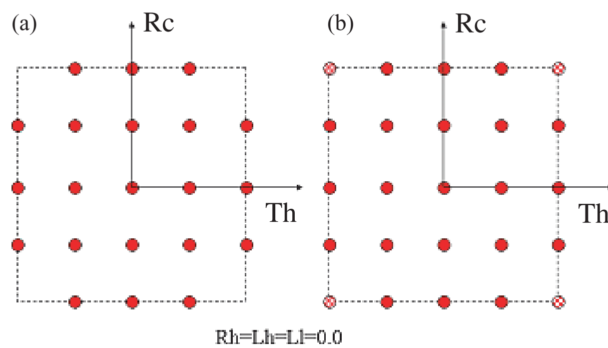

Figure 8. (*a*) Cross-section of previous 68 runs set by Rh = Lh = Ll = 0.0 hypersurface; (*b*) Augmentation by four corner runs to complete the set.

7. Number of defects

The number of defects is the most dynamically changing and important response value. In real silica, the normal concentration of internal defects is very low, around one per million atoms, and so for a 648-atom model it is reasonable to consider the absence of any defects as the most realistic case. The number of defects has been determined for each of 68 runs. DOE statistical analysis unambiguously reports that Ndef depends on Rh, Th, Rc and slightly on Ll but does not depend on Lh, which denotes that even a low level of Lh (40 ps) is enough to completely equilibrate the system at that condition. Even after eliminating Ll, the dimension of the parameter space (4) is high enough to require a lot of additional runs to fill the internal space of a four-dimensional hypercube with a fine mesh to obtain an accurate functional form of Ndef. Therefore, it was decided to restrict analysis to the two-dimensional case where {Th, Rc} can vary, while {Rh, Ll} are kept at zero values in coded terms. Figure 8(*a*) shows

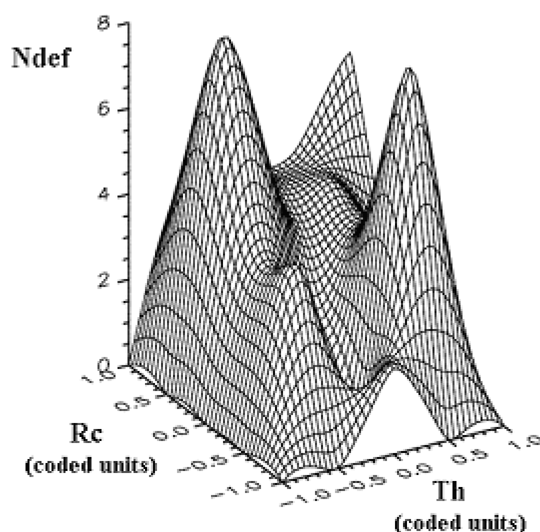

Figure 9. Ndef as a function of {Th, Rc} at Rh = Lh = Ll = 0.0 (coded units).

a cross-section of the previously performed 68 run set by the Rh = Lh = Ll = 0.0 (coded units) hypersurface. It can be seen that there is a fine mesh, but it lacks simulations at four corners. These four simulations were performed to augment the previous design as shown in figure 8(*b*).

The resulting set of Ndef values as a function of {Th, Rc} was interpolated by a minimum curvature algorithm, and the surface obtained is presented in figure 9.

As seen in figure 9, the surface of Ndef is quite irregular and consists of several hills and hollows. The final system is defect-free at low Th values for almost any Rc value within the low and high limits. Increasing Rc increases the number of defects in the system for almost all Th values except very low ones. Increasing Th has the same effect and results in formation of additional internal defects. Two pronounced peaks are found at {Th; Rc} = {−0.5; 0.5} and {Th; Rc} = {0.5; −0.5}. The structure has a defect-free point at {Th; Rc} = {0.0; −0.5}, which violates the general trend. Another unexpected phenomenon is defect formation at {Th; Rc} = {0.0; −1.0}, although it can be explained by random noise. In general, a simultaneous increase of Th and Rc increases the number of defects in the system. In other words, either annealing at lower temperatures or slow cooling results in the fewest number of defects.

To investigate the influence of the initial sample geometry on the number of defects after the preparation cycle (figure 1), the final defect-free a-SiO_2 sample prepared at {Rh = 1.0 K fs^{-1}; Th = 10 000 K; Lh = 80 ps; Rc = 0.1 K fs^{-1}; Ll = 20 ps} was used as an initial system for a set of new 25 simulations performed according to figure 1 the preparation sequence of figure 1 at {Rh = 0.0; Lh = 0.0; Ll = 0.0} (coded units), and with Th and Rc chosen according to figure 8(*b*). This is similar to the simulations presented in figure 9, with the only difference being the initial sample geometry. The resulting number of defects on the {Rc; Th} mesh was interpolated by minimum-curvature algorithms and presented in figure 10.

A comparative analysis of figures 9 and 10 demonstrates a general similarity between the surfaces. For both systems, Ndef is minimized at lower temperatures or lower cooling rates. The small differences between the surfaces may be either a minor affect of the initial sample geometry or random noise. Figure 10 demonstrates a decreased number of internal defects probably because the initial structure is the defect-free amorphous silica sample.

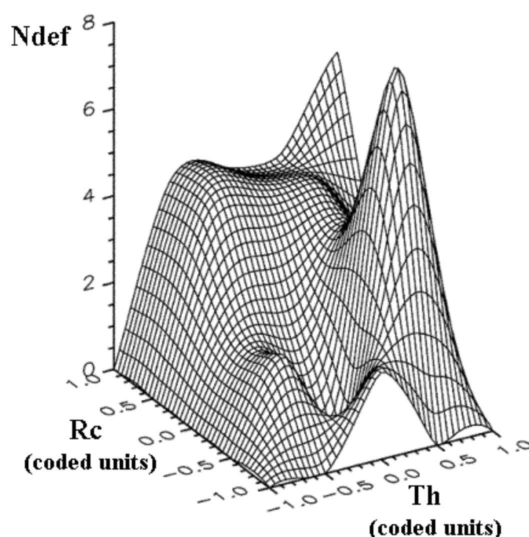

Figure 10. Number of defects, Ndef, as a function of {Th, Rc} at Rh=Lh=Ll=0.0 in the case of the amorphous sample as the initial system.

The common features are the broad defect-rich area near {Th = 0.0, Rc = 0.0} (coded units) and unambiguous growing of Ndef at {Th = 1.0; Rc = 1.0} (coded units). Increasing Lh can completely eliminate the possible minor effect of the initial geometry on the number of defects produced, but it will increase the computational cost.

Besides the impact of the different initial geometry on the number of produced defects (Ndef), we analysed concurrent variation of other response values. It was found that the change of the initial system had a negligible effect on the other responses. It can be explained by the fact that these properties are much more affected by the utilized inter-atomic potential and obtained by averaging over the much larger statistical ensemble.

The possible minor effect of the initial geometry does not undermine the validity and efficiency of the DOE approach as well as prepared sample quality. DOE techniques applied to an initial multidimensional mesh of response values are capable of interpolating response values and reliable prediction of the points of the optimal response magnitudes.

The real amorphization of the sample is not significantly affected by a possible weak initial geometry dependence. In general, the glass transition T° can be determined as the point of bending of the enthalpy curve vs temperature, which is a function of the cooling rate. To estimate this temperature for our cooling rates, we used the analytical form fit by Vollmayr *et al* [1], which gives 3178 (K) and 4056 (K) as glass transition temperatures for the lowest ($0.1\,\mathrm{K\,fs^{-1}}$) and highest cooling rates ($1.0\,\mathrm{K\,fs^{-1}}$), respectively. These temperatures should be accepted as estimates only since we use a different inter-atomic potential; however, they demonstrate that our Th values are able to produce an amorphous sample at the given cooling rates (table 1), even if the sample demonstrates a weak effect of the initial geometry.

It could be informative to filter out random oscillations of Ndef(Th, Rc) in the case of the crystalline sample as the initial system and to obtain an analytical form of its general behaviour as a function of all five parameters. The fact that the amplitude of the Ndef oscillations is comparable with the typical function magnitude complicates this task. One of the ways to do it is to fit the interpolated data by a polynomial of lower degree, which cannot follow all the oscillations and performs a sort of data averaging reproducing the general trend adequately.

To do this, overall 72 simulated points in the case of the crystalline sample as the initial system have been fit by a five-dimensional polynomial of the second order, which is presented below:

$$\mathrm{Ndef} = 3.89 + 1.84 * \mathrm{Th} + 1.48 * \mathrm{Rc} + 4.12 * \mathrm{Rh} * \mathrm{Rh} - 2.28 * \mathrm{Th} * \mathrm{Th} - 1.74 * \mathrm{Ll} * \mathrm{Ll} + 1.32 * \mathrm{Th} * \mathrm{Rc} + \varepsilon \quad (6)$$

Several terms that were found to be much less important than the other ones were omitted. As can be seen, the resulting analytical form (6) has a clear physical sense. For example, increasing Ll decreases Ndef, meaning that long equilibration can repair the bond network, though this effect is relatively weak. Growth of Rh increases the number of defects, which is a consequence of the weak memory of the initial heating stage. The terms with Rc emphasize the obvious expectation that steep cooling should increase the number of produced defects. The Ndef dependence on Th has a negative parabolic behaviour, which increases up to some temperature and then begins to decrease. However, the temperature of maximum is close to the Th high limit, which denotes a gradual average growth of defects for almost all investigated Th values. Since the simplified analytical form is only a polynomial of the second degree, it does not provide a very close approach to the simulated data on a five-dimensional hypersurface and should not be used for prediction of the number of defects. Its primary usefulness is in clarifying the general trends in the process of defect formation.

8. Conclusions

We present a systematic DOE approach that is capable of optimizing a MD preparation sequence to create defect-free a-SiO_2 samples with good RDF and ADF main peak positions and reasonable FWHM. The simulated silica sample has somewhat lower $\mathrm{FWHM}_{\mathrm{Si–O–Si}}$ and $\mathrm{FWHM}_{\mathrm{O–Si–O}}$ values than experimental ones. However, it should be taken into account that the experimental values of FWHM include broadening due to measurements of the neutron or x-ray structure factor and its subsequent Fourier transformation. The simulated $\mathrm{FWHM}_{\mathrm{Si–O}}$, $\mathrm{FWHM}_{\mathrm{O–O}}$ and $\mathrm{FWHM}_{\mathrm{Si–Si}}$ response values are greater than the experimental ones, especially $\mathrm{FWHM}_{\mathrm{O–O}}$. The discrepancy between experimental and simulated FWHM values probably is caused by the two-body part of the inter-atomic potential being too soft near the point of equilibrium.

The DOE technique provides a high degree of control over the number of defects produced as well as other properties and can result in defect-free samples. The total computational efficiency of the DOE approach is comparable with traditional preparation methods.

The response values became the target of statistical analysis, which revealed that each parameter or its product with another parameter affected at least one response value and therefore had to be preserved.

The DOE statistical analysis confirmed the conjecture that some properties of a-SiO_2 such as the RDF and ADF main peak positions are primarily determined by the inter-atomic potential, while others, such as Ndef, are controlled mainly by the preparation sequence. The FWHM of RDF and ADF peaks are dependent both on preparation sequence and inter-atomic potential. Simplified analytical expressions established quantitative dependences between the preparation parameters and some properties of the a-SiO_2 sample.

It was found that $\mathrm{FWHM}_{\mathrm{Si–O–Si}}$ and $\mathrm{FWHM}_{\mathrm{O–Si–O}}$ depend only on Th and Rc and do not depend on the other parameters. The effect of Th is much stronger than the effect of Rc and demonstrates a threshold behaviour. As Th increases, the $\mathrm{FWHM}_{\mathrm{Si–O–Si}}$ and $\mathrm{FWHM}_{\mathrm{O–Si–O}}$ sharply increase from 7° to 17° and 22°, respectively, reaching a maximum at $\mathrm{Th}_{\mathrm{max}} = 6750$ (K) and after that begin to decrease slowly and start to stabilize around 18–20° with minor oscillations.

The number of defects, Ndef, depends on Rh, Th, Rc and Ll. Increasing Ll weakly decreases the number of defects by repairing the bond network. Increase the cooling rate, Rc, and a high temperature, Th, produce more internal defects, though there can be some significant oscillations. Increasing Rh increases the number of defects in the system too, which is a consequence of the minor residual memory of the initial preparation stage.

$FWHM_{Si–Si}$ and $FWHM_{O–O}$ response values depend on Lh, Th, Rh and partially on Rc. Increasing Lh increases $FWHM_{Si–Si}$ and $FWHM_{O–O}$ since the longer the system is kept at high temperature, Th, the more the bond lengths smear. Increasing Rc results in increased $FWHM_{Si–Si}$ and $FWHM_{O–O}$ because the larger portion of high-temperature bond length smearing is transferred from the melt into the solid phase. Increasing Rh weakly decreases the smearing of O–O and Si–Si bond lengths. Increasing Th steadily increases $FWHM_{O–O}$, whereas $FWHM_{Si–Si}$ grows until $Th_{max} = 8600$ (K) and after that begins to decline slowly.

The Si–O–Si ADF main peak position depends on all five parameters. The fit functional form (5) provides an adequate description of Si–O–Si angle behaviour. It reflects the fact that a high temperature, Th, increases the Si–O–Si angle, while equilibration at a low room temperature shrinks it back. A high cooling rate, Rc, transfers the expanded Si–O–Si angle from the high-temperature melt into the solid SiO_2 bond network. The Si–O–Si angle is decreasing up to $Th = 8675$ (K), but after that it begins to grow slowly. This minimum temperature (8675 K) is quite close to the transition temperature of $FWHM_{Si–Si}$ (8600 K), which can be explained by the geometrical dependence between the Si–O–Si angle and Si–Si bond length.

It is important to emphasize that although our investigation was restricted to the generation of an a-SiO_2 model, the proposed approach is very general and may be effectively used for development and optimization of arbitrary amorphous systems. DOE techniques, which are actively and successfully used in many fields of engineering, can be extremely useful in solving non-trivial problems of computational physics and materials science.

Acknowledgments

This work was funded by Grant NSF DMI 9974381. In addition, we would like to thank National Centre for Supercomputer Applications (NCSA) for provided supercomputer time.

References

[1] Vollmayr K, Kob W and Binder K 1996 *Phys. Rev.* B **54** 15808
[2] Oligschleger C 1999 *Phys. Rev.* B **60** 3182
[3] Binder K 2000 *J. Non-Cryst. Solids* **274** 332
[4] Trachenko K O, Dove M T, Harris M J and Heine V 2000 *J. Phys.: Condens. Matter* **12** 8041
[5] Yamahara K, Okazaki K and Kawamura K 2001 *J. Non-Cryst. Solids* **291** 32
[6] Litton D A and Garofalini S H 2001 *J. Appl. Phys.* **89** 6013
[7] Kieffer J 1999 *J. Phys. Chem.* B **103** 4153
[8] Jin W, Kalia R K and Vashishta P 1994 *Phys. Rev.* B **50** 118
[9] Palin E J, Trachenko K O and Dove M T 2002 *J. Phys.: Condens. Matter* **14** 4857
[10] Bembenek S D and Laird B B 2001 *J. Chem. Phys.* **114** 2340
[11] Timpel D, Schaible M and Scheerschmidt K 1999 *J. Appl. Phys.* **85** 2627
[12] Valle R G D and Venuti E 1996 *Phys. Rev.* B **54** 3809
[13] Benoit M, Ispas S, Jund P and Jullien R 2000 *Eur. Phys. J.* B **13** 631
[14] Tse J S and Klug D D 1992 *Phys. Rev.* B **46** 5933
[15] Sarnthein J, Pasquarello A and Car R 1995 *Phys. Rev. Lett.* **74** 4682
[16] Pasquarello A and Car R 1997 *Phys. Rev. Lett.* **79** 1766
[17] Grossman I N 1985 *Sov. J. Glass Phys. Chem.* **11** 384

[18] Shiff V K 1990 *J. Non-Cryst. Solids* **123** 36
[19] Bell R J and Dean P 1972 *Phil. Mag.* **25** 1381
[20] Gladden L F and Elliot S R 1987 *Phys. Rev. Lett.* **59** 908
[21] Roder A, Kob W and Binder K 2001 *J. Chem. Phys.* **114** 7602
[22] Horbach J, Kob W and Binder K 1999 *J. Phys. Chem.* B **103** 4104
[23] Montgomery D C 2000 *Design and Analysis of Experiments* (New York: Wiley)
[24] Wyckoff R W G 1982 *Crystal Structures* (Malabar, FL: Krieger) p 36
[25] Wright A C 1994 *J. Non-Cryst. Solids* **179** 84
[26] Keen D A 1997 *Phase Transit.* **61** 109
[27] Andersen H C 1980 *J. Chem. Phys.* **72** 2384
[28] Duffrene L and Kieffer J 1998 *J. Phys. Chem. Solids* **59** 1025
[29] Johnson P A V, Wright A C and Sinclair R N 1983 *J. Non-Cryst. Solids* **58** 109
[30] Pettifer R F, Dupree R, Farnan I and Sternberg U 1988 *J. Non-Cryst. Solids* **106** 408
[31] Coombs P G, De Natale J F, Hood P J, McElfresh E K, Wortman R S and Schackelford J F 1985 *Phil. Mag.* **51** L39