

## Continuous Symmetry Measures of Density Maps<sup>†</sup>

Mark Pinsky,<sup>\*‡§</sup> David Danovich,<sup>||§</sup> and David Avnir,<sup>\*||§</sup>

Institute of Earth Sciences, Institute of Chemistry, and The Lise Meitner Minerva Center for Computational Quantum Chemistry, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

Received: March 10, 2010; Revised Manuscript Received: May 6, 2010

Density maps are commonly used for data presentation, and their symmetry is a central characteristic. Quite often near symmetry can be detected but not the exact one. Here we present a method for estimating the symmetry content of density maps. We thus extend the application of the continuous symmetry measure (CSM) methodology to many-points problems. We demonstrate the applicability of the method on electron-density probability maps and use, as a model, a series of copper clusters: <sup>1</sup>Cu<sub>3</sub><sup>+</sup>, <sup>1</sup>Cu<sub>4</sub> (planar and tetrahedral), <sup>2</sup>Cu<sub>7</sub>, and <sup>1</sup>Cu<sub>6</sub>. These were distorted in various ways, and the electron density maps of the distorted clusters were analyzed with respect to the content of the elementary symmetry point groups, namely reflection, inversion rotation, and improper rotation. In general, it has been found that symmetry distortions of nuclei arrangement are amplified in the electronic arrangement.

### 1. Introduction

Quantitative evaluation of the degree of symmetry<sup>1,2</sup> has evolved into a useful concept for establishing correlations between this structural property, and molecular properties that depend on it. An early demonstration of the applicability of this approach is due to Ratner et al.<sup>3</sup> A quantitative relation between the first hyperpolarizability,  $\beta$ , of noncentrosymmetric structures, and the quantitative degree of centrosymmetry content of such structures, was shown for the first time: For a series of systematic distortions of benzene, a good monotonic relationship between  $\beta$  and the inversion-symmetry measure, was found, thus adding a new tool for the rational design of nonlinear optical materials. Since the publication of this application, numerous correlations with quantitative symmetry, or the related quantitative chirality or polyhedral shape measures, have been established.<sup>4</sup>

The definition of the continuous-symmetry measures (CSM),  $S(G)$ , is based on estimating the minimal distance between two structures: the original one, represented by a set of points,  $\mathbf{Q}_k$ , and the searched, nearest  $G$ -symmetric structure, represented too by a set of points  $\hat{\mathbf{Q}}_k$ . The measure  $S(G)$ , is defined as<sup>1</sup>

$$S(G) = 100 \times \frac{1}{nd^2} \sum_{k=1}^n |\mathbf{Q}_k - \hat{\mathbf{Q}}_k|^2 \quad (1)$$

where  $d$  is the root-mean-square size of the original mass-centered structure ( $\mathbf{Q}_k$ ) and  $n$  is the number of points in the structures. The CSM has been estimated by either using numerical algorithms (such as the folding-unfolding algorithm<sup>5</sup>) or employing analytical methods.<sup>2,6,7</sup> The main limitation of these methods has been the need to search for optimal permutations between  $\mathbf{Q}_k$  and  $\hat{\mathbf{Q}}_k$  to minimize  $S(G)$ . Since the number of all possible permutations grows as  $n!$ , estimation of

the CSM of large molecules could be carried out only if information on connectivity (which remains unchanged in  $\hat{\mathbf{Q}}_k$ ) has been used as an input;<sup>8</sup> large assemblies of disconnected points, such as density maps, remained out of scope of the CSM methodology. Density maps can represent a wide range of physical phenomena, of statistical data, and of processed images (all in the context of symmetry analysis), and hence the interest in solving this computational problem. Here we propose and demonstrate approximate methods for that computational task and have selected electron density maps for that purpose. Such electron presentations are visual interpretations of the electronic wave functions, and their analysis should not be mixed with symmetry analyses of the electronic wave functions themselves.<sup>2</sup>

In particular, for a series of copper clusters in their electronic ground states, we estimate for their electron density functions the CSM of inversion-symmetry point group ( $C_i$ ), of rotational symmetry ( $C_n$ ), and of the improper symmetry ( $S_n$ ). The electron density functions are interpreted as electron-probability distribution functions ((pdf's),  $W(x,y,z)$ ) of random coordinates. It is of relevance to mention here that extensions of the CSM to electron densities and to molecular orbitals have been proposed before,<sup>9</sup> yet despite their importance as pioneering efforts, they are not as general as one may wish, and they do not adhere to the strict definition of the CSM as a distance function to an a priori unknown reference structure (which is searched). In a recent development<sup>2</sup> we have generalized the CSM for calculating *analytically* the degree of symmetry content for any mathematical entity that is part of a metric space such as vectors, matrices, operators, and functions, including electron density functions. Using this approach, the symmetry analysis of electron densities was further developed and demonstrated by Alemany et al.<sup>6</sup> The density-map analysis and the functions analytical approach are complementary: The former analyzes the graphical (vectorial) presentation of the density; the latter analyzes the function itself.

The focus of our report is on the computational tool, and we have selected to demonstrate it on a problem of practical relevance, namely, the structure and symmetry of metal clusters.<sup>10</sup> In the framework of this Festschrift, we recall Ratner's studies on alkali-metal clusters,<sup>11</sup> where, interestingly, he uses

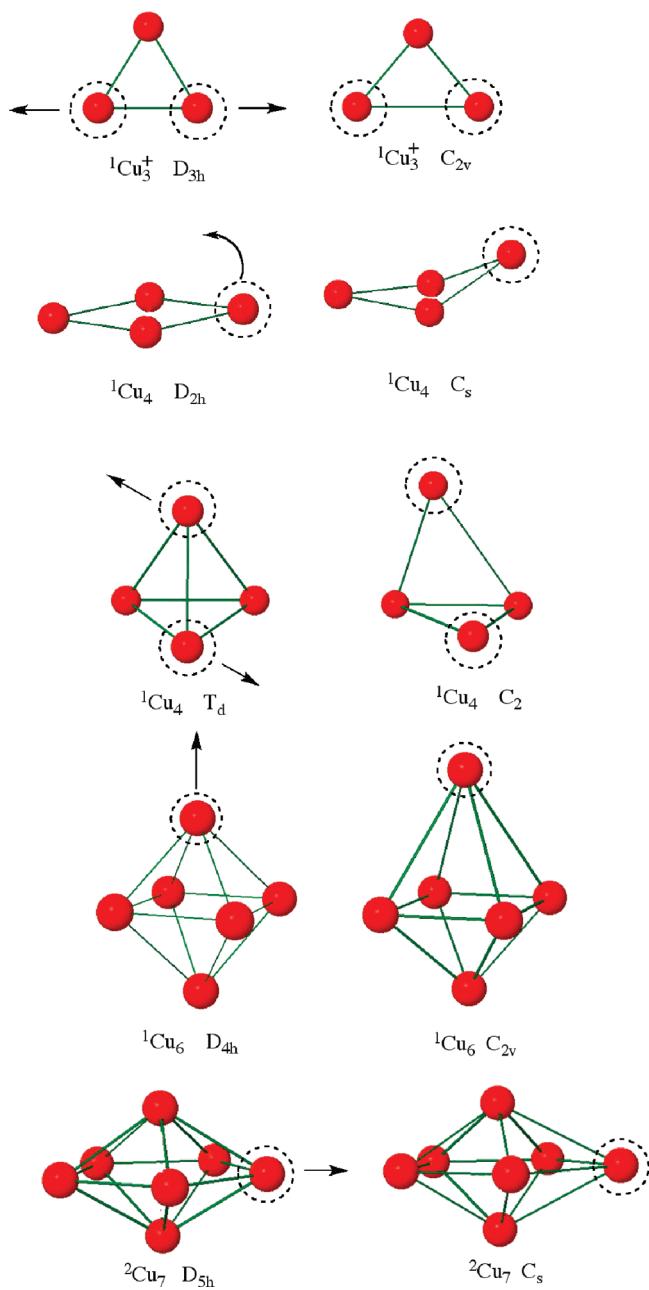
<sup>†</sup> Part of the "Mark A. Ratner Festschrift".

\* Corresponding authors. E-mail: M.P., pinsky@cc.huji.ac.il; D.A., david@chem.ch.huji.ac.il.

<sup>‡</sup> Institute of Earth Sciences.

<sup>§</sup> The Lise Meitner Minerva Center for Computational Quantum Chemistry.

<sup>||</sup> Institute of Chemistry.



**Figure 1.** Copper clusters used in this study. Left: original symmetric shapes and their symmetry point groups; arrows and dashed circles indicate the distortion mode (see text for details). Right: shape of the distorted cluster and the residual symmetry point group after distortion.

the terms “substantially non-centrosymmetric”, “reduced symmetry”, and “more highly symmetric”—all of which can now be replaced by the language and scale of “degree of symmetry”, as demonstrated in sections 3–5; first, however, we describe the copper clusters we use as a model.

## 2. Copper Clusters, Their Distortions, and Their Electron Density Probability Distribution Maps

**Copper Clusters.** The following copper clusters were studies to illustrate the general approach for estimating the symmetry of a large assembly of disconnected points (Figure 1):  ${}^1\text{Cu}_3^+$  ( $D_{3h}$  point group symmetry), two  ${}^1\text{Cu}_4$  clusters (in  $T_d$  and  $D_{2h}$  symmetries),  ${}^1\text{Cu}_6$  ( $D_{4h}$ ), and  ${}^2\text{Cu}_7$  ( $D_{5h}$ ; see Supporting Information for Cartesian coordinates of all clusters before and after

distortion). The geometries of the clusters were fully optimized using B3P86 (and UB3P86 for  ${}^2\text{Cu}_7$ , because this is an open shell system) density functional with Stuttgart relativistic effective core potential with extended valence basis set (with 8s7p6d/6s5p3d contractions)<sup>12</sup> as implemented in the Gaussian03 package;<sup>13</sup> that is, only valence electrons were used for this part of the calculation. To ensure that the obtained geometries are local minima on the potential energy surface, all calculations are the result of a full geometry optimization without symmetry restrictions, followed by the frequency calculations. In all copper clusters, the exact symmetries indicated above, were obtained.

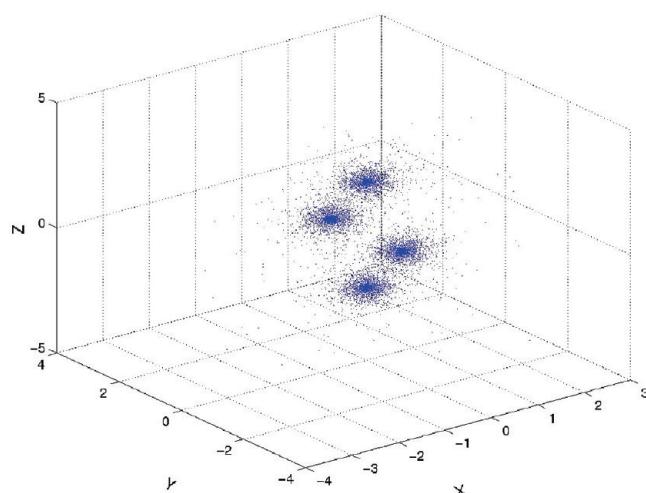
**Distorted Clusters.** These clusters were distorted as indicated by the arrows in Figure 1, as follows:

- ${}^1\text{Cu}_3^+$  was distorted from  $D_{3h}$  symmetry to  $C_{2v}$  symmetry by changing one of the angles from  $30^\circ$  to  $80^\circ$  while keeping the bond lengths of the two bonds that form this angle, unchanged.
- ${}^1\text{Cu}_4$  was distorted from  $D_{2h}$  symmetry to  $C_s$  symmetry by bending one atom in the  $XY$  direction (the plane of the page), accompanied by a small bond elongation (which is why the structure is of  $C_s$  symmetry and not of  $C_{2v}$  symmetry; see Supporting Information). Four structures were thus obtained, with a maximal distortion of  $34.84^\circ$ .
- Another  ${}^1\text{Cu}_4$  cluster was distorted from its  $T_d$ -symmetric version to  $C_2$  symmetry by moving one atom  $0.6 \text{ \AA}$  and the second one  $-0.6 \text{ \AA}$  along the  $z$ -axis, relative to the  $T_d$  structure, a bond elongation of 18%.
- ${}^2\text{Cu}_7$  was distorted from  $D_{5h}$  symmetry to  $C_s$  symmetry by moving one atom  $0.4 \text{ \AA}$  along the  $y$ -axis (a bond elongation of 16%), so that the two bonds connected to this atom are elongated unequally (which is why the structure is of  $C_s$  symmetry and not of  $C_{2v}$  symmetry; see Supporting Information).
- ${}^1\text{Cu}_6$  was distorted from  $D_{4h}$  symmetry ( $O_h$  is not of minimal energy in this case) to  $C_{2v}$  symmetry by moving one Cu atom  $1.0 \text{ \AA}$  along the  $y$ -axis, relative to the  $D_{4h}$  structure, a 37% bond elongation.

We emphasize that the distorted clusters were not reoptimized and therefore should not be treated as locally stable structures, but as models for analyzing the mathematical tools developed here for analysis of the symmetry density maps.

**Electron-Density Probability Maps.** The electron density maps of the optimized and distorted clusters were calculated with wave functions obtained in the single point calculations with the restricted Hartree–Fock (RHF) method for the clusters with singlet electronic ground state and the unrestricted Hartree–Fock (UHF) method for the clusters with doublet electronic ground state. In these calculations all-electron split valence (SV) basis set as implemented in the ORCA program<sup>14</sup> was used to obtain self-consistent wave functions. For one of the clusters ( $D_{2h}$   ${}^1\text{Cu}_4$ ) we used the much larger 6-31G\* basis set (136 contracted basis functions vs 84) to ensure the validity of the choice of the smaller basis set (which indeed proved justified; see below). For the self-consistent field (SCF) procedure extreme convergence criteria were used. The input file for the ORCA program, including SCF (with extreme convergence criteria to obtain accurate wave function) and PLOTS (3D plot with 500 points resolution in each direction) sections can be also found in the Supporting Information.

For calculating the symmetry measure one converts the electron probability density functions (pdf's) into a set of random points (coordinates). This was done using a generator of random numbers from given 3D (or 2D) pdf as described in the Appendix. Large sets of disconnected points  $\{x_i^*, y_i^*, z_i^*; i = 1, \dots, n\}$  (typically,  $n \approx 10^5$  to  $10^6$ ) (the density maps) are thus



**Figure 2.** Series of 10 000 three-coordinate sets, randomly selected from the wave function of the  $^1\text{Cu}_4$  ( $D_{2h}$ ) cluster.

obtained. Note that the coordinates of the points are noncorrelated. Figure 2 demonstrates the electron density map of  $^1\text{Cu}_4$  in its  $D_{2h}$ -symmetric form.

### 3. Solution to the Problem of Symmetry Evaluation of Density Maps

We present our solution to the problem of evaluating the symmetry of a large collection of points using a specific case: We calculate the degree of inversion symmetry,  $S(C_i)$ , of the electron-density maps of distorted  $^1\text{Cu}_4$  clusters (Figure 1, second row). Inversion symmetry was selected for this section, because the location of that symmetry element is easy to calculate: It is at the geometric center of mass of the map. In general, for CSM evaluation one needs to know the location of the symmetry element, the measure of which one estimates, and in the examples that follow in the next sections, we treat cases where the determination of the location and orientation of the symmetry element is less obvious.

As explained above, the requirement for CSM estimation to examine all of the possible permutations between points is practically impossible to carry out. This was solved as follows: Suppose one wishes to estimate the  $S(C_i)$  of the  $10^6$  points of the cluster of Figure 2 or of a distorted version of it. A first point,  $\mathbf{Q}_1$ , is selected randomly and  $C_i$  is operated on it, resulting in an inverted point. The point in the density map which is *nearest* to this inverted coordinate,  $\tilde{\mathbf{Q}}_1$ , is identified and selected; the first pair of  $\mathbf{Q}_1$  and  $\tilde{\mathbf{Q}}_1$  is thus obtained. Note that if the pair is of exact  $C_i$  symmetry, then the inverted coordinates coincide with  $\tilde{\mathbf{Q}}_1$ . Repetition of this procedure for the remaining points allows one to divide the set of the  $n$  points into  $n/2$  orbits. The last  $C_i$  orbits can be far from that symmetry, but if the total number of the orbits is large enough, then their influence on the resulting estimation becomes negligible. Likewise, as we show below, this method is insensitive to the selection of the first point.

A practical question in the estimation of the symmetry measure of density maps is, how many points does one need to render the symmetry estimation accurately enough? That is, beyond what  $n$  is a plateau in the estimated symmetry measure reached? For that purpose,  $\hat{S}(G,n)$ , the dependence of the CSM estimation (eq 1) on the number of points, is plotted. Let us take as an example the first distorted  $^1\text{Cu}_4$  structure (0.25 Å movement along the  $x$ -axis and  $-0.157107$  Å movement along the  $y$ -axis, with a  $7.6863^\circ$  distortion for a bending angle) (Figure

3, square dots): It is seen that at first, as  $n$  increases,  $\hat{S}(C_i,n)$  drops (the estimation gets better as more points are used), until a slightly descending semiplateau is reached beyond  $\sim 10^4$  points. One can either average the  $\hat{S}(G,n)$  values of the plateau or select the highest  $n$  ( $10^6$  in Figure 3) to report the estimated symmetry measure,  $\hat{S}(G)$ ; in this report we take the latter option ( $\hat{S}(C_i)$  is the estimation of the true symmetry measure,  $S_0(C_i)$ ; see eqs 2–4 below.) To check the sensitivity of results on the sequence of orbits, we carried out an additional calculation of  $\hat{S}(C_i,n)$  with another random sequence of orbits. It was found that the relative error does not exceed 1% for  $n > 3 \times 10^4$ . It is informative to apply the proposed methodology on the case of exact nuclear  $D_{2h}$  symmetry for which  $S(C_i) = 0$  and analyze its electron density map. It is seen (Figure 3) that the  $\hat{S}(C_i,n)$  function shows a near power-law behavior, never reaching  $\hat{S}(C_i) = 0$ , but providing an  $\hat{S}(C_i) > 0$  estimation of the true  $S(C_i) = 0$  value; the estimation is  $\hat{S}(C_i) = 0.045$ . Note that in the distorted case there are two domains: For low  $n$  values there is an approximate power-law line, which then straightens to a plateau for higher  $n$  values. We now explain these observations as arising from errors in the CSM estimation, which are due to random fluctuations of generated points' coordinates.

In an earlier study<sup>15</sup> we have shown that random fluctuations of point coordinates lead to both a positive bias and a dispersion of CSM estimation. The following equations for the mean value and the variation of the CSM estimation were obtained under an assumption that the points' coordinates are normally distributed.

$$\langle \hat{S} \rangle = 100 \times \left( S_0(G) + \frac{3\sigma^2}{d^2} \right) \quad (2)$$

$$\sigma_s^2 = \frac{2 \times 10^4 \sigma^2}{nd^2} \left( 2S_0(G) + \frac{3\sigma^2}{d^2} \right) \quad (3)$$

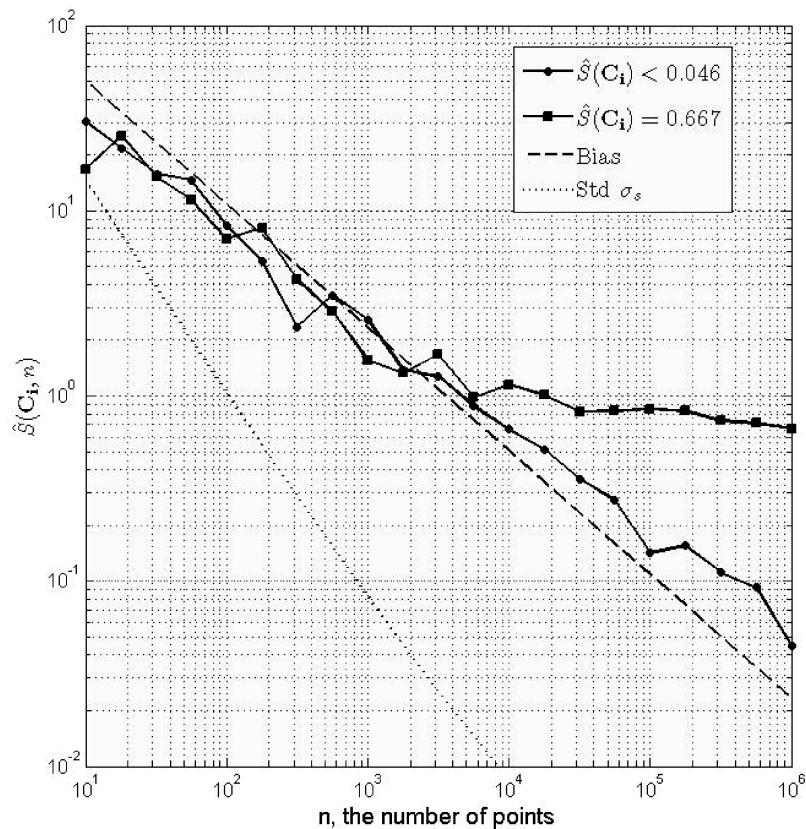
Here  $S_0$  is the true value of CSM,  $d$  is the rms radius of the structure,  $3\sigma^2$  is the coordinate variation that originates from the transition from a function to a density map (section 2), and  $3\sigma^2/d^2$  is the bias of estimation. In our density map application the variation  $3\sigma^2$ , can be evaluated in the following way: The volume of the structure can be approximately evaluated as  $(4/3)\pi d^3$ , where the mean volume occupied by one point is  $(4/3n)\pi d^3$ . Its linear scale,  $l = (4/3n)\pi^{1/3}d$ , can be assumed to determine the average coordinates variation, namely,  $3\sigma^2 = l^2$ . Using these evaluations and eqs 2 and 3, one obtains

$$\langle \hat{S} \rangle = 100 \times \left[ S_0 + \left( \frac{4\pi}{3} \right)^{2/3} n^{-2/3} \right] \quad (4)$$

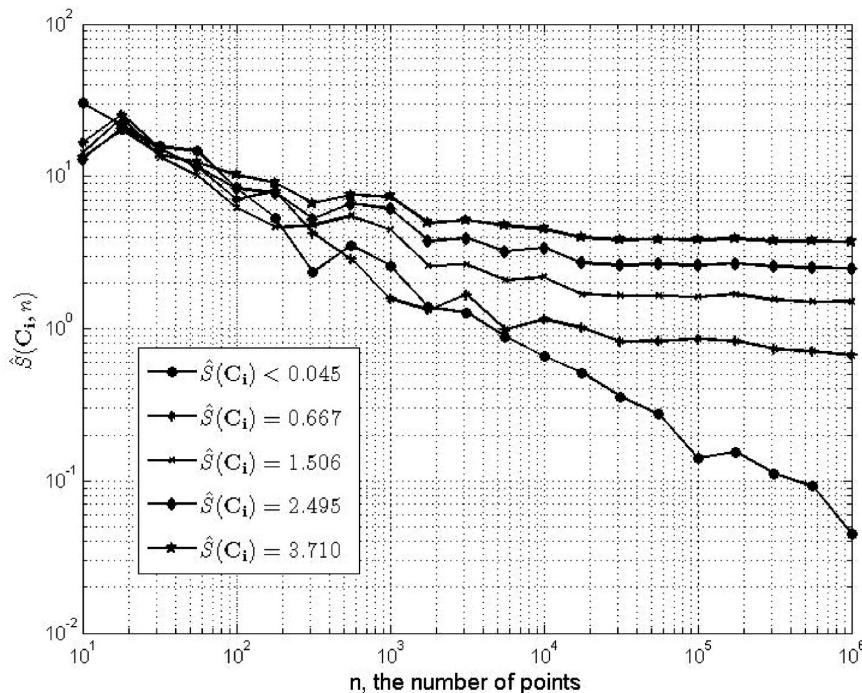
$$\sigma_s^2 = 10^4 \times \left\{ \frac{2}{3} \left( \frac{4\pi}{3} \right)^{2/3} \left[ 2S_0 + \left( \frac{4\pi}{3} \right)^{2/3} n^{-2/3} \right] n^{-5/3} \right\} \quad (5)$$

Equations 4 and 5 show that the bias and the variation of the CSM estimation do not depend on the size of the structure. Furthermore, these equations show also that the bias and the variation decrease with  $n$  as  $n^{-2/3}$  and  $n^{-5/3}$ , respectively (Figure 3).

Returning now to the distorted and nondistorted structures of  $^1\text{Cu}_4$  (Figure 3), one can see that at relatively small values of  $n < 5 \times 10^3$ ,  $\hat{S}(C_i,n)$  has a slope of  $-2/3$  for both cases, namely it actually represents the bias of estimation. At larger values of  $n$ ,  $\hat{S}(C_i,n)$  of the distorted structure turns into a plateau, where the estimation exceeds the bias by a factor of  $\sim 30$  (with



**Figure 3.** Inversion-symmetry estimation,  $\hat{S}(C_i, n)$ , of the electron density maps of distorted (squares) and nondistorted (circles)  $D_{2h}^{-1}\text{Cu}_4$  clusters as a function of the number of points,  $n$ . The dependences of the bias and of the standard deviation on  $n$  (eqs 4 and 5) are shown as well. In the inset, the  $\hat{S}(C_i)$  values of the two clusters are given.



**Figure 4.** Inversion-symmetry estimation,  $\hat{S}(C_i, n)$ , of the electron density maps of distorted and nondistorted  $D_{2h}^{-1}\text{Cu}_4$  clusters as a function of the number of points,  $n$ , for various levels of distortions. From bottom to top the distortions are 0.25 Å movement along the  $x$ -axis and  $-0.157107$  Å movement along the  $y$ -axis, with a  $7.68^\circ$  distortion for a bending angle; 0.5 Å movement along the  $x$ -axis and  $-0.314207$  Å movement along the  $y$ -axis, with a  $16.26^\circ$  distortion for a bending angle; 0.75 Å movement along the  $x$ -axis and  $-0.471407$  Å movement along the  $y$ -axis, with a  $25.46^\circ$  distortion for a bending angle; 1.0 Å movement along the  $x$ -axis and  $-0.628507$  Å movement along the  $y$ -axis, with a  $34.84^\circ$  distortion for a bending angle.

an error of  $\sim +3\%$ ). The random part of the errors characterized by standard deviation  $\sigma_s$  is shown in the figure as well. Note

that the variation, which decreases as  $n^{-5/6}$  with  $n$ , produces negligible errors.

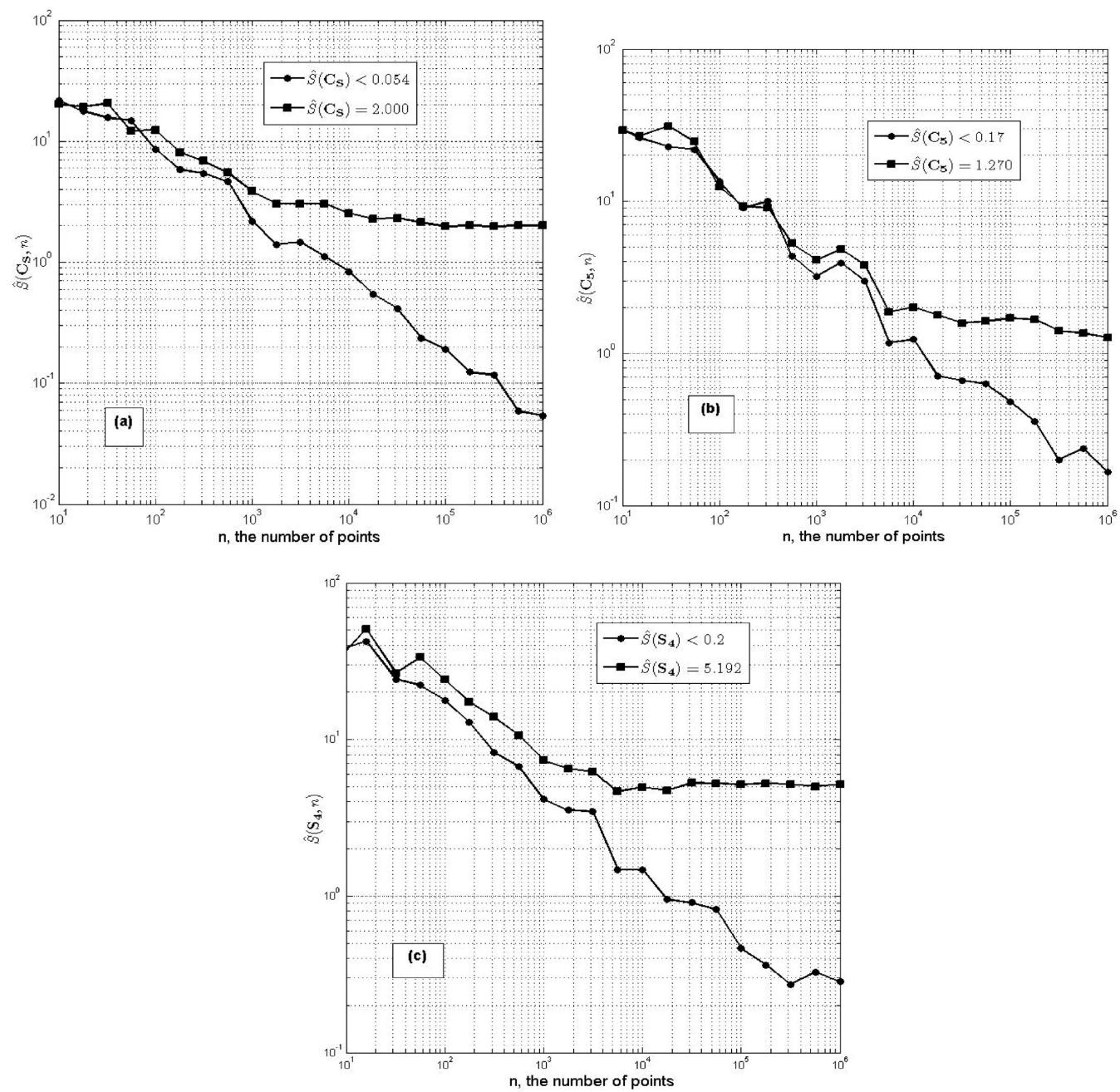
**TABLE 1: Comparison of the Inversion-Symmetry Measure Values for the Electron Density Maps and for the Cu Nuclei Location**

CSM of density map <sup>a</sup>	CSM of nuclei
0.046 (0.047)	0.000
0.669 (0.651)	0.195
1.507 (1.469)	0.790
2.495 (2.438)	1.783
3.710 (3.652)	3.140

<sup>a</sup> In parentheses, the results of using the larger 6-31G\* basis set are given.

Next, we follow in Figure 4 how the behavior of  $\hat{S}(C_i, n)$  varies with the degree of distortion to develop an intuition about the geometrical meaning of the CSM values. It is seen that the

higher the distortion, the earlier  $\hat{S}(C_i, n)$  departs from the  $n^{-2/3}$  behavior, as expected in view of the interpretation proposed above. We also use this series of  $\hat{S}(C_i)$  values to make an important comparison: Practically all of the earlier CSM analyses were based on the positions of the nuclei—how do these values compare with the electron density CSM values? Table 1 shows that consistently the nuclei distortion leads to a higher distortion in the electron density maps (which may seem counterintuitive). Some of this trend can be attributed to the inherent statistical errors in the density maps, but not all of it: Note that for exact nuclei  $C_i$  symmetry the error in estimation is of the order of 0.05, which can be taken as the error range of the entries in Table 1. Also given in Table 1 are the results of using the 6-31G\* basis set: As seen there, for our purposes, the smaller basis set is good enough.



**Figure 5.** Electron density maps symmetry analyses of symmetry-distorted copper clusters (upper lines) and of their nondistorted origins (lower lines). (a) Reflection-symmetry estimation,  $S(C_4)$ , of  $^{1\text{Cu}_4}$  distorted from its  $T_d$  symmetry to  $C_2$ . (b) Rotational-symmetry estimation,  $S(C_5)$ , of  $^{2\text{Cu}_7}$  distorted from  $D_{5h}$  symmetry to  $C_s$  symmetry. (c) Roto-reflection-symmetry estimation,  $S(S_4)$ , of  $^{1\text{Cu}_6}$  distorted from  $D_{4h}$  symmetry to  $C_{2v}$ .

**TABLE 2: CSM Values for Different Symmetry Types, Calculated from the Atom Coordinates and 3D Density Maps of the Wave Function**

cluster, original symmetry	symmetry analyzed	CSM - electron density (error)	CSM-nuclei
$^1\text{Cu}_4, T_d$	$C_s$	2.000 (0.054)	1.018
$^2\text{Cu}_7$	$C_5$	1.270 (0.17)	0.302
$^1\text{Cu}_6$	$S_4$	5.192 (0.30)	2.198
$^1\text{Cu}_4, T_d$	$C_2$	0.0718 (0.07)	0.000
$^1\text{Cu}_4, T_d$	$C_i$	17.77 (0.07)	25.28
$^1\text{Cu}_4, T_d$	$S_4$	3.995 (0.07)	2.025

#### 4. Symmetry Evaluation of Density Maps of Copper Complexes: Rotation, Reflection, Roto-reflection, and Chirality

In the previous detailed example, inversion symmetry was used, and it was noted that its location (the center of mass) is available. For the estimation of other symmetry measures, a prior information about the optimal direction of the symmetry element (rotation axes and/or the direction of the normal to the reflection plane) is required. In its absence, the current wisdom is to use further approximations for its determination. The approximation we use here is to determine the optimal direction of the symmetry element from the CSM analysis of the atoms nuclei arrangement; computational tools for that purpose have been developed.<sup>6</sup> It leads to some error in the electronic CSM estimation, but in many cases we do not expect it to be large. We demonstrate the applicability of this approach with the estimation of the three most elementary symmetry groups (in addition to inversion), namely mirror symmetry  $S(C_s)$  (Figure 5a,  $^1\text{Cu}_4$  distorted from its  $T_d$  symmetry to  $C_2$ ; see Figure 1, second row), rotational symmetry  $S(C_5)$  (Figure 5b,  $^2\text{Cu}_7$  distorted from  $D_{5h}$  symmetry to  $C_s$  symmetry; see Figure 1, last row), and roto-reflection  $S(S_4)$  (Figure 5c,  $^1\text{Cu}_6$  distorted from  $D_{4h}$  symmetry to  $C_{2v}$ ; see Figure 1, fourth row). In all three cases the symmetries of the nondistorted clusters were calculated as well; all  $S(G)$  values are collected in Table 2 (first three rows) and it is seen that as in Table 1, here too, the electron density

CSM values are always higher than the values obtained from the Cu-atom locations.

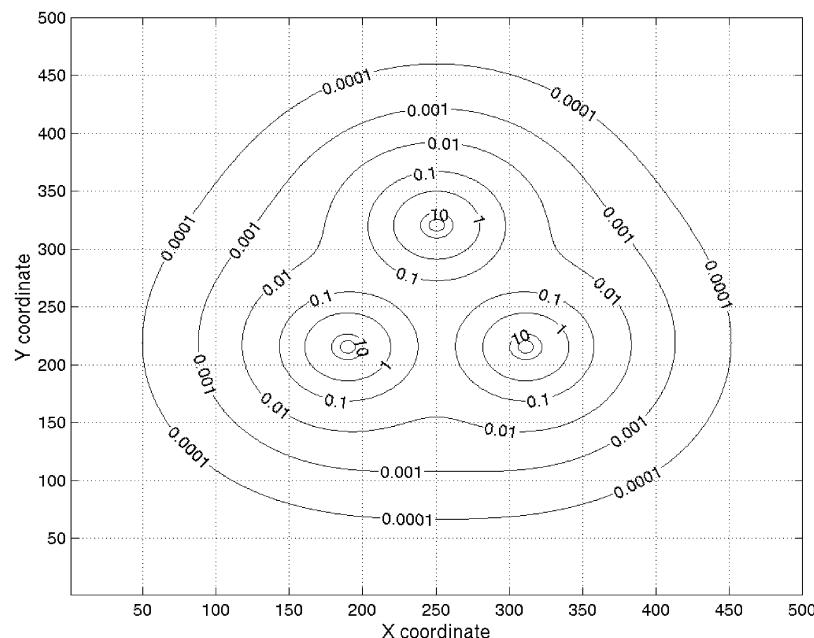
Table 2 contains additional analyses performed on  $T_d$   $^1\text{Cu}_4$  (Figure 5a): First, since the distortion is to another (lower) symmetry,  $C_2$  in this case for the nuclei, it was of interest to estimate the  $S(C_2)$  value, as a test of the methodology. The result (fourth row), 0.0718, compared to 0.0000 for the nuclei  $S(C_2)$  analysis, indicates the bias of estimation. Second, since the distorted cluster is of  $C_2$  symmetry, it is a *chiral cluster*. According to the continuous chirality measure (CCM) approach, CSM analysis provides an estimation of the degree of chirality: It is the minimal distance to achirality, namely, the minimum of all  $S(n)$  values ( $n = 0, 2, 4, \dots$ ). From Table 2 ( $C_s$ ,  $C_i$ , and  $S_4$ ; first, fifth, and sixth rows in the Table 2) it is clear that  $\hat{S}(\text{chiral}) = \hat{S}(C_s) = 2.000$ . Finally, note the very high value of  $\hat{S}(C_i)$ , both for the electron density and for the nuclei; this is because  $C_s$ ,  $C_2$ , and  $S_4$  are subgroups of  $T_d$  so the distance to these symmetries in the distorted cluster is small, but  $C_i$  is not a subgroup of  $T_d$ , so the distance to the nearest structure with that symmetry is large.

#### 5. Planar $^1\text{Cu}_3^+$ : 2D Analyses of 3D Density Maps

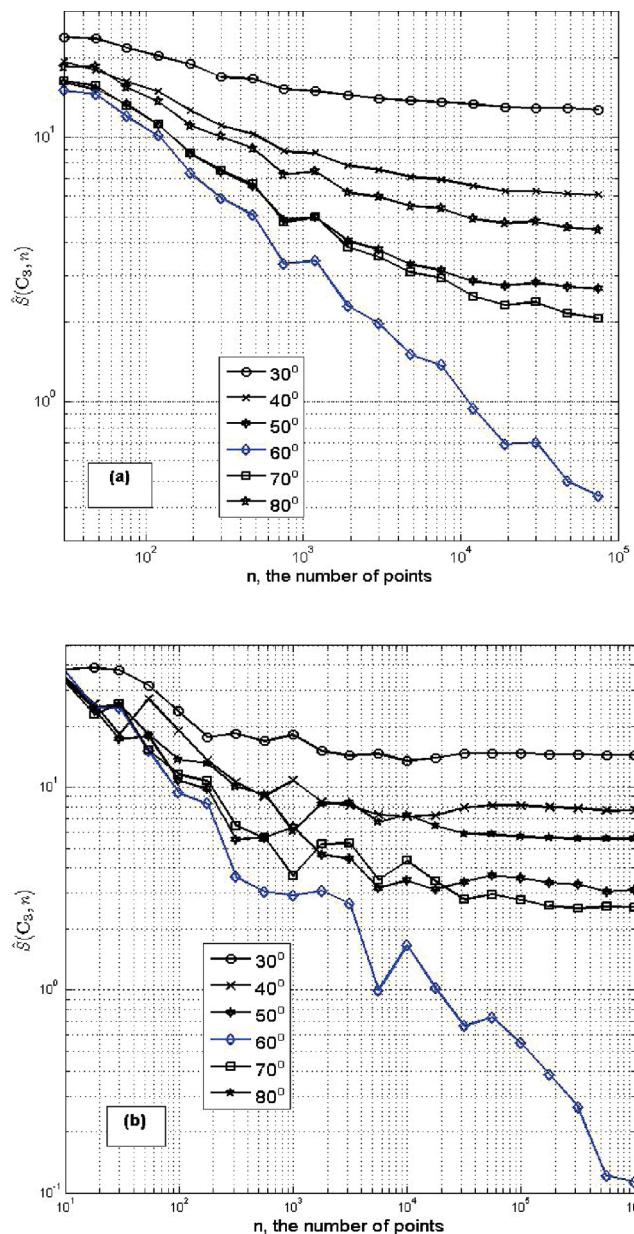
Analyses of 2D cross sections of 3D objects, when relevant, may provide savings in computation time and ease in locating the symmetry element (center point instead of an axis). We take the density map of the planar  $D_{3h}$   $^1\text{Cu}_3^+$  cluster (Figure 6) and its planar distortions ( $C_{2v}$  symmetry; Figure 1, top row) to evaluate this possibility, by comparing 2D (Figure 7, top) and 3D (Figure 7, bottom)  $\hat{S}(C_3)$  analyses. The results (Table 3) indicate that, indeed, the 2D analyses, whenever relevant, can be used particularly if the identification of trends is more important than the value itself (as in this case). It is also seen that except for  $60^\circ$ , the case of exact  $C_3$  symmetry, the 2D density values are closer to the nuclei values, compared to the 3D values. This may reflect the additional degree of freedom for deformation out of the plane that is possible in the 3D case.

#### 6. Conclusion

As most of the molecular properties depend on the electronic structure, the extension of the continuous-symmetry measure



**Figure 6.** 2D cross-section contour map of the iso-surface total electron density of the fully symmetric  $D_{3h}$   $^1\text{Cu}_3^+$  cluster.



**Figure 7.**  $\hat{S}(C_3, n)$  dependencies on the angular distortion of  $^{1}\text{Cu}_3^+$  for  $\hat{S}(C_3)$  analysis: (a) 2D; (b) 3D.

**TABLE 3: Values of  $\hat{S}(C_3)$  for the 2D and 3D Cases**

angle (deg)	$\hat{S}(C_3)$ -2D density <sup>a</sup>	$\hat{S}(C_3)$ -3D density <sup>a</sup>	$\hat{S}(C_3)$ nuclei
30	12.68	14.39	11.98
40	6.043	7.707	5.055
50	2.668	3.113	1.079
60	0.442	0.113	0.00
70	2.071	2.568	0.848
80	4.472	5.559	3.223

<sup>a</sup> 2D:  $n = 10^5$ , error 0.44. 3D:  $n = 10^6$ , error 0.11 (smaller error than 2D because of the larger  $n$ ).

(CSM) methodology from the analysis of atom-arrangement geometry to electron densities was highly needed. Toward this goal we presented here a method for estimating the symmetry content of density maps, particularly of molecular electron-density probability maps. We demonstrated the applicability of the method on a series of copper clusters in their ground electronic states, from  $\text{Cu}_3$  to  $\text{Cu}_7$  and note that the symmetry distortions of the nuclei arrangement results in higher symmetry

distortions of the electronic structure. Exploration of a variety of symmetry dictated molecular properties by this approach is in progress.

**Acknowledgment.** We are grateful to Chaim Dryzun for many useful comments. Supported by the German-Israel Binational Foundation (grant 983/2007). Supported by the Israel Science Foundation (grant 307/08).

### Appendix. Generator of the Random Coordinates

For simplicity we detail the random number generator for the 2D case and then extend it to the 3D case:

#### Step 1

Let us define  $W_2(x,y)$  as a given 2D probability distribution function (pdf) of random coordinates. This pdf can be represented in the following form

$$W_2(x,y) = W_1(x) W(y|x) = W_1(y) W(x|y) \quad (\text{A1})$$

where  $W_1(x)$  and  $W_1(y)$  are the 1D pdfs and  $W(y|x)$  and  $W(x|y)$  are the conditional pdfs. We calculate  $W_1(x)$  and  $W(y|x)$  from a given 2D pdf using the following formulas

$$W_1(x) = \int_{-\infty}^{\infty} W_2(x,y) dy \quad W(y|x) = \frac{W_2(x,y)}{W_1(x)} \quad (\text{A2})$$

#### Step 2

The 1D integral pdf,  $F_1(x)$ , and the conditional integral pdf,  $F(y|x)$ , are calculated by integration of the corresponding pdfs:

$$F_1(x) = \int_{-\infty}^x W_1(x) dx \quad F(y|x) = \int_{-\infty}^y W(y|x) dy \quad (\text{A3})$$

Examples of these integral pdfs, corresponding to the wave function shown in the Figure 6 are presented in Figure 8.

#### Step 3

Using the basic congruent generator,<sup>16</sup> the random number  $\xi$ , which is uniformly distributed inside the interval [0 1], is generated. Then, the first random coordinate  $x_1^*$  is calculated by solving the equation

$$F_1(x_1^*) = \xi \quad (\text{A4})$$

#### Step 4

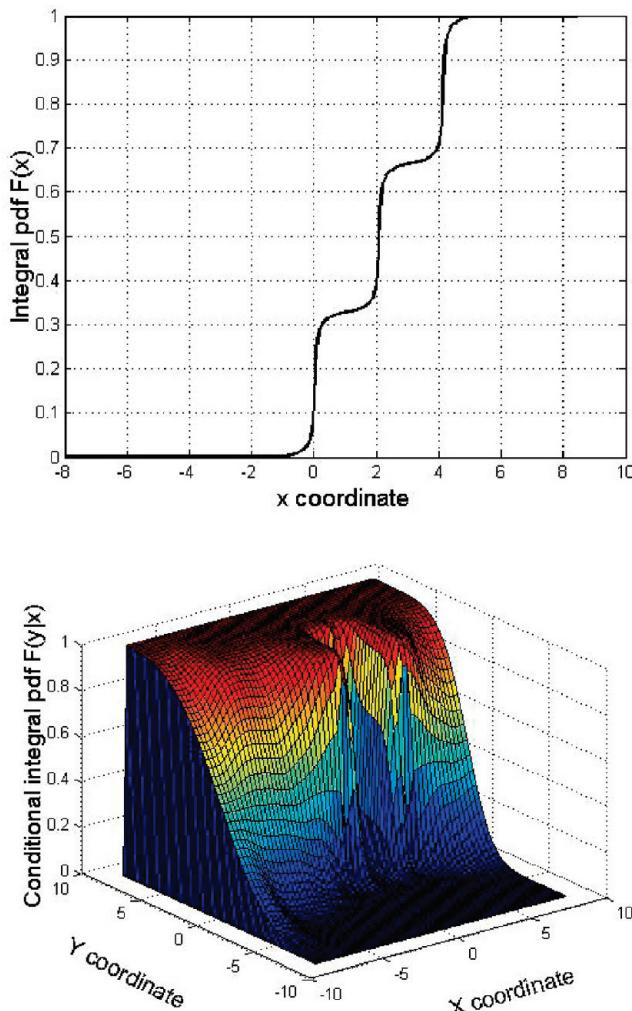
A new uniformly distributed random number  $\eta$  is formed, and then the second random coordinate  $y_1^*$  is calculated from the equation

$$F(y_1^*|x_1^*) = \eta \quad (\text{A5})$$

#### Step 5

Repetition of steps 3 and 4 the necessary number times allows one to obtain a set of random coordinates  $\{x_i^*, y_i^*, i = 1, \dots, n\}$ .

Similarly, for the generator in 3D case the 3D pdf of random coordinates  $W_3(x,y,z)$  is represented as



**Figure 8.** 1D integral pdf  $F_1(x)$  (top) and the conditional integral pdf  $F(y|x)$  (bottom) of the coordinates of  $^{1}\text{Cu}_3^+$  3D density map. The steps in the upper figure correspond to peaks of the density map (see Figure 6), which are located in the vicinity of nuclei centers.

$$W_3(x, y, z) = W_1(x) W(y|x) W(z|x, y) \quad (\text{A6})$$

All the three pdfs in the right side of (A6) can be calculated from a given 3D pdf  $W_3(x, y, z)$ :

$$\begin{aligned} W_1(x) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_3(x, y, z) dy dz \\ W_2(x, y) &= \int_{-\infty}^{\infty} W_3(x, y, z) dz \quad W(y|x) = \frac{W_2(x, y)}{W_1(x)} \\ W(z|x, y) &= \frac{W_3(x, y, z)}{W_2(x, y)} \quad (\text{A7}) \end{aligned}$$

Then, the corresponding integral pdfs,  $F_1(x)$ ,  $F(y|x)$ , and  $F(z|x, y)$ , are calculated by integration. Using the basic random number generator, three uniformly distributed random numbers  $\xi$ ,  $\eta$ , and  $\mu$  are produced. These numbers are used for calculation of random coordinates by solving the following equations

$$\begin{aligned} F_1(x_1^*) &= \xi \\ F(y_1^*|x_1^*) &= \eta \\ F(z_1^*|x_1^*, y_1^*) &= \mu \end{aligned} \quad (\text{A8})$$

To generate a set of random coordinates, the procedure is repeated the necessary number times.

Note that the integral pdf's (Figure 8, top) display sharp jumps. For more complicated structures consisting of many atoms, the density map fills the volume more homogeneously which would blur such jumps. In these situations the random number generator is more accurate and therefore application of the proposed method to more complicated density maps may result in lower error ranges.

**Supporting Information Available:** Coordinates of the original clusters and their distorted versions. Input file for the ORCA program for producing electron density maps. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## References and Notes

- (1) Zabrodsky, H.; Peleg, S.; Avnir, D. *J. Am. Chem. Soc.* **1992**, *114*, 7843. *Nature* editorial. Fowler, P. W. *Nature* **1992**, *360*, 626.
- (2) Dryzun, C.; Avnir, D. *Phys. Chem. Chem. Phys.* **2009**, *11*, 9653.
- (3) Kanis, D. R.; Wong, J. S.; Marks, C. J.; Ratner, M. A.; Zabrodsky, H.; Keinan, S.; Avnir, D. *J. Phys. Chem.* **1995**, *99*, 11061.
- (4) (a) Some recent papers using this methodology: Estrada, E.; Carbo-Dorca, R. *Match-Comm. Math. Computer Chem.* **2009**, *62*, 105. Cirera, J.; Ruiz, E.; Alvarez, S.; Neese, F.; Kortus, J. *Chem. Eur. J.* **2009**, *15*, 4078. Chang, H. Y.; Sivakumar, T.; Ok, K. M.; Shiv Halasyamani, P. *Inorg. Chem.* **2008**, *47*, 8511. Ergaz, I.; Toscano, R. A.; Delgado, G.; Steinberg, A.; Glaser, R. *Cryst. Growth Design.* **2008**, *8*, 1399. Wei, Y.; Sastry, G. N.; Zipse, H. *Org. Lett.* **2008**, *10*, 5413. Janssens, S.; Borgoo, A.; Van Alsenoy, C.; Geerlings, P. *J. Phys. Chem. A* **2008**, *112*, 10560. Farrugia, L. J.; Middlemiss, D. S.; Sillanpa, R.; Seppa, P. *J. Phys. Chem. A* **2008**, *112*, 9050. Milner, D.; Raz, S.; Hel-Or, H.; Keren, D.; Nevo, E. *Pattern Recognition* **2007**, *40*, 2237. (b) Some recent papers from our group: Dryzun, C.; Mastai, Y.; Shvalb, A.; Avnir, D. *J. Mater. Chem.* **2009**, *19*, 2062. See highlight in *Science* **2009**, *323* (5919), 1266. Zayit, A.; Pinsky, M.; Elgavi, H.; Dryzun, C.; Avnir, D. *Chirality*, in press. Tuvi-Arad, I.; Avnir, D. *J. Mat. Chem.*, in press.
- (5) Salomon, Y.; Avnir, D. *J. Math. Chem.* **1999**, *25*, 295.
- (6) Pinsky, M.; Casanova, D.; Alemany, P.; Alvarez, S.; Avnir, D.; Dryzun, C.; Kizner, Z.; Sterkin, A. *J. Comput. Chem.* **2008**, *29*, 190. Pinsky, M.; Dryzun, C.; Casanova, D.; Alemany, P.; Avnir, D. *J. Comput. Chem.* **2008**, *29*, 2712.
- (7) Casanova, D.; Alemany, P.; Alvarez, S. *J. Comput. Chem.*, in press.
- (8) Keinan, S.; Avnir, D. *J. Am. Chem. Soc.* **2000**, *122*, 4378.
- (9) For example: Grimme, S. *Chem. Phys. Lett.* **1998**, *297*, 15. Lipkowitz, K. B.; Gao, D.; Katzenelson, O. *J. Am. Chem. Soc.* **1999**, *121*, 5559. Bellarosa, L.; Zerbetto, F. *J. Am. Chem. Soc.* **2003**, *125*, 1975.
- (10) For example: Chang, C. M.; Chou, M. Y. *Phys. Rev. Lett.* **2004**, *93*, 133401.
- (11) Dahlseid, T. A.; Pople, J. A.; Ratner, M. A.; Kappes, M. M. *J. Phys. Chem.* **1994**, *98*, 8851.
- (12) Dolg, M.; Stoll, H.; Preuss, H.; Pitzer, R. M. *J. Phys. Chem.* **1993**, *97*, 5852.
- (13) Frisch, M. J., et al. *Gaussian 03*, Revision C.02; Gaussian Inc.: Wallingford, CT, 2004.
- (14) Neese, F. *ORCA, an ab initio, density functional and semi-empirical program package*; University of Bohn: Germany, 2008.
- (15) Pinsky, M.; Yoge-Einot, D.; Avnir, D. *J. Comput. Chem.* **2003**, *24*, 786.
- (16) Abramowitz, M.; Stegun, A. *Handbook of mathematical functions*; Dover Publications: New York, 1972; 1046 pp.