Appendices

to accompany

An evaluation of spatial correlation functions in textural analysis of metamorphic rocks

David M. Hirsch, Richard A. Ketcham and William D. Carlson

Appendix A. Correct production of null-hypothesis envelopes

Importance

One conclusion of this study that deserves special emphasis concerns the production of correct envelope simulations, which produce the interface-controlled null-hypothesis regions to which each rock's statistics may be compared. Errors in the production of envelopes may produce nonsensical results, which are easily detected, or may produce results that, while not obviously wrong, are nonetheless incorrect, and lead to incorrect conclusions.

The shape of the sample must be replicated in the envelope simulations. If this is omitted, then the edge-correction methods will have differing effects in the sample as compared to the envelope simulations.

The set of crystals, including the radius of each, must be replicated in the envelope simulations.

The envelope simulations must reflect any limitations on crystal observability that are present in the natural data. The observability criterion must be tuned for the data-collection measures used to obtain the spatial data from the rock sample. If this is not performed, then the statistical measures of the rock may show reduced or excess ordering relative to the envelope simulations, and diffusional control may be mistakenly concluded or mistakenly rejected (Fig. A1). Observability criteria are discussed in detail below.

The crystals in the envelope simulations must be placed with the restrictions enforced by the interface-controlled growth rate. If this is omitted, then, depending on the placing method used, the simulations may show either excess ordering or excess clustering relative to the rock, and conclusions drawn will then be incorrect.

Calculation

The production of null-hypothesis regions for each statistic is vital, and is quite exacting. Great care must be taken in their calculation, for testing has shown that even moderately sophisticated envelope-production methods can produce erroneous conclusions. The method used in the present work to produce correct envelope simulations is detailed below; it has been

extensively tested against a large database of controlled simulations of both interface-controlled and diffusion-controlled nucleation and growth produced by time-explicit methods.

Crystal placement. Each envelope simulation is an artificial array of crystals designed to model the statistical characteristics of a rock with a given crystal size distribution that formed in an interface-controlled episode of nucleation and growth. By this it is meant that as many parameters as possible for the rock under analysis are mirrored in the envelope simulation. This is accomplished by creating a set of crystals whose sizes match those of the crystals in the rock under analysis and placing those crystals in a bounding box of identical size and shape to the rock, from largest to smallest, in random locations (as expected in an interface-controlled nucleation environment). The volume effect – that is, the impossibility of a crystal nucleating inside a pre-existing crystal – is taken into account as follows. If any crystal, when randomly placed, is found to overlap a previous crystal, a test is performed to check whether the later-placed crystal could have nucleated in that location relative to the earlier one, or whether that nucleation site would have been within the pre-existing crystal. Because for interface-controlled growth, radii are proportional to the elapsed time since nucleation, this check is easily performed by removing from both crystals the radius of the smaller one, and testing whether the nucleation point lies within the body of the earlier-nucleated crystal as it existed at the nucleation time of the later crystal. This requirement is expressed mathematically as follows:

$$d \geq r_L - r_S \,, \qquad \text{(A1)}$$

where d is the center-to-center distance between the two crystals, and r_L and r_S are the radii of the larger and smaller crystals, respectively. This requirement stems from the linear relation between radius and time elapsed since nucleation under an interface-controlled growth law. If this inequality is true, then the location is allowed; if not, a new random location is selected.

Volume Fraction. The possibility of matching the volume fraction between the sample and the envelope simulations was explored. One would have to compromise between a precise match to crystal sizes and a precise match to volume fraction, unless the degree of overall impingement of crystals were also matched. However, as degree of impingement is strongly linked to the short-range ordering features of the sample, doing so would always produce an envelope with ordering features matching the data set. This would invariably lead to the erroneous conclusion of interface control in every case.

A number of test runs were performed using various simulations, both interface- and diffusion-controlled, and the difference was minimal between matching volume fraction moderately closely, and not matching volume fraction at all. Attempting to match volume fraction between the sample and the null-hypothesis simulations was judged to be less than rigorous, insofar as it produced envelopes whose crystal size distributions did not precisely match those of the sample data sets; the additional computation time required is a further disadvantage to this approach.

Observability. Another key issue to consider is observability. In an interface-controlled nucleation and growth regime, highly interpenetrating pairs (and clusters) of crystals may exist. These highly overlapped crystals are generated in time-explicit interface-controlled nucleation

and growth simulations, and are permitted by the interface-controlled placement criterion given in equation (A1). In natural samples, however, they might not be observed if the interpenetration produces shapes for the compound crystals that cannot be resolved as clusters and separated into individual crystals. This observability problem will arise during the analysis of sample data collected by many techniques, including the serial sectioning and optical scanning technique (Raeburn, 1996; Daniel and Spear, 1999) and the technique of computed X-ray tomography we have used.

Pairs of highly intergrown crystals in a rock may sometimes be misidentified as single, larger crystals. When this occurs, the data set extracted from the rock will contain too few crystals; those crystals will have exaggerated sizes and the mean nearest-neighbor distance will be too long. These errors will invalidate comparisons made to envelope simulations in which all crystals are regarded as separately observable, regardless of how close to one another they may be. To remedy this problem, the following observability criterion is employed in the envelope simulation runs applied to natural data sets.

The observability criterion is has two parts: the pair must satisfy both parts in order to be considered separately observable. The first part of the criterion is

$$d \geq (0.85)\, d_I\,, \tag{A2}$$

where d is the center-to-center distance between the two crystals, and d_I is the distance from the center of the larger crystal to the plane that contains the circle of intersection of the (spherical) crystal surfaces. The second part of the criterion is

$$l \geq 3\, r_S\,, \tag{A3}$$

where l is the total length of the pair of crystals and r_S is the radius of the smaller crystal. If the pair fails either of these tests, then it is concluded that the two crystals cannot be distinguished as separate from one another, and the later-placed crystal is repositioned randomly.

The values of the numerical constants in these tests were obtained by a tuning process that subjected actual data sets obtained from both CT imagery and time-explicit nucleation and growth simulations to the observability criterion. The numerical values were varied in order that the number of crystal pairs classed as “inseparable on observation” is maximized in the simulations and minimized in the CT-derived data sets. This is clearly appropriate because all the crystals in the CT-derived data sets were, in fact, separately observed and must therefore be clearly distinct from their neighbors.

This filter typically rejects a large proportion of attempted crystal placings as inseparable on observation, although this proportion is dependent on volume fraction. The mean percentage of rejections based on each criterion made in producing envelope simulations for each rock sample is given in Table A1.

After all crystals have been placed, the simulation is the best estimate for what the rock under analysis would look like if it had originated in an interface-controlled process. In this way, the

set of envelope runs becomes the null-hypothesis result of interface-controlled nucleation and growth to which one can legitimately compare the results for the rock under analysis.

Diffusion-controlled envelopes

Although, as stated in the text, we believe that the production of statistically rigorous null-hypothesis regions representing diffusion-control is at present computationally prohibitive, it is appropriate to discuss instances of these regions presented in the literature (Raeburn, 1996; Daniel and Spear, 1999). We believe that the algorithm used to produce these diffusion-controlled simulations is oversimplified, and fails to take into account important factors in the diffusion-controlled nucleation and growth process, such as impinging diffusional domains, and radial growth rates that vary with time.

The production of these regions is very similar to the production of the interface-controlled regions detailed above, but the placing criterion is adjusted to attempt to account for the expanded volume around each growing crystal in which nucleation is prohibited in a diffusion-controlled nucleation and growth regime. The adjusted placement criterion is:

$$d \geq (R_{dpl} / R) (r_L - r_S), \quad \text{(A4)}$$

where R_{dpl} is the radius of the depleted zone around a garnet at the conclusion of growth, R is the radius of the garnet, and other variables are the same as those in equation (A1). The ratio (R_{dpl}/R) is related to the volume fraction by:

$$(R_{dpl} / R) = (VF)^{-1/3}, \quad \text{(A5)}$$

where VF is the volume fraction of porphyroblast material in the sample. Note that in Daniel and Spear (1999), they use a volume fraction in this calculation modified from that given in their Table 1 to account for the portion of the sample volume occupied by quartz veins (Daniel, 1999, pers. comm.).

In order to test this method of production of “diffusion-controlled” envelope simulations, we subjected a simulation produced by a thermally accelerated, diffusion-controlled nucleation and growth algorithm to statistical analysis, but using equations (A4) and (A5) for production of the null-hypothesis region. As is clear from Figure A2, the envelopes differ significantly from the simulated data set, known to be diffusion-controlled, suggesting that the method of calculation of these “diffusion-controlled” null-hypothesis regions is incorrect. Although in Figure A2 the functional values for the known diffusion-controlled simulation fall below the envelopes calculated in this way, it is not clear that this will always be the case; the difference between the two may depend on factors such as crystal number density, volume fraction, the relative rates of nucleation and diffusion, and perhaps heating rate.

Table A1. Percent of crystal placings required by observability criteria to be repeated during production of envelope simulations.

	Volume	Percent re-placings	
Rock	fraction	Criterion 1	Criterion 2
PM1	0.380	7.4	15.8
PM2	0.108	0.4	1.1
WR1	0.245	1.7	1.3
WR2	0.420	10.9	6.1
WR3	0.269	7.5	6.9
WR4	0.264	2.4	1.1
MD	0.076	6.4	19.5

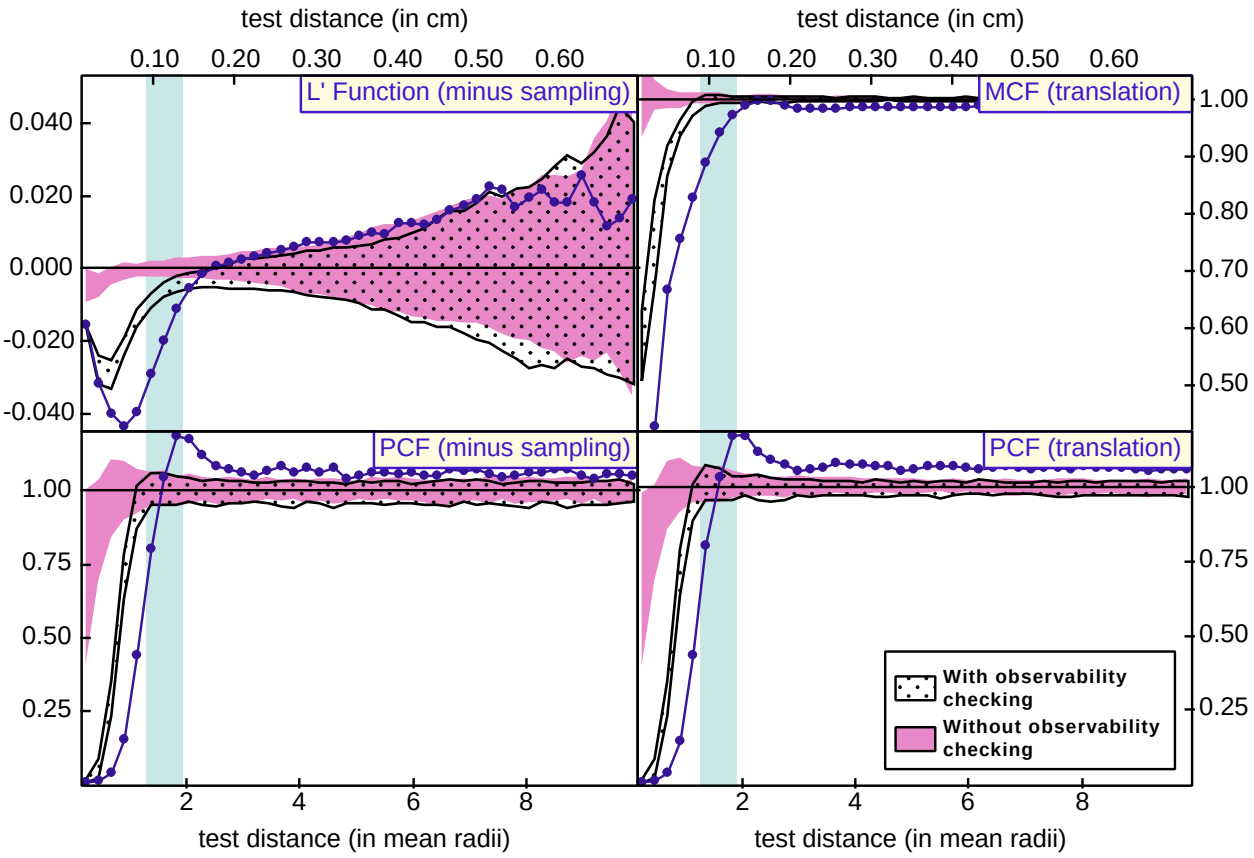

Figure A1. Correlation functions measured on sample PM-1 both with and without ensuring that the envelope simulations have the same observability criteria as the sample. The envelopes that incorporate observability criteria show significantly lower statistical values than those without. This indicates that the observability criteria produce more conservative conclusions, *i.e.*, they are less prone to provide spurious indications of ordering.

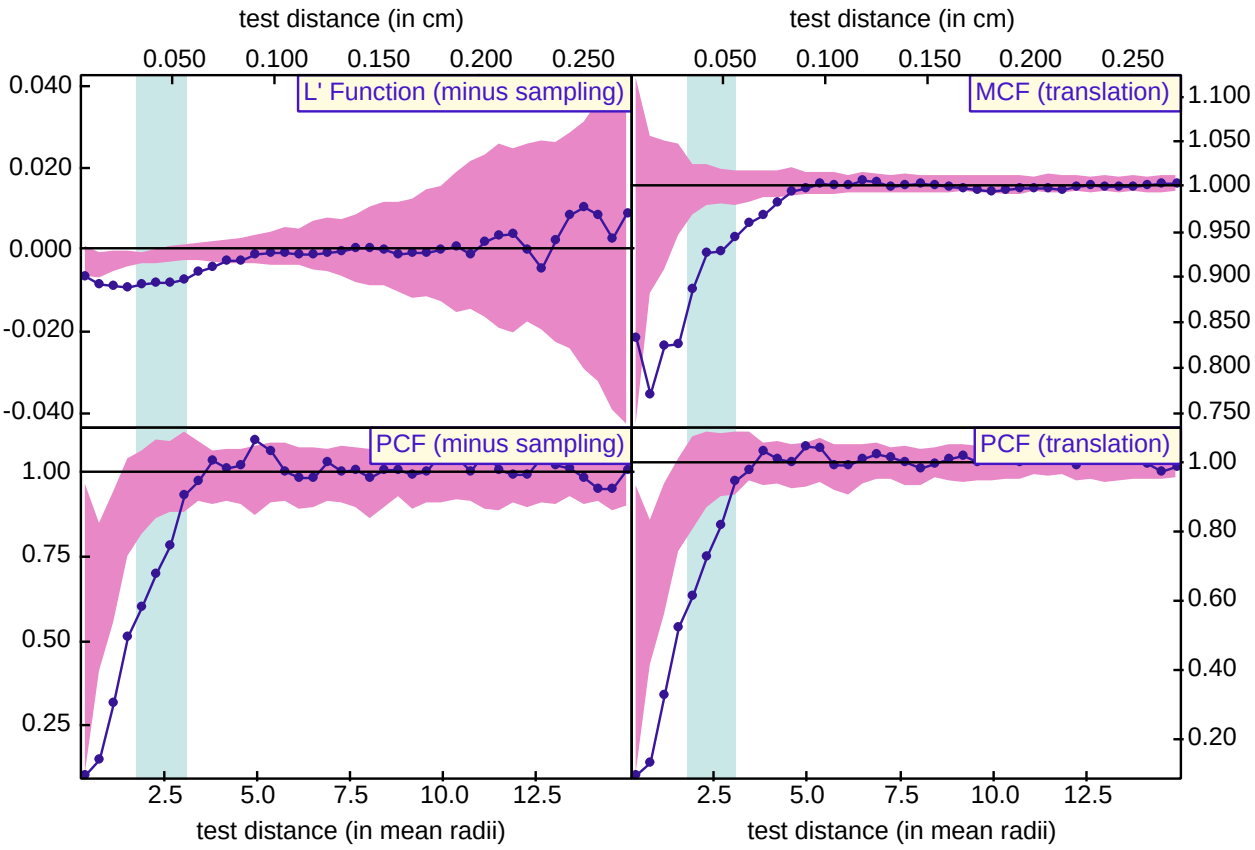

Figure A2. Correlation functions measured on thermally-accelerated nucleation and growth simulation DC-1, also shown in Figure 12. Here the shaded envelope is produced not by making a large number of interface-controlled simulations, but by the method given in Daniel and Spear (1999) for their "diffusion-controlled" null-hypothesis envelope. The disparity between the statistical values measured on a known diffusion-controlled simulation and those measured on the envelope cast doubt on the validity of this method for calculating a diffusion-controlled envelope.

Appendix B. Artificial arrangements of crystals

Case I: Random simulation. The default case that the statistics must reproduce is a random array of points, but because that has no geological relevance, we test the distribution most similar to that: an array of randomly-placed crystals with limited overlap. A random simulation was made of 1000 crystals in a 1-cm^3 cube, with random radii up to 0.05 cm and with later-placed crystals allowed to overlap earlier-placed crystals by an amount dictated by the interface-controlled growth criterion, that is, no crystal can have nucleated inside the volume of a pre-existing crystal (see Appendix A for a complete description). The results for this simulation are given in Figure B1. As expected, for both edge-correction methods, the L' -function is near the Poisson-distribution value of zero, and the PCF and MCF are near unity. The data fall within the 2-σ interface-controlled envelope, as one would expect, as this simulation is very nearly the same as those used in the envelope calculation. Just as one would expect based on the 2-σ criterion, few points fall outside the null-hypothesis envelopes. The minus-sampling edge-correction method has an increasing amount of noise relative to the translation method as the test distance increases, as it is being calculated with decreasing numbers of crystals.

Case IIa: Ordered simulation, no noise. To confirm that the statistics can reveal ordering trends in the data, an array of hexagonal-closest-packed crystal centers was tested, the most ordered array possible. An ordered simulation of 938 crystals in a 1-cm^3 cube was produced in which crystal centers are located in a hexagonal-close-packed array with a 0.001 cm offset in a random direction (for computational purposes), giving an inter-crystal distance of ~0.11 cm. As above, the radii are random up to 0.05 cm. The results for this simulation are given in Figure B2. Because these simulations are nearly perfectly ordered, they produce results atypical of natural samples. At small test distances, there are no crystals in the calculation. The L' function gives $-r$ for this region, and the PCF is undefined (here shown as a value of zero). These regions represent extreme ordering. The values for the L' -function and the PCF are high when the examination region first includes the twelve neighbors nearest to each crystal, and are periodic as the test distance increases, causing the examination region to encounter each “shell” of crystals followed by the empty region surrounding that shell. The low or undefined values at small r, together with periodic nature of the functions, distinguish these ordered simulations from the interface-controlled envelope simulations shown in Figure B1. Because the radii are random, the MCF gives a nearly constant and uninformative value; the positive excursion at ~7 mean radii is random noise accentuated by the sparsity of crystals at that separation.

Case IIb: Ordered simulation with random “noise”. To examine the sensitivity of the statistics to ordering effects, a number of simulations were performed using the above parameters, but displacing the crystals by an increasingly large vector in a random direction. The crystals were prevented from overlapping too much as dictated by the interface-controlled growth criterion (see Appendix A). As can be seen in Figure B3, the ordering trends in the data can be observed even after perturbing all the crystals by about half the original mean nearest-neighbor distance, 0.06 cm (this is also near the value of the mean nearest-neighbor distance in the perturbed array).

Cases IIIa-e: Clustered simulations. One of the strengths of correlation-function statistics is their ability to identify ordering and clustering that occur together but at different scales. Geologically, this might correspond to clustering at large scales induced by compositional layering (or other heterogeneities that affect the locations of potential nucleation sites), on which is superimposed shorter-range ordering induced by diffusional controls on crystallization. In order to determine the reliability and sensitivity of the correlation functions in such circumstances, simulations that combine clusters, ordering and randomness have been produced and analyzed. (Here and in the discussion below, clustering is discussed in the context of compositional layering, although the clustering in the simulations and figures is spherical for mathematical tractability. Because the examination regions are spherical and the imposed clustering is spherical, the trends observed in the data are likely to be more pronounced than those observed for layered samples. Still, the conclusions drawn from the data do not strictly rely on the spherical nature of the clusters, and will also apply to porphyroblasts that occur in layers or other kinds of clusters.)

Case IIIa: Ordered clusters, random within. A simulation of ordered clusters made up of randomly disposed crystals was analyzed (Fig. B4). This might correspond to interface-controlled growth in a highly layered sample, in which the layering was periodic. Although the distribution is random at small scales (within the clusters), the crystal number density within each cluster is greater than would be found if the complete distribution were random, and therefore at scales smaller than the clustering, the L'-function and PCF encounter more crystals than in the interface-controlled envelope simulations. Thus the values of these function lie above the interface-controlled envelopes. At test distances near the cluster radius (= 5.5 mean crystal radii), we begin to observe in the PCF the effects of ordering in the locations of the clusters.

Case IIIb: Random clusters, random within. Functional values for a simulation of randomly disposed clusters, each of which is made up of randomly disposed crystals, are given in Figure B5. This might correspond to interface-controlled growth in a highly layered sample, in which the layering was random. As expected, the data show the strong clustering present in the simulation. The ordering signal at distances greater than about ten mean radii is due to the relative sparsity of crystals at scales greater than that of a cluster; because the statistics are normalized to the bulk crystal number density, then just as within a cluster the local crystal number density is higher than the bulk value, leading to a positive excursion outside the envelope, so outside clusters the local crystal number density is lower than the bulk value, leading to a negative excursion outside the envelope.

Case IIIc: Ordered clusters, ordered within. Figure B6 presents functional values for a simulation of ordered clusters made up of ordered crystals. This might correspond to diffusion-controlled growth in a highly layered sample, in which the layering was periodic. At small test distances, up to that of the nearest-neighbor separation, ordering effects are observed. In this region, the functions take on their minimum values, zero for the PCF and $-r$ for the L'-function, because they encounter zero crystals separated by these distances. At intermediate distances, the clustering is observed in the data, shown most strongly by the L'-function. At scales near half the inter-cluster separation, ordering is again observed in the PCF, reflecting the larger-scale ordering in the data. As the scale of measurement expands to include the neighboring cluster,

clustering is again observed in the data. Thus we observe in this crystal array different types of distribution at three different scales.

Case IIId: Random clusters, ordered within. Functional values for a simulation of randomly disposed clusters made up of ordered crystals are plotted in Figure B7. This might correspond to diffusion-controlled growth in a highly layered sample, in which the layering was random. At the smallest test distances, smaller than the nearest-neighbor separation, we find ordering, as expected. At the nearest-neighbor separation, the positive excursion reflects the incorporation of the first nearest neighbors, and the periodicity that is present at greater test distances reflects the successive "shells" of neighbors at greater separations. The clustering effects are partially obscured by the existence of multiple clusters in contact, as can be observed in Figure B7a, but are evident in the L' function, whose gradual decline from about 0.20 cm to the mean nearest cluster neighbor distance at about 0.32 cm displays the clustering signature.

Case IIIe: Random clusters, ordered within, random "noise". To examine the sensitivity of the statistics to ordering and clustering effects, a number of simulations were performed, displacing the crystals in the simulation with the parameters listed in the preceding paragraph by increasingly large vectors in a random direction. As can be seen in Figure B8, the ordering trends in the data can still be detected after displacing the crystals by about 10 percent of the original nearest-neighbor distance. This is much smaller than the perturbation allowed in the similar case lacking clustering, case IIa above, emphasizing the negative effects inflicted by clustering on the ability to extract useful statistical conclusions from samples.

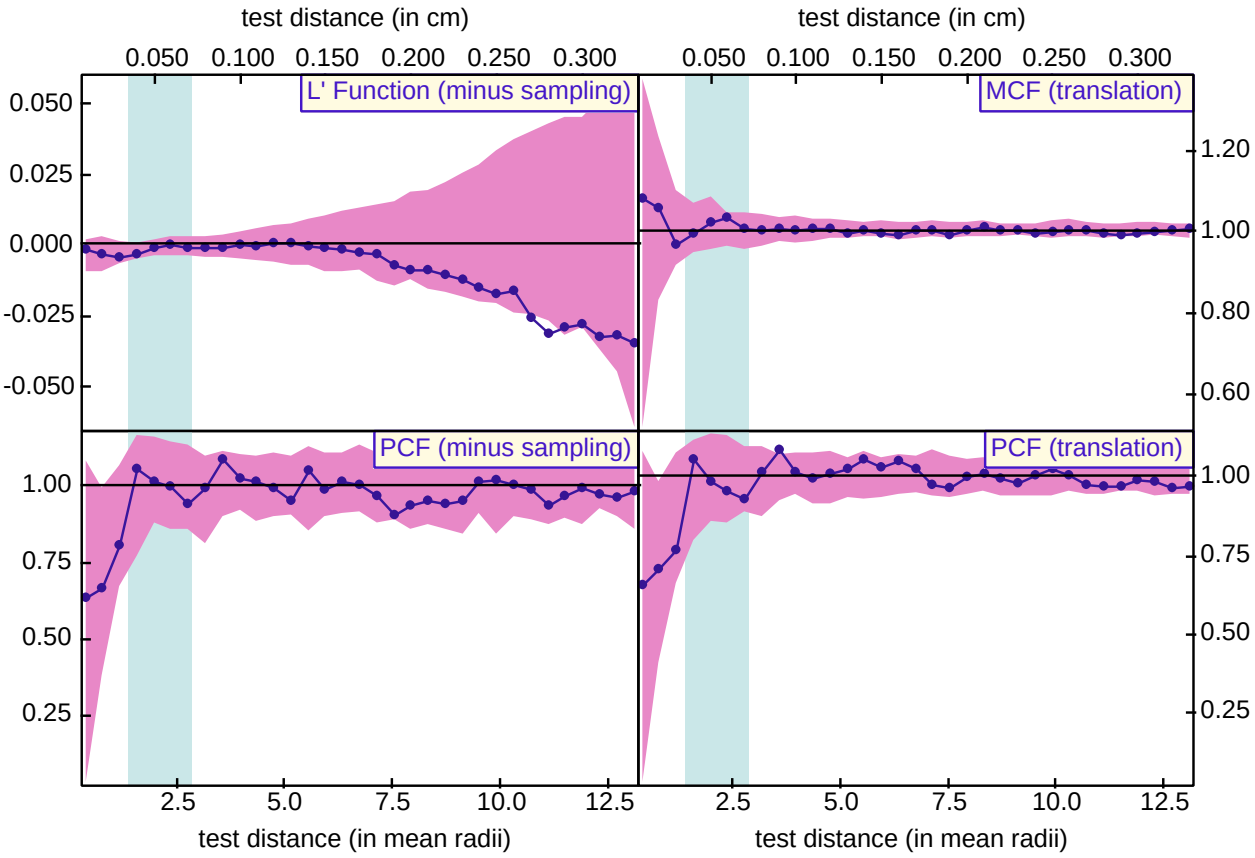

Figure B1. Correlation functions measured on a random array of 1000 crystals. Maximum overlap between crystals is given by the interface-controlled growth criterion. Radii are random over the interval [0, 0.05 cm]. Sample volume is 1 cm^3. The data fall within the 2-σ envelope, as expected.

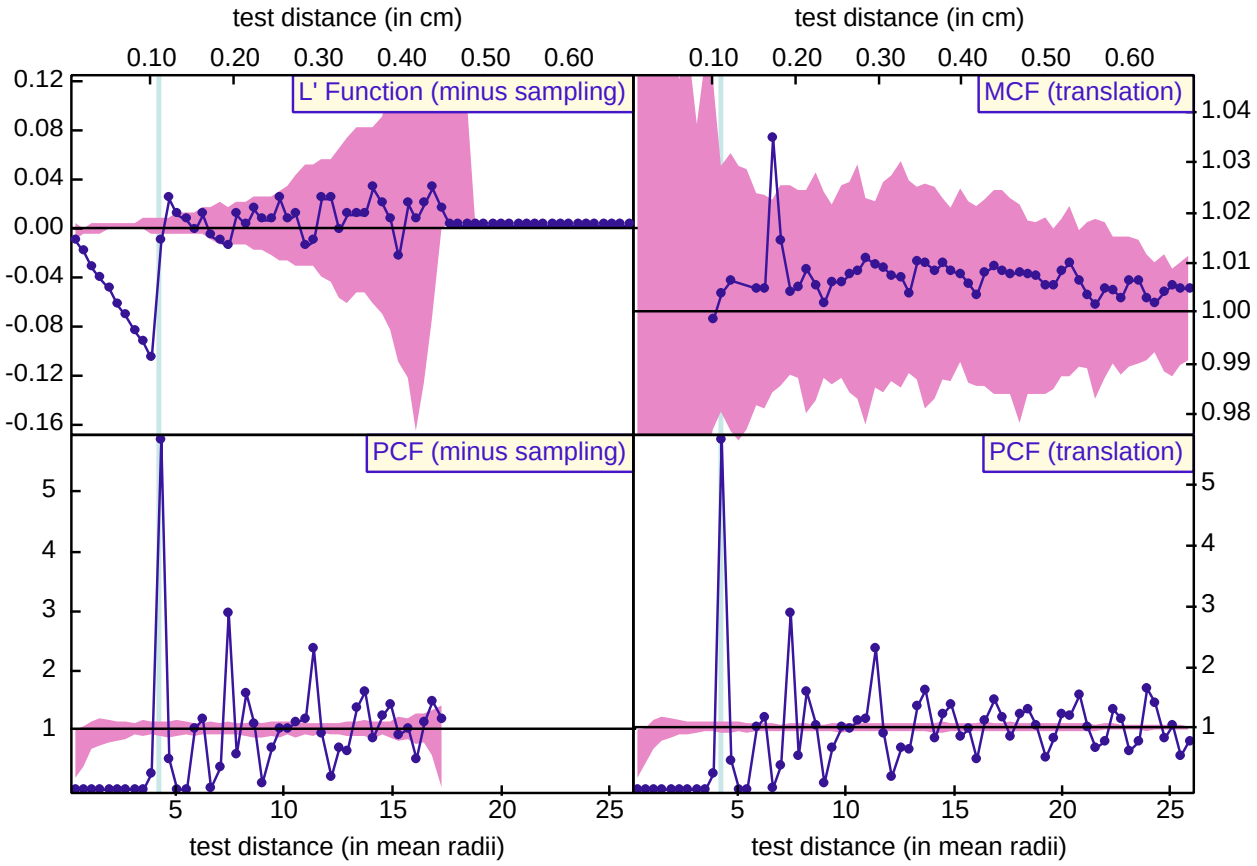

Figure B2. Correlation functions measured on an ordered array of 938 crystals. Crystal locations are dictated by hexagonal closest packing, with random offsets of 0.001 cm. Maximum overlap between crystals is given by the interface-controlled growth criterion. Radii are random over the interval [0, 0.05 cm]. Sample volume is 1 cm^3. The data show strong negative excursions outside the 2-σ envelopes, except at scales of the first, second, …, *n*th nearest neighbors.

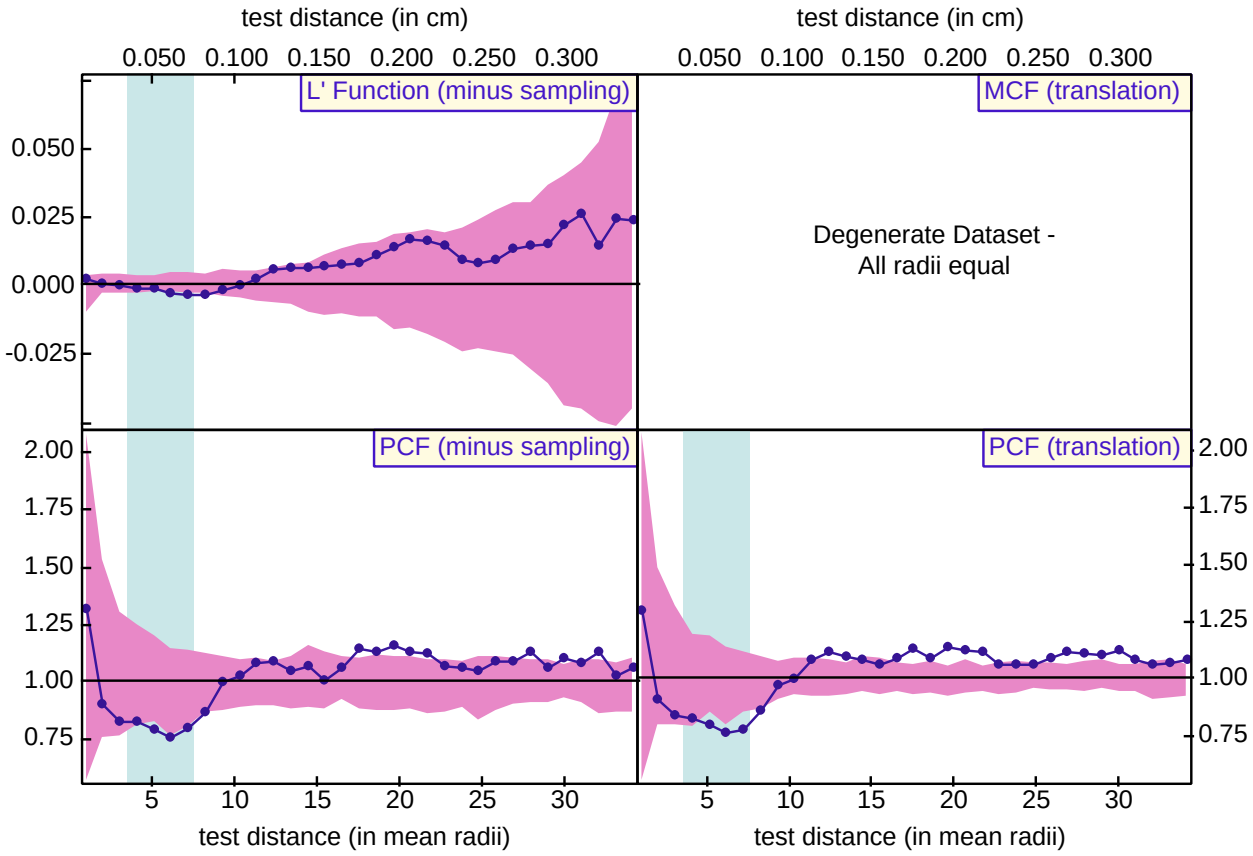

Figure B3. Correlation functions measured on the crystal array of Figure B2, but with each crystal displaced by a distance of 0.06 cm in a random direction in order to obscure the ordering signal, continue to show ordering effects at the scale of about the mean nearest-neighbor distance. In this model, all crystals have the same radius, so the MCF is degenerate.

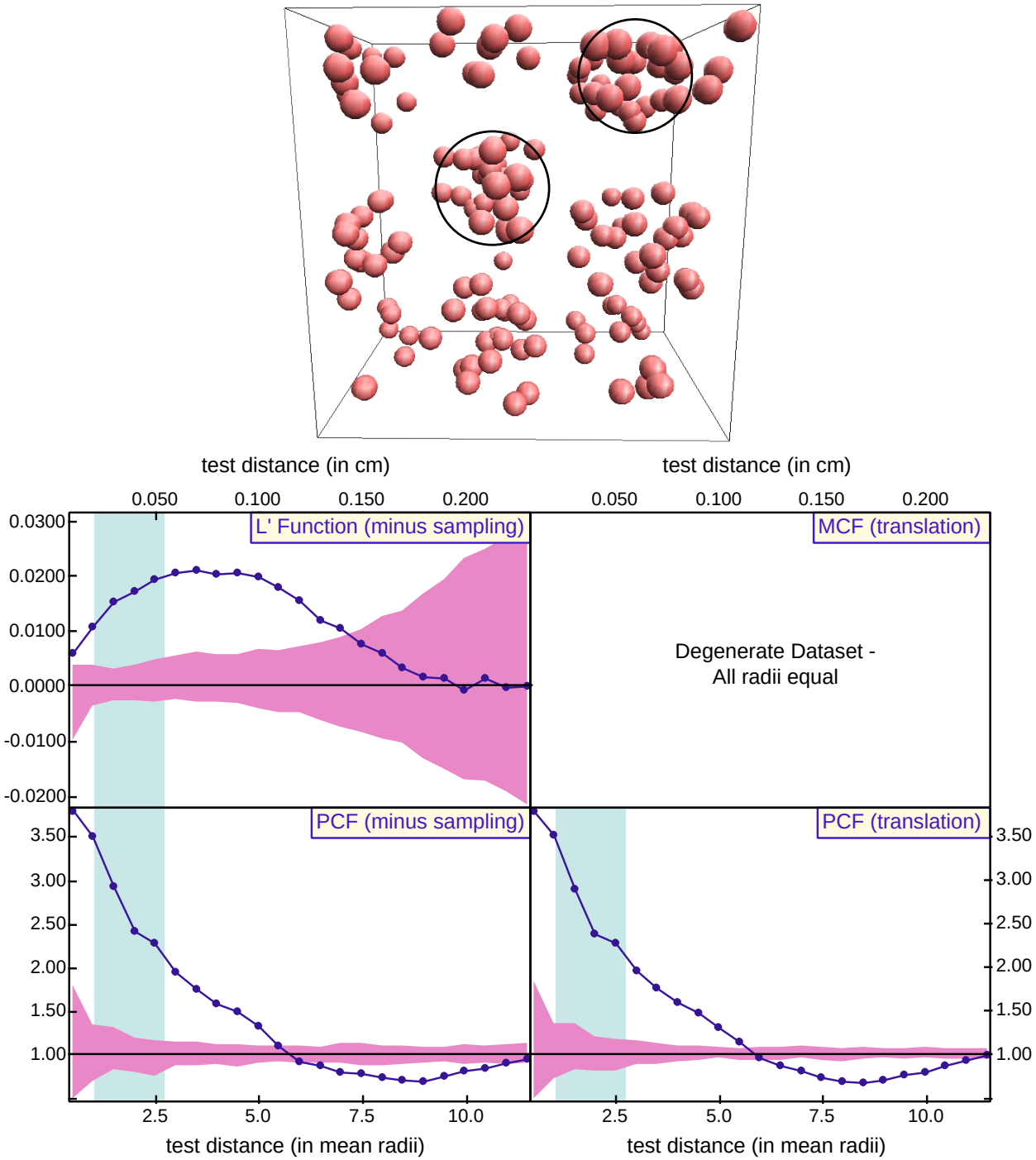

Figure B4. (a) Rendering of sub-volume of clustered simulation in which cluster centers are ordered using hexagonal closest packing, and crystal locations are random within each cluster; two clusters are circled. Cluster radii are 0.11 cm, with a center-to-center separation of 0.3 cm. There are 20 crystals per cluster. Crystal radii are all 0.02 cm (making the MCF degenerate). Sample volume is 1 cm^3. (b) Correlation functions measured on a clustered array of 980 crystals. Strong positive excursions outside the envelopes reflect the clustering of the crystals in the array, with the PCF showing ordering near the scale of the cluster radius (5.5 mean radii).

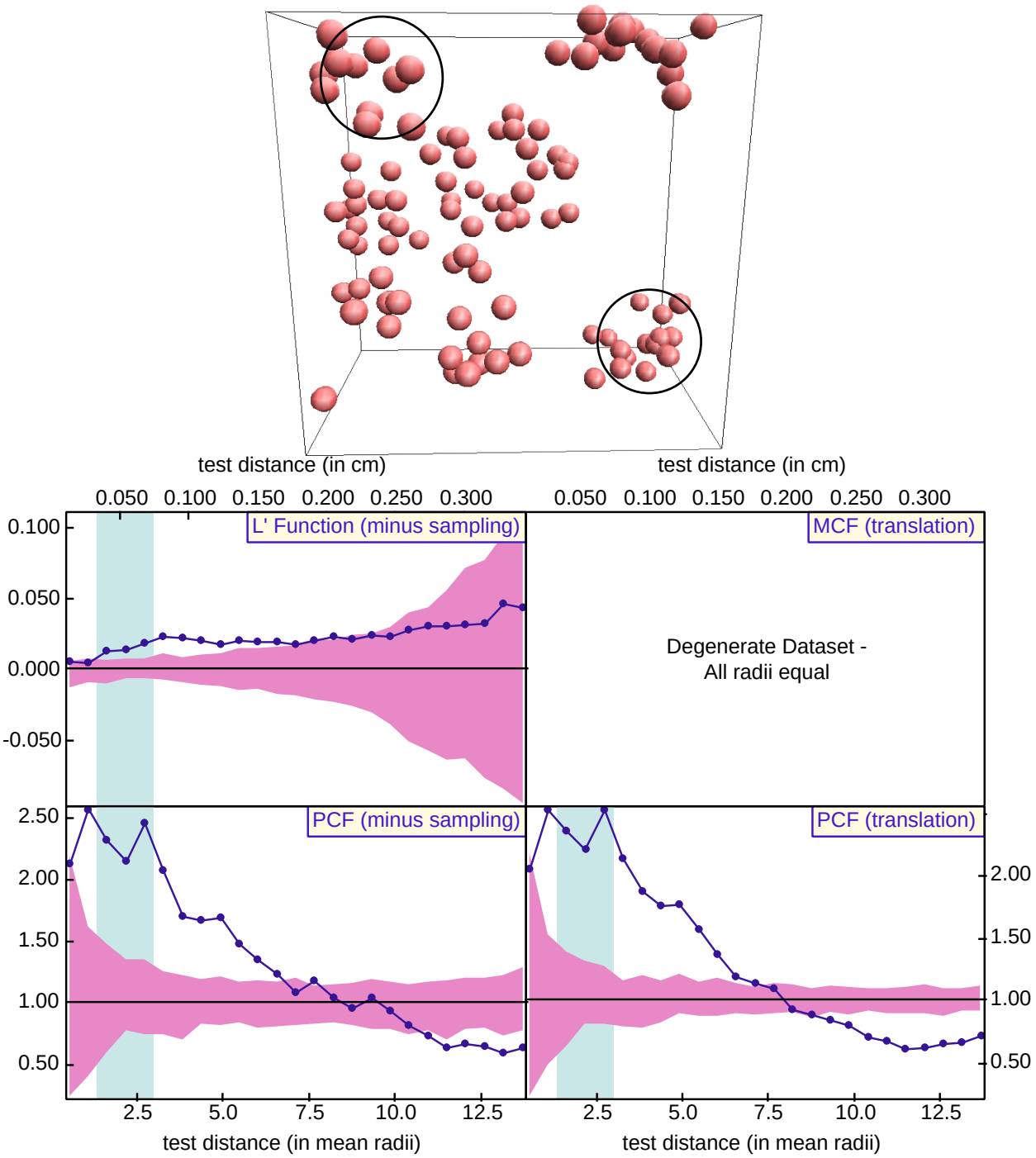

Figure B5. (a) Rendering of sub-volume of clustered simulation in which cluster centers are randomly disposed, as are crystal locations within each cluster; two clusters are circled. Cluster radii are 0.18 cm, and there are 20 clusters and 30 crystals per cluster. Crystal radii are all 0.025 cm (making the MCF degenerate). Sample volume is 1 cm^3. (b) Correlation functions measured on a clustered array of 388 crystals. Strong positive excursions outside the envelopes reflect the clustering of the crystals in the array, and the negative excursion in the PCF at large test distances reflects the sparsity of crystals outside a cluster relative to the mean crystal density.

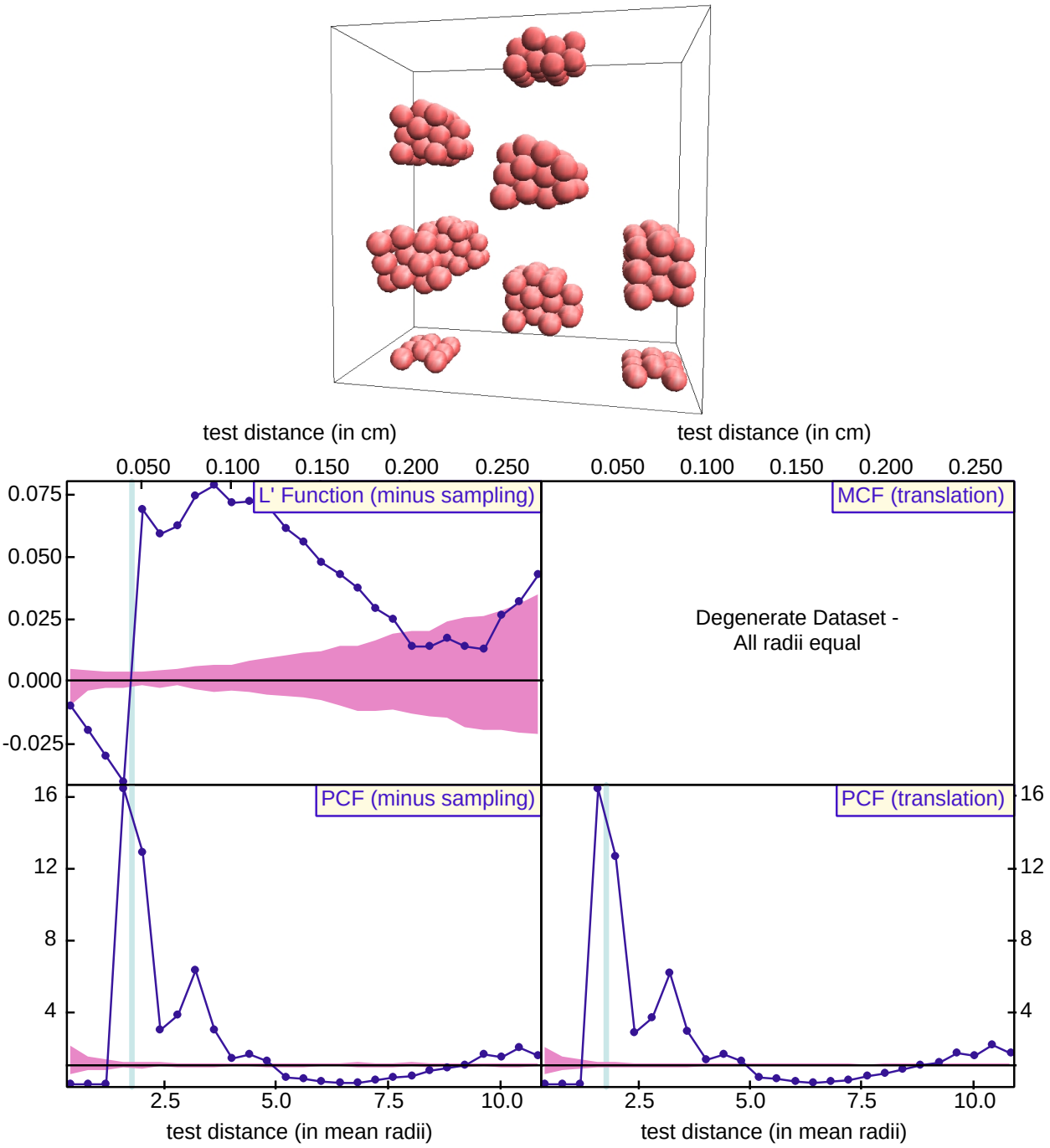

Figure B6. (a) Rendering of sub-volume of clustered simulation in which cluster centers are ordered using hexagonal closest packing, as are crystal locations within each cluster; there are 30 crystals per cluster. Cluster radii are 0.08 cm and center-to-center separation is 0.3 cm. Crystal radii are all 0.025 cm (making the MCF degenerate). Sample volume is 1 cm^3. (b) Correlation functions measured on a clustered array of 987 crystals. The data show alternating excursions below and above the 2-σ envelope; at the smallest scales of measurement, the excursion is below the envelope. This pattern reflects first the small-scale ordering within the clusters (below), then the clustering (above), then the space between the clusters (below) then the nearest neighboring cluster (above), and so on.

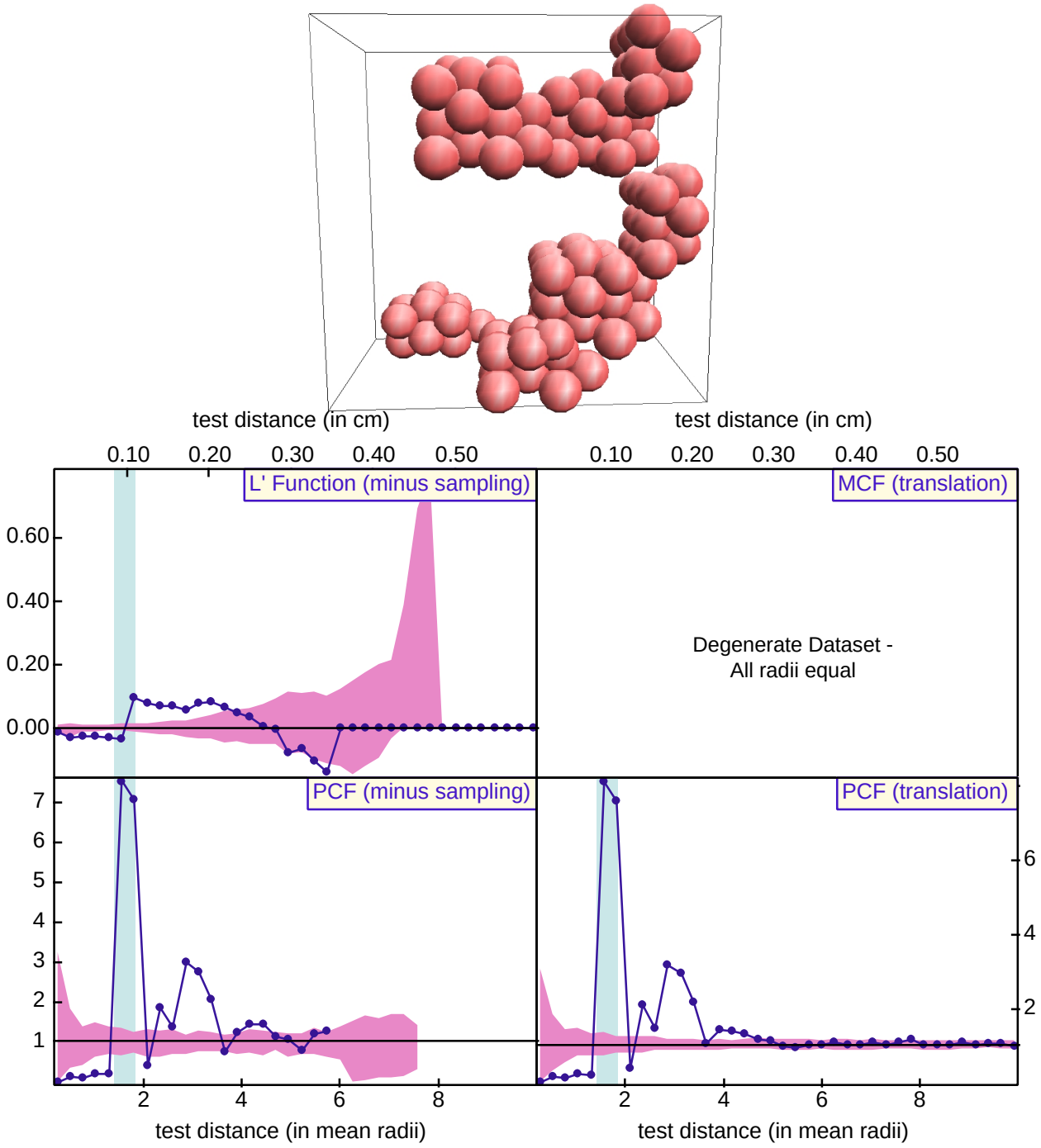

Figure B7. (a) Rendering of sub-volume of clustered simulation in which cluster centers are randomly disposed, and crystal locations are ordered within each cluster using hexagonal closest packing; there are 12 clusters, and 30 crystals per cluster. Cluster radii are 0.18 cm, and crystal radii are all 0.06 cm (making the MCF degenerate). Sample volume is 1 cm^3. (b) Results of correlation functions measured on a clustered array of 260 crystals. The strong positive excursion in the PCF reflects the inclusion of the first nearest neighbor in the calculation, the negative excursion reflects the absence of crystals separated by distances slightly greater than the nearest neighbor distance, the next positive excursion reflects the second nearest neighbor, and so on.

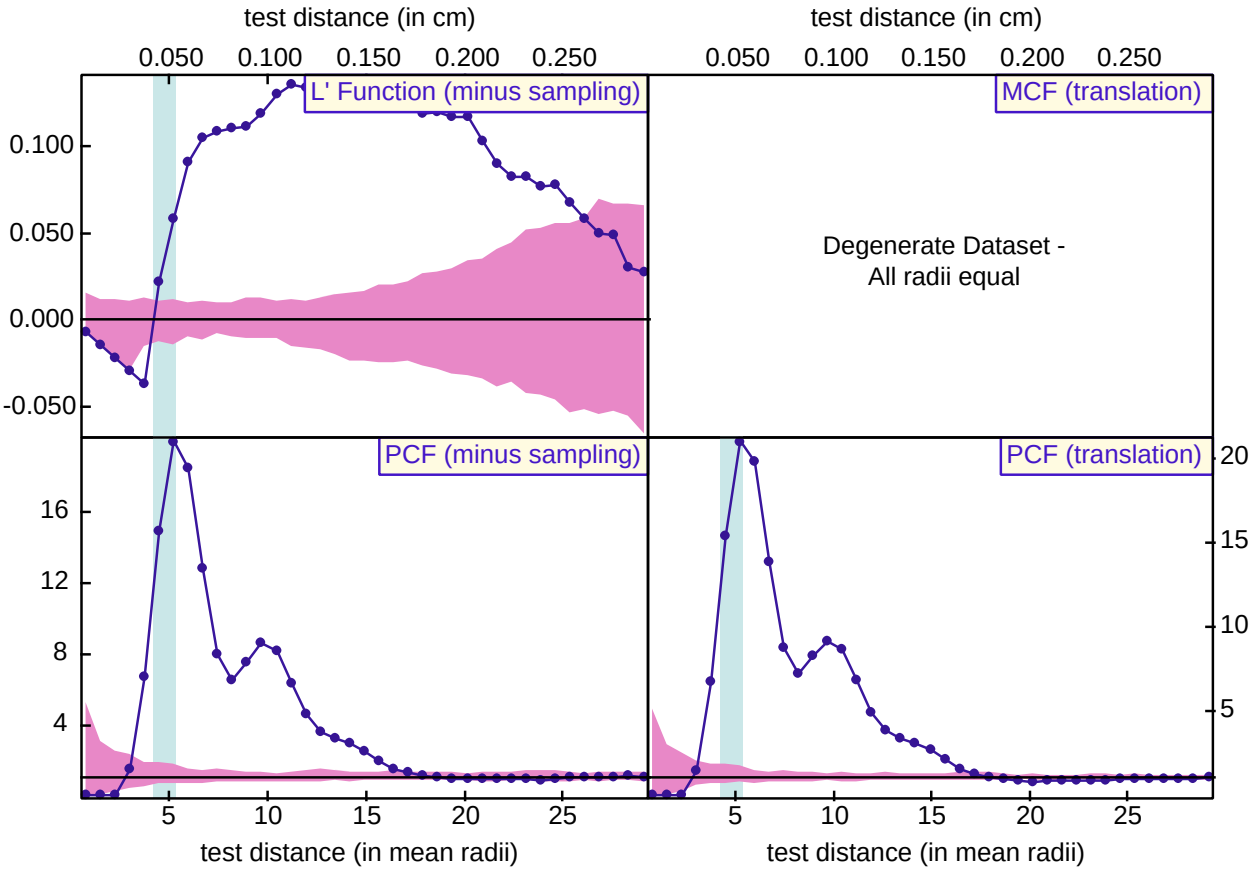

Figure B8. Correlation functions measured on the crystal array of Figure B7, but with each crystal displaced by 0.01 cm in order to obscure the ordering/clustering signal. The data continue to show the same basic pattern, with small-scale ordering within larger-scale clustering. The small-scale ordering is demonstrated not only by the single point in the *L*' -function that falls below the envelope, but also by the fact that at the smallest scales, both the *L*' -function and the PCF take on theit smallest possible values, indicating that there were no crystals observed at that separation.

Appendix C. Tests of robustness of correlation functions

Denison *et al*. (1997, Appendix 1) describe limitations on sample characteristics for three single-valued statistics: the ordering index, clustering index, and impingement index. They conclude that because of edge effects, a minimum of 1000 crystals is necessary to obtain reliable values for these statistics. A similar set of analyses was performed to evaluate appropriate sample sizes for application of correlation-function statistics.

Numbers of crystals

In order to determine the sensitivity of correlation functions to low crystal numbers, outer portions of the diffusion-controlled simulation shown in Figure 12 were progressively removed, and the resulting data repeatedly analyzed to determine at what amount of reduction the diffusional effects were no longer reliably observable. Because the locations and sizes of the crystals in the interior are controlled by the same factors as those near the edges that were removed, the value of the statistics would be expected to remain constant, although the noise should increase.

The results for this type of analysis are shown in Figure C1. They show that the arrays may be reduced to a small fraction of the original number of crystals without changing the conclusion reached by the statistical analysis using correlation functions. In the results shown, the number of crystals is reduced from 2971 to 184 without loss of the diffusion-control signal. Further reduction in the number of crystals leads to a different conclusion, or the data degenerates too far to be usable. This finding suggests that reliable results may be obtained from data sets as small as a few hundred crystals, although greater crystal numbers will greatly strengthen the confidence that can be placed in the inferences drawn.

Aspect ratios

A similar set of analyses was performed to examine the sensitivity to sample aspect ratio, which determines the magnitude of the edge effects: in this case, rather than the outer 10% of the sample being removed, only the top 10% was removed in successive calculations.

The results from this analysis (Fig. C2) are also encouraging: for the results shown, the number of crystals is reduced from 2971 to 310, and the sample aspect ratio increased from 1:1:1 to 1:10:10 without loss of the diffusion-control signal. Further reduction in the number of crystals, or increase in the aspect ratio, leads to a different conclusion, or the data degenerates too far to be usable.

It is important to note that these results are obtained with simulations, which have no complicating factors that may be found in real rocks, such as inhomogeneity of nucleation sites or nutrients for crystal growth. These factors may introduce enough noise into the data that at small numbers of crystals or large aspect ratios, any diffusion-controlled signal that may be present will be obscured.

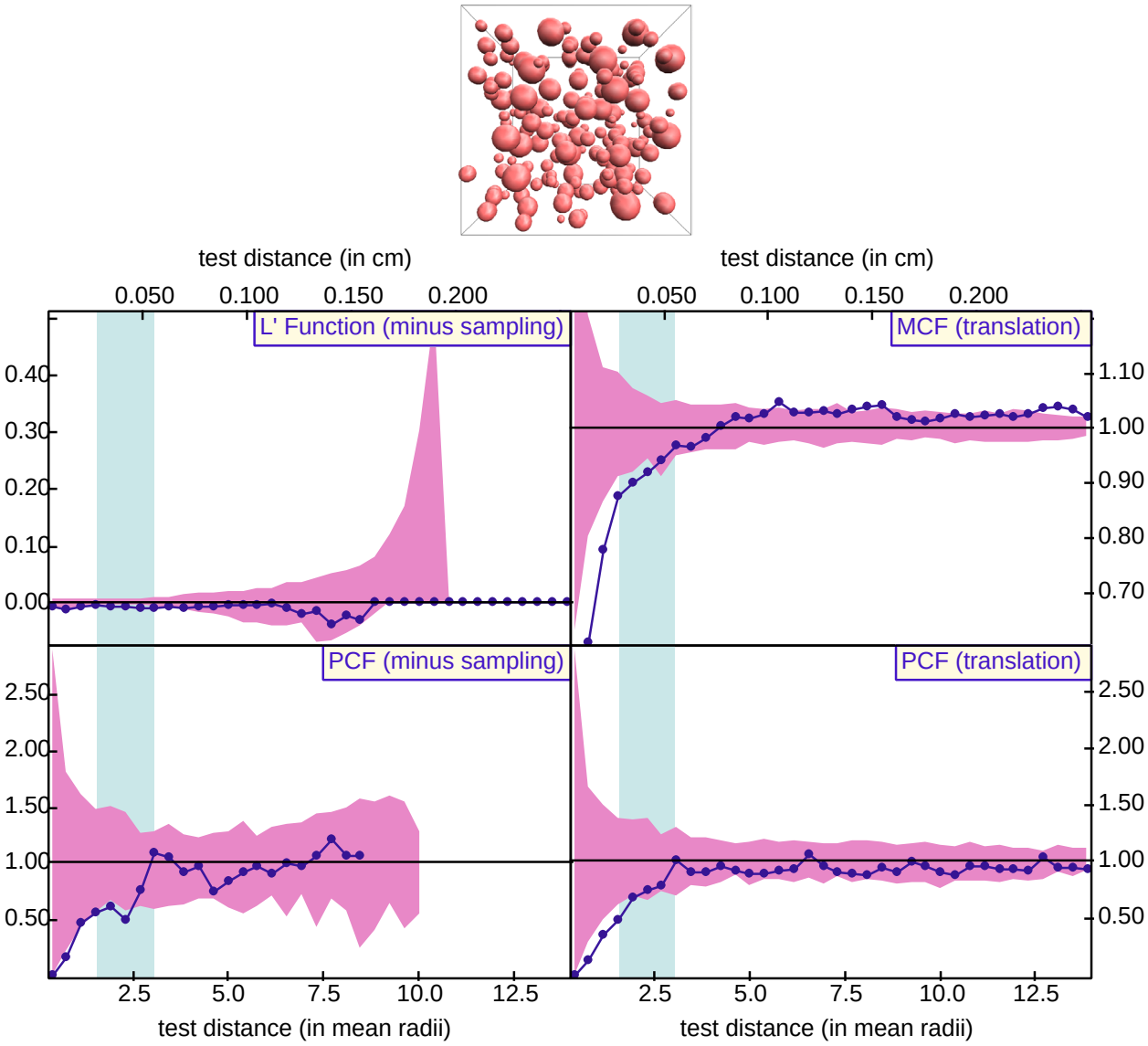

Figure C1. (a) Rendering of final remaining volume in DC-1; cube edge is 0.4 cm. (b) Correlation functions measured on remaining volume after removing outer crystals. DC-1 has 184 crystals remaining, yet the diffusional-control conclusion is still clear from the data.

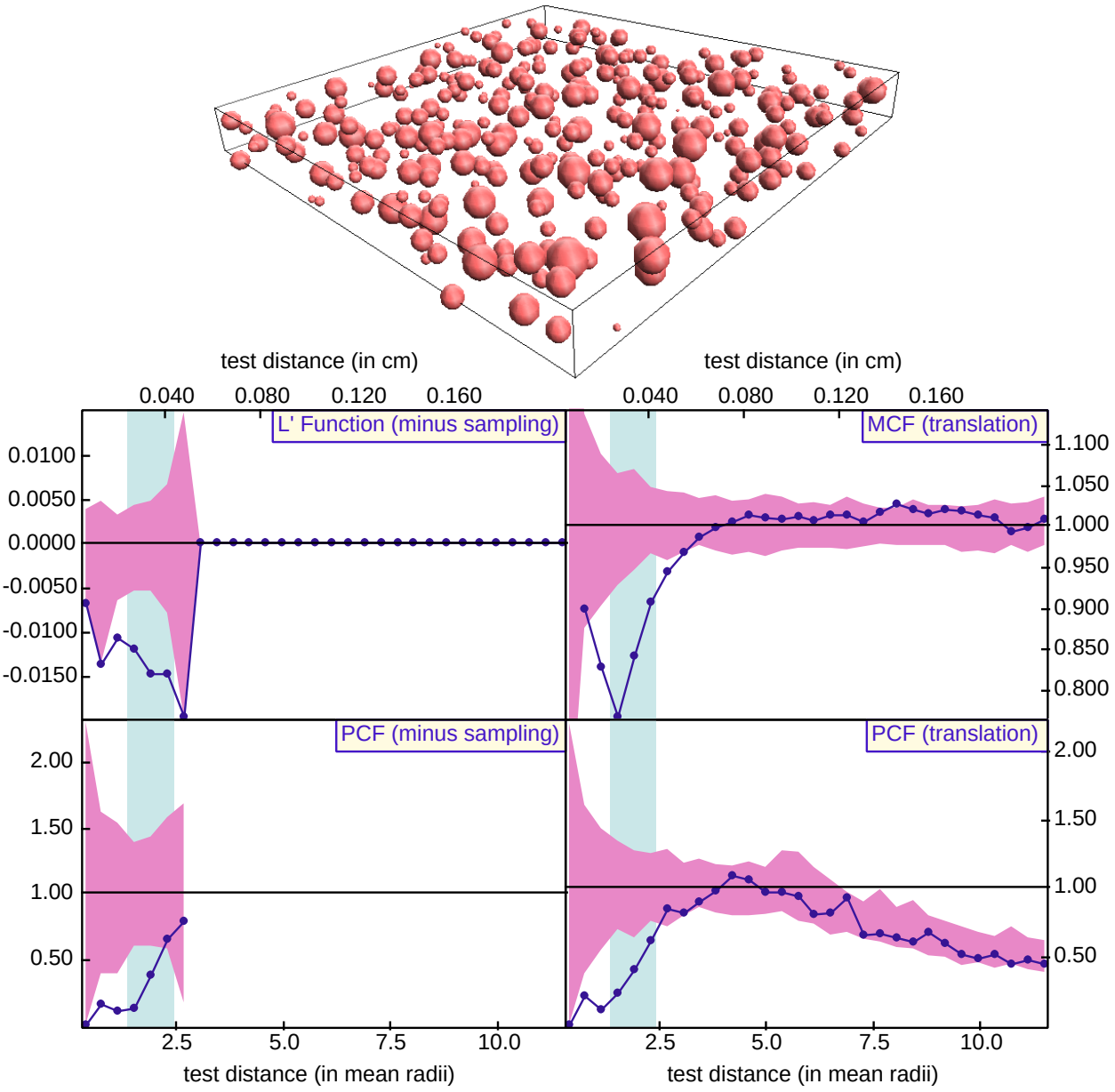

Figure C2. (a) Rendering of remaining volume in simulation DC-1. Vertical dimension is 0.1 cm while both horizontal dimensions are 1.0 cm. (b) Correlation functions measured on remaining volume after removing crystals from the top of simulation DC-1. DC-1 has 310 crystals remaining, yet the diffusional-control conclusion is still clear from the data.