

Received 28 January 2025

Accepted 16 March 2025

Edited by J. Ilavsky, Argonne National
Laboratory, USA

This article is part of a collection of articles
related to the 19th International Small-Angle
Scattering Conference (SAS2024) in Taipei,
Taiwan.

Keywords: structure factors; random packing;
parallel cylinders; X-ray/neutron scattering
intensity simulation.

Structure factors of random hard disk packing in 2D by explicit modeling

Yoshiharu Nishiyama*

Univ. Grenoble Alpes, CNRS, CERMAV, 38000 Grenoble, France. *Correspondence e-mail: yoshi@cermav.cnrs.fr

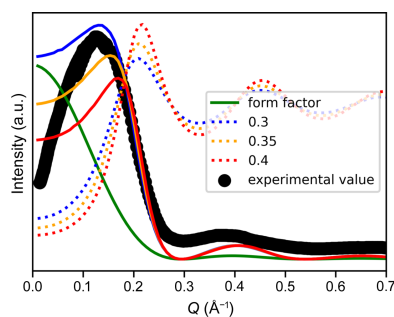
This paper describes a method that can (1) generate random packing of hard disks in 2D using Monte Carlo simulation, (2) extract the corresponding pair distribution function using normalization by disk line picking probability and (3) convert it to the structure factor. The generated structure factor agrees well with the analytical form based on the Percus–Yevick equation at a low area fraction (that is, within 1% at an area fraction below 0.2 and 2% at an area fraction of 0.3) but differs at a higher area fraction with more pronounced peaks and oscillations. Above an area fraction of 0.69, the hexagonal packing feature appears as sharp peaks at low Q , which are absent in the analytical solution. The structure factors up to an area fraction of 0.65 as a function of QR and the area fraction are stored in table form. The structure factor table can be combined with the cylinder form factors to simulate the X-ray/neutron scattering intensity of wood cell wall scattering.

1. Introduction

Determining the shape/dimension of the particles and their arrangement in a dense system of particles by small-angle scattering techniques is challenging since the characteristic features of form factors and structure factors appear on the same length scale, sometimes leading to misinterpretation of X-ray/neutron scattering data. An example is the packing of cellulose microfibrils in wood cell walls with a volume fraction of typically 30–50% (Rowell *et al.*, 2012). Ultimately, there is no unique solution to simultaneously determine the particle shape and structure factor if neither is known.

The arrangement of hard disks, the equivalent of hard spheres in 2D, is what Metropolis studied in his seminal work on the Monte Carlo algorithm in 1953 (Metropolis *et al.*, 1953). Since Monte Carlo simulation that involved explicitly calculating the arrangement of particles was expensive at the time, more analytical approaches using integral equations were also developed (Lado, 1968). The Percus–Yevick (PY) approximation is one such method that leads to an analytical pair distribution function (PDF) which shows relatively good agreement with the Monte Carlo simulation, except for the underestimation of the population at contact (Lado, 1968).

Rosenfeld (1990) proposed an even simpler analytical form that reproduces the structure factor of the PY approximation at low number density but fails at higher density. Here the structure factor



$$S_{\text{Rosenfeld}} = \frac{1}{1 + 4\phi \{A[J_1(Q)/Q]^2 + B[J_0(Q)J_1(Q)/Q + GJ_1(2Q)/Q]\}}, \quad (1)$$

where ϕ is the area fraction, J_n is the n th-order Bessel function and



OPEN ACCESS

Published under a CC BY 4.0 licence

$$G = (1 - \phi)^{-1.5}, \quad (2)$$

$$A = \frac{1 + (2\phi - 1)\chi + 2\phi G}{\phi}, \quad (3)$$

$$B = \frac{(1 - \phi)\chi - 1 - 3\phi G}{\phi}, \quad (4)$$

$$\chi = \frac{1 + \phi}{(1 - \phi)^3}. \quad (5)$$

Another possibility is to start with hexagonal packing and introduce an analytical form of disorder of the second kind (*i.e.* paracrystallinity) as proposed by Hashimoto *et al.* (1994). Penttilä *et al.* (2019) incorporated this in the WoodSAS model to fit X-ray scattering from wood. This allows the extraction of parameters such as lattice constants and disorder parameters. This paracrystalline model is adapted for a crystallizing system with well defined crystalline peaks that broaden at higher resolution. However, Penttilä *et al.* (2019) had to artificially trim the central scattering that tends to significantly deviate toward high intensity at low Q using the original mathematical model.

Since the structure factor of a hard disk liquid is a smooth function of the scattering vector magnitude Q , one can also establish a sparse tabulation to accurately reproduce the structure factor by a spline function. Here, we report a method to explore the liquid structure factor via Monte Carlo simulation and store the data in a readily utilizable form, which can be a more physical model than the ‘paracrystalline hexagonal packing’ in many cases. This type of packing is typical of fibrillar systems without a strong long-range interaction, for example sterically stabilized colloidal systems (Grelet & Rana, 2016) or biological systems such as collagen fibril packing (Meek & Boote, 2009).

2. Methods

2.1. Generation of random distribution

The Monte Carlo simulation performed here is very similar to what was originally used by Metropolis, with small modifications. The disks are initially positioned on a hexagonal array within a square periodic boundary box, and they are then moved in random steps under the condition of not colliding with another particle. In the following, we take the radius of the disk as a unit of length. In the case of a radius that is not one unit long, Q can be replaced by QR where R is the radius. Given a target volume (area) fraction ϕ_0 , the center-to-center distance, a , between disks in a hexagonal packing is

$$a = \sqrt{\frac{2\pi}{\phi_0\sqrt{3}}}. \quad (6)$$

We can fill a square with side length l , chosen as $l = 500$ here, with

$$n_x = \lfloor l/a \rfloor \quad (7)$$

disks in the horizontal direction and

$$n_y = \left\lfloor \frac{2}{\sqrt{3}} \frac{l}{a} \right\rfloor \quad (8)$$

discs in the vertical direction. This corresponds to a real volume fraction

$$\phi = \frac{\pi n_x n_y}{l^2}, \quad (9)$$

which is slightly smaller than ϕ_0 . The step size v_i along the x and y directions follows a normal probability function:

$$P(v_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{v_i^2}{2\sigma^2}\right) \quad [i = x, y]. \quad (10)$$

At each step, one randomly chosen disk is moved by a step (v_x, v_y) and we check for a collision with other disks. If the distance between two disks is smaller than two, this move is rejected and the disk returns to the initial position. σ is taken as $a/2 - 1$ but is reduced if more than 50% of the movements are rejected: 10N moves were applied here, where $N = n_x n_y$ is the number of disks, to check the rejection ratio. When the rejection rate was greater than 50%, σ was multiplied by 0.8 and the rejection rate was checked again. This operation was repeated until the rejection rate was below 50%.

2.1.1. Collision check

To minimize the number of distance calculations for collision check, the periodic boundary is divided into tiles with side length $\sqrt{2}$, which can accommodate only one disk. A unique serial number is assigned to each disk whose current coordinates are stored in memory. In total $(\lceil l/\sqrt{2} \rceil^2)$ tiles are represented in an integer array that hosts -1 if there is no disk at the position and the serial number of the disk if there is one. To check the collision, we only need to verify 20 tiles whose center-to-center distance is within a distance of 4, excluding the tile on which the disk that is moved sits, and calculate the distance only if there is a disk on the tile.

2.2. Pair distribution function

A histogram of the distances between particles for all pairs of particles is calculated within a circle centered on the periodic boundary box with a diameter of l . This corresponds to $T(N_0) = N_0(N_0 - 1)/2$ distance calculations, where N_0 is the number of particles in the circle. This histogram was normalized by $T(\overline{N}_0)$ using a theoretical number $\overline{N}_0 = (\phi/l^2)/4$ instead of N_0 , which fluctuates around the value $\overline{N}_0$. The width of the histogram bin w_b was 0.1, with the total number of bins $N_b = 1/w_b$. The normalized histogram is an array of probabilities p_i for each bin i , which sum to 1:

$$\sum_{i=1}^{N_b} p_i = 1.$$

The probability density of the distance between two random points, $\rho(r)$, in a circle with radius R , is (Mathai, 1999)

$$\rho(r) = \frac{4r}{\pi R^2} \left[\cos^{-1} \left(\frac{r}{2R} \right) - \frac{r}{2R} \sqrt{1 - \frac{r^2}{4R^2}} \right], \quad (11)$$

which can be further simplified in discrete form with $x_i = \delta_r r_i$ as

$$n_i = n(x_i) = \frac{16\delta_r}{\pi} \left(\cos^{-1} x_i - x_i \sqrt{1 - x_i^2} \right), \quad (12)$$

where n_i is the expected count in the i th bin and δ_r is the bin width divided by $2R$, or $1/N_b$.

Finally, the discrete PDF is obtained by dividing each histogram element by $n(x_i)$:

$$g(r_i) = \frac{p_i}{n_i}. \quad (13)$$

n_i only needs to be calculated once, and thus the calculation time is essentially spent obtaining the histogram. The discrete PDF was calculated and stored every $50N$ moves.

2.3. Structure factor calculation

The structure factor S of a two-dimensional isotropic or cylindrically symmetric system in the radial direction is (Oster & Riley, 1952)

$$S(Q) = 1 - \nu \int_0^\infty 2\pi r [1 - g(r)] J_0(Qr) dr, \quad (14)$$

$$= 1 + \nu \int_0^\infty 2\pi r h(r) J_0(Qr) dr, \quad (15)$$

where $g(r)$ is the PDF and $h(r) = g(r) - 1$ is the correlation function, J_0 is the Bessel function of zero order, and ν is the number density of the disks. $g(r)$ tends to 1 at large distance r , and thus $h(r)$ tends toward 0. Thus, integration can be stopped at some finite distance r . In the discrete form,

$$S(Q_i) = 1 + 2\pi\nu \sum_{j=0}^{j_{\max}} r_j J_0(Q_i r_j) h(r_j). \quad (16)$$

$r_j J_0(Q_i r_j)$ can be stored in a matrix

$$\mathbf{A} = (a_{ij}) = r_j J_0(Q_i r_j), \quad (17)$$

such that the structure factor array $\mathbf{S}$ can be obtained by simple matrix operation:

$$\mathbf{S} = 1 + 2\pi\nu \mathbf{A} \mathbf{h}, \quad (18)$$

where the i th element of $\mathbf{h}$ is $h(r_j)$ of equation (16). The extent of r , $r_{\max}$, was typically chosen to be $l/2$. $g(r)$ was linearly scaled down from $l/4$ to reach 0 at $l/2$ to reduce noise at low Q , but was also varied between 5 and 200 to check the influence on the calculated structure factor.

2.4. Spline fitting

Due to the oscillation in the very low Q region, the numerically calculated structure factors below $Q = 0.6$ were fitted to a parabolic function

$$S = m_2 Q^2 + m_0 b, \quad (19)$$

with the weighting proportional to Q^2 , and the direct results from the simulations were replaced with the values according to (19). The one-dimensional curves were smoothed with a Savitzky–Golay filter (Savitzky & Golay, 1964) using a window width of 29 and using the Python package *savgol_filter* from the *SciPy* library (Virtanen *et al.*, 2020) to decrease the numerical noise. Curves over the whole Q range from 0 to 10 and for an area fraction from 0.05 to 0.65 were approximated with a bivariate smoothing second-order B spline using the *bisplrep* function from *SciPy*. The weight was set to unity for all points with smoothing factor 0.1.

3. Results

3.1. Pair correlation function

Fig. 1 shows the averages of 100 consecutive pair correlation functions calculated for an area fraction of 0.6 [Fig. 1(a)] and 0.71 [Fig. 1(b)], sequentially selected from 1000 data points. There is an almost perfect superposition up to $r = 250$ but, at larger distances, slow fluctuation can be seen in addition to the white noise close to 500 because of the small sampling number. Also, due to the periodic boundary conditions, disks at a distance well beyond half the box size correlate as they are at a shorter distance on the other side of the boundary. Thus, a small oscillation can be seen close to the maximum distance limit.

The corresponding short-distance part of the pair correlation function is enlarged in Figs. 1(c) and 1(d), in which all curves exactly superpose for a given system.

3.2. Influence of integration limit on the structure factors

Fig. 2 shows the average structure factors corresponding to the pair distributions in Fig. 1 calculated according to equation (18) using the different extents of r , equivalent to truncating the right side of the matrix $\mathbf{A}$ and the bottom of the vector $\mathbf{h}$. Structure factors were calculated for 1000 PDFs and then averaged. At an area fraction of 0.6 [Figs. 2(a) and 2(c)], there is almost no difference beyond the integration limit of 25, except at very small Q where the oscillation started at a smaller Q with a larger integration limit. However, at higher density with an area fraction of 0.71, taking a smaller limit than 100 starts to smear out the peak around $Q = 5.6$ [Fig. 2(d)] and also introduces oscillation around the main peak.

3.3. Structure factors as a function of area fraction

Fig. 3 shows the structure factors of the disks as a function of area fraction, which exhibit more and more pronounced features with increasing area fraction. At an area fraction of 0.71, three sharp peaks at $Q = 3.24$, 5.61 and 6.47, which

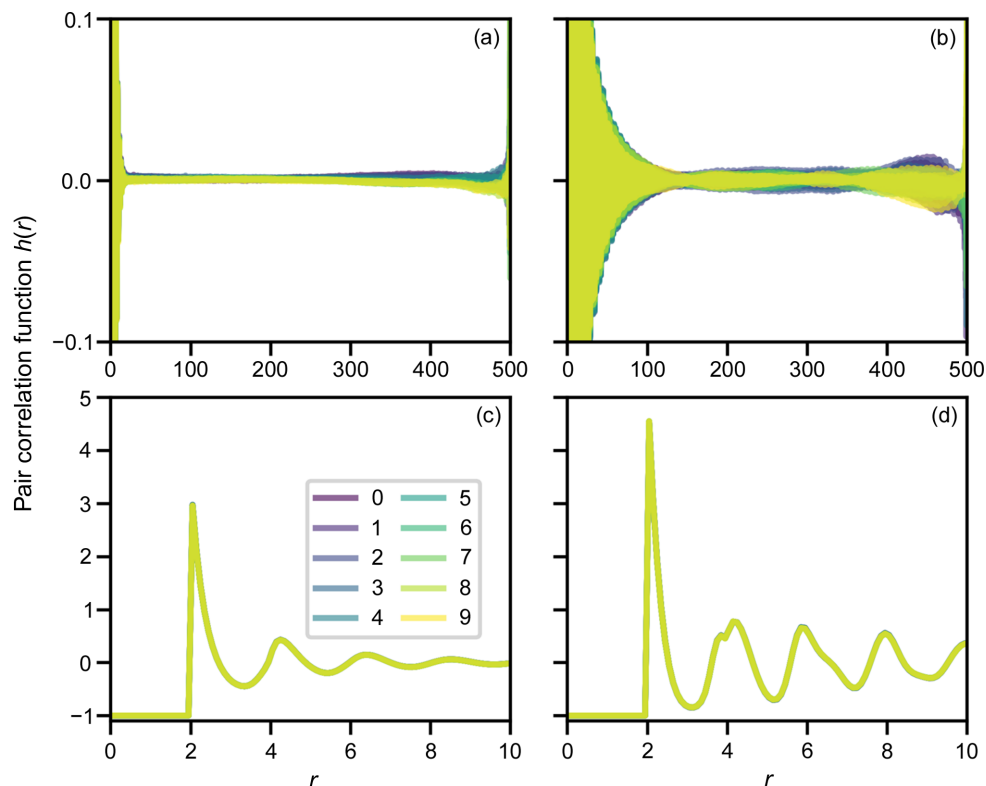


Figure 1 Pair correlation function of random disks with the area fractions (a) 0.6 and (b) 0.71. Each curve is an average of 100 consecutive pair correlation functions; (c) and (d) are enlargements of the small-distance regions of (a) and (b), respectively. The 10 lines correspond to 10 segments of the run.

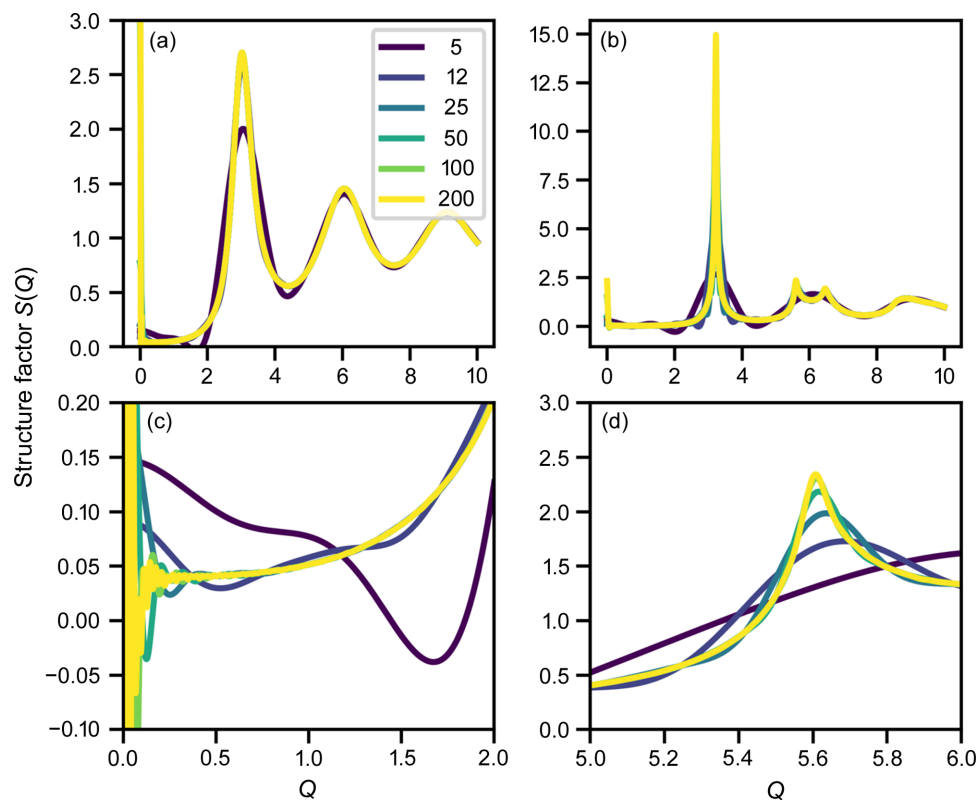


Figure 2 Structure factors corresponding to Fig. 1, with the area fractions (a) 0.6 and (b) 0.71, calculated with different integration limits in r , from 5 to 200, distinguished by color; (c) and (d) are zoomed-in images of (a) and (b), respectively. $Q = QR$ as $R = 1$.

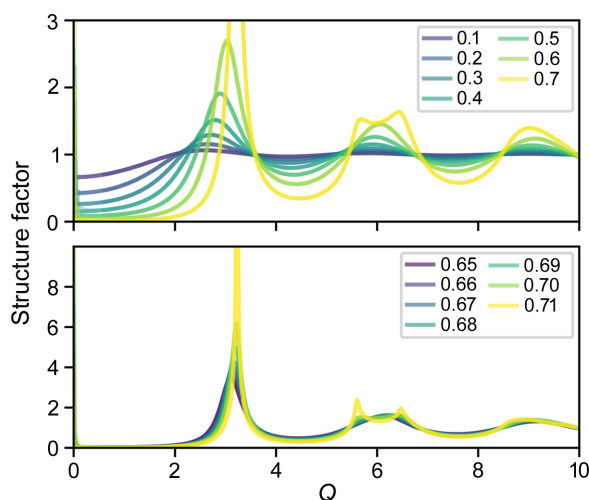


Figure 3
Structure factors of random disks at different area fractions. Spline fitting curves are superposed on the simulation data as faint solid lines (visible only at low Q).

correspond to the typical Q ratios of 1, $\sqrt{3}$ and 2 of a hexagonal lattice, are present. The diffraction features fade away at higher Q .

4. Discussion

4.1. Comparison with the previous structure factor calculations

The PDF and the structure factor at an area fraction of about 0.6 agree well with that reported by Lado (1968) using a 192-molecule constant-pressure system. The structure factor based on this calculation reported by Rosenfeld (1990) at $\phi = 0.6$ also agrees with our calculation of the structure factor. Using a larger system than Lado (1968) is unnecessary for this area fraction because the pair correlation function fades out quickly, as seen in Fig. 1.

Above $\phi = 0.69$, a larger system size was necessary to capture the structure since the pair correlation function extends over a long distance, giving rise to a few sharp peaks corresponding to the hexagonal lattice. In addition, the structure evolves slowly at this level of crowding and might be jammed into a state dependent on the sample's history. A more disordered state might be achievable by other methods, such as the expansion of disks from a given configuration to achieve a high area fraction (Weber *et al.*, 1995; Torquato *et al.*, 2000), which is not addressed in this work.

Fig. 4(a) shows a general view of the interpolated structure factors within the simulated range. The tabulated structure factor data are provided in binary form to be directly loaded as a *numpy* array. Fig. 4(b) shows the ratio of the interpolated structure factor over $S_{\text{Rosenfeld}}$ of equation (1). The structure factor ratio at $\phi = 0.4$, marked in black, shows that the relative difference from the analytical form is within 5% at this area fraction. At smaller ϕ , the agreement is even better: below 2% when $\phi < 0.3$ and below 1% when $\phi < 0.2$. The

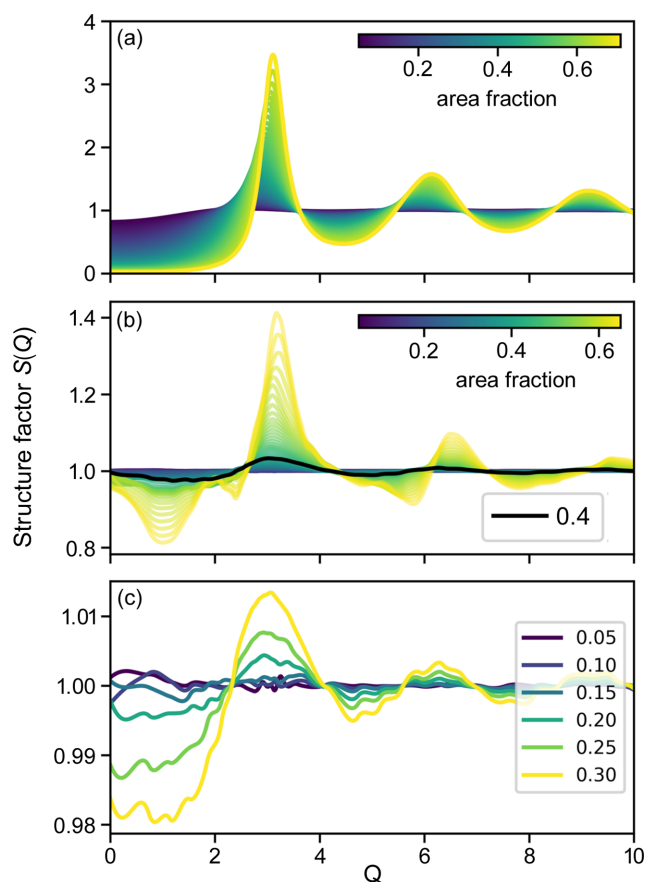


Figure 4
Smoothed and interpolated curves of structure factors as a function of area fraction from 0.05 to 0.7. (a) Interpolated structure factors. (b) Relative difference from the analytical function provided by Rosenfeld (1990) in the ϕ range 0.05 to 0.65. (c) Same as (b) but for ϕ below 0.3.

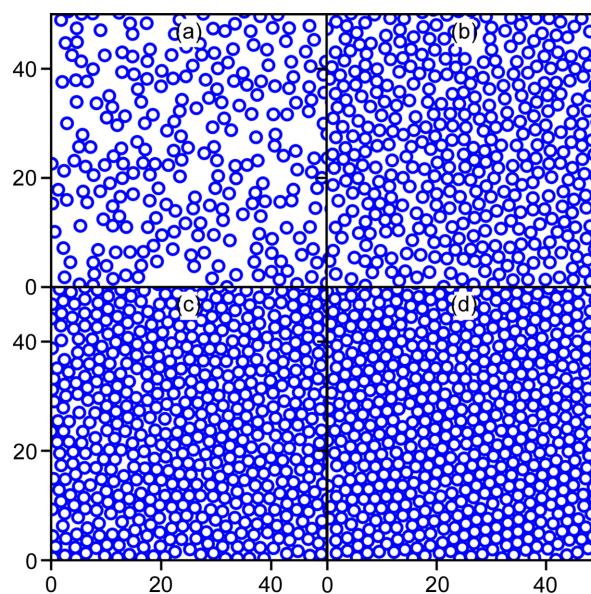


Figure 5
Snapshots of randomly distributed disks at the area fractions (a) 0.35, (b) 0.5, (c) 0.65 and (d) 0.7.

deviation becomes more and more pronounced at higher disk densities.

One of the advantages of explicit simulation is visualization in real space. Fig. 5 shows snapshots of the disk dispositions at different disk densities. It can be seen that up to an area fraction of 0.5 [Figs. 5(a) and 5(b)] there is not much occurrence of hexagonal arrangements. At higher densities [Figs. 5(c) and 5(d)], the figures show some regularity and impression of the ‘crystalline’ zone, but even at such regular filling, the structure factor fades away quickly as seen in Figs. 3 and 4(a).

4.2. Interpretation of X-ray correlation peak

Fig. 6 shows the experimental X-ray scattering of native birch wood. When the anisotropic component is isolated and the orientation correction applied, there is a clear correlation peak at $Q = 0.15$. This peak is often interpreted as representing the interfibrillar distances. However, this peak can also be reproduced by multiplying the structure factor with the cylinder form factor

$$P(RQ) = \left[\frac{J_1(QR)}{QR} \right]^2, \quad (20)$$

drawn in green in the figure arbitrarily assuming $R = 1.3$ nm. The dotted profiles show the structure factors calculated above, replacing Q with QR for different area fractions. The expected intensity profiles

$$I = P(RQ)S(RQ) \quad (21)$$

are shown as solid lines. The peak position of the scattering does not correspond to the peak of the structure factor, which is about $0.21 \text{ \AA}^{-1}$ (2.75 on the QR scale) in this area fraction range. However, the peak of the product of the structure factor and form factor is much closer to that of the experi-

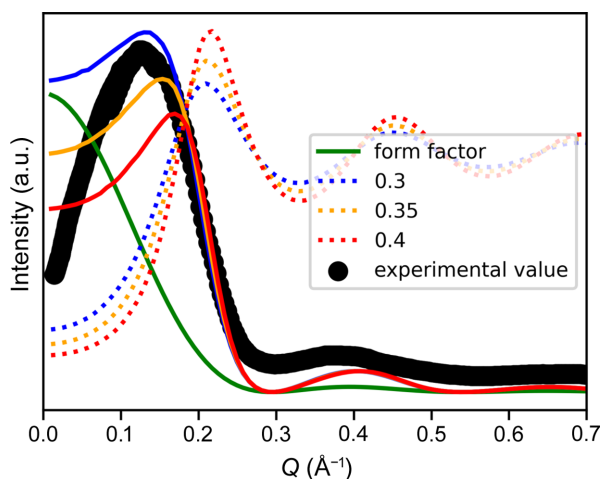


Figure 6

Comparison of X-ray small-angle scattering data of birch wood reported by Chen *et al.* (2021) with calculated structure factors and the form factor assuming a random packing of infinitely long cylinders with a radius of 13 Å. The intensity is multiplied by Q to compensate for the measured intensity being an average over an azimuthal angle, while the model assumes a line trace of a perfectly aligned system.

mental scattering data, and the position shifts to lower Q for a lower volume fraction. The same type of small-angle peak is reported (Kuribayashi *et al.*, 2023) in a wide range of wood samples with some variation in peak positions, which can also be changed by hydrothermal treatment.

5. Conclusions

An exhaustive simulation of the random packing of hard disks in a plane confirmed the validity of Rosenfeld’s analytical expression of the structure factor at small area fractions $\phi < 0.3$. However, the analytical form deviates at higher volume fractions. The interpolation of structure factors calculated for a limited number of models provides an alternative to the analytical form and allows further interpretation of scattering from dense systems where the form factor and structure factor interfere. This approach can be further extended for softer interparticle interactions with small modifications in the Monte Carlo procedure.

Acknowledgements

The author thanks Dr Doru Constantin from the Institut Charles Sadron and Dr Qiang Zhang from the Laboratoire Léon Brillouin for fruitful discussions and the corrected form of Rosenfeld’s structure factor formula, and Professor Pan Chen for providing the X-ray scattering data in numerical form.

Data availability

The programs that generate the structure factors and the structure factor table in binary form are available at https://github.com/yoshi-CERMAV/hard_disk_rdf.

References

- Chen, P., Li, Y., Nishiyama, Y., Pingali, S. V., O’Neill, H. M., Zhang, Q. & Berglund, L. A. (2021). *Nano Lett.* **21**, 2883–2890.
- Grelet, E. & Rana, R. (2016). *Soft Matter*, **12**, 4621–4627.
- Hashimoto, T., Kawamura, T., Harada, M. & Tanaka, H. (1994). *Macromolecules*, **27**, 3063–3072.
- Kuribayashi, T., Ogawa, Y., Morfin, I., Matsumoto, Y. & Nishiyama, Y. (2023). *Cellulose*, **30**, 8405–8413.
- Lado, F. (1968). *J. Chem. Phys.* **49**, 3092–3096.
- Mathai, A. M. (1999). *An introduction to geometrical probability*. OPA.
- Meek, K. M. & Boote, C. (2009). *Prog. Retin. Eye Res.* **28**, 369–392.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. (1953). *J. Chem. Phys.* **21**, 1087–1092.
- Oster, G. & Riley, D. P. (1952). *Acta Cryst.* **5**, 272–276.
- Penttilä, P. A., Rautkari, L., Österberg, M. & Schweins, R. (2019). *J. Appl. Cryst.* **52**, 369–377.
- Rosenfeld, Y. (1990). *Phys. Rev. A*, **42**, 5978–5989.
- Rowell, R., Pattersen, R., Han1, J. S., Rowell, J. S. & Tshabalala, M. A. (2012). *Handbook of wood chemistry and wood composites*, 2nd ed., pp. 35–74. CRC Press.
- Savitzky, A. & Golay, M. J. E. (1964). *Anal. Chem.* **36**, 1627–1639.
- Torquato, S., Truskett, T. M. & Debenedetti, P. G. (2000). *Phys. Rev. Lett.* **84**, 2064–2067.

- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., Vijaykumar, A., Bardelli, A. P., Rothberg, A., Hilboll, A., Kloeckner, A., Scopatz, A., Lee, A., Rokem, A., Woods, C. N., Fulton, C., Masson, C., Häggström, C., Fitzgerald, C., Nicholson, D. A., Hagen, D. R., Pasechnik, D. V., Olivetti, E., Martin, E., Wieser, E., Silva, F., Lenders, F., Wilhelm, F., Young, G., Price, G. A., Ingold, G., Allen, G. E., Lee, G. R., Audren, H., Probst, I., Dietrich, J. P., Silterra, J., Webber, J. T., Slavič, J., Nothman, J., Buchner, J., Kulick, J., Schönberger, J. L., de Miranda Cardoso, J. V., Reimer, J., Harrington, J., Rodríguez, J. L. C., Nunez-Iglesias, J., Kuczynski, J., Tritz, K., Thoma, M., Newville, M., Kümmerer, M., Bolingbroke, M., Tartre, M., Pak, M., Smith, N. J., Nowaczyk, N., Shebanov, N., Pavlyk, O., Brodtkorb, P. A., Lee, P., McGibbon, R. T., Feldbauer, R., Lewis, S., Tygier, S., Sievert, S., Vigna, S., Peterson, S., More, S., Pudlik, T., Oshima, T., Pingel, T. J., Robitaille, T. P., Spura, T., Jones, T. R., Cera, T., Leslie, T., Zito, T., Krauss, T., Upadhyay, U., Halchenko, Y. O. & Vázquez-Baeza, Y. (2020). *Nat. Methods*, **17**, 261–272.
- Weber, H., Marx, D. & Binder, K. (1995). *Phys. Rev. B*, **51**, 14636–14651.