

CHARACTERISING 3D SPHERICAL PACKINGS BY PRINCIPAL COMPONENT ANALYSIS

Tingting Zhao¹ Y. T. Feng^{1,2,*} Yuanqiang Tan²

¹ Zienkiewicz Centre for Computational Engineering, Swansea University, UK

² The Institute of Manufacturing Engineering, Huaqiao University, China

Abstract

A novel packing characterising system based on principal component analysis (PCA) is developed in our previous work [*Computer Methods in Applied Mechanics and Engineering* 340: 70-89, 2018] to provide a quantitative approach that can reveal some fundamental features of particle packings. The current work extends this methodology to 3D cases. Gaussian Quadrature is adopted to obtain the volume matrix representation of a particle packing. Then the digitalised image of the packing is obtained by converting cross-sectional images along one direction to column vectors of the packing image. Both a principal variance function and a dissimilarity coefficient are proposed to characterise differences between different packings (or images). Comprehensive investigations for several sets of purposefully generated particle packings are conducted to understand relationships of their principal variances with packing features. Differences between two packings with different packing features can be revealed by the principal variances (PV) and dissimilarity coefficient (DC). Furthermore, the values of PV and DC can indicate different levels of effects on packing caused by configuration randomness, particle distribution, packing density and particle size distribution. The uniformity and isotropy of a packing can also be investigated by this PCA based approach.

KEYWORDS: 3D particle packing, digitalised image, principal component analysis, principal variance, spatial homogeneity and isotropy

1 Introduction

Particulate systems have attracted significant attention over the last several decades due to their wide existence in nature, ranging from nano to cosmic scales, and may involve either hard particles (such as sand) or soft ones (such emulsions and foams) [1, 2, 3, 4]. The particle packing plays an important role in leading the physical behaviour of a particle system. Therefore the spatial-statistical analysis of the geometrical structure of the system is of great scientific and engineering interests. As the topology of the system is highly complex, it is difficult to observe the way particles packed around each other by experiments [5]. With the development of various particle based methods, such as Molecular Dynamics (MD) and the Discrete Element Method [6, 7, 8], more detailed information on the internal structure of particulate systems can be accessed.

Currently, the techniques applied to investigate the features of particle packings focus on the packing density, orientations of the particle contacts, and internal (topological) structures of

*corresponding author; e-mail: y.feng@swansea.ac.uk

packings. Conventional means to characterise a particle packing are limited especially for loose packings without obvious contact chains. The classical characteristics such as particle size distribution and packing density cannot consider the spatial distribution of a packing. Besides, the method involving the use of a radial function is less effective [9]. When a packing is subject to external loading and generates contacts between particles, more features can be checked such as the coordination number, contact force distribution and the contact orientation distribution [10]. More advanced techniques have been developed to characterise a packing, see for instance [11].

These techniques may be classified into different categories from different perspectives. From the perspective of problem scale, there are macroscopic based (stress, strain and critical void ratio [12]) and microscopic based (coordination number and stress tensor [13]). While from the perspective of the specific information to be extracted from the particle system, the methods can be classified into two categories: One is focus on the void ratio (or the solid fraction) [14, 15] of the packing and the other pays attention on the contact network [16].

A particle packing can be viewed as a specific spatial variation of solid material (assigned a value 1) and void (a value of 0). The packing can be represented as a digitalised image with grey-scale ranging from 0 to 1 on a regular grid. Then the resulting packing image can be analysed by techniques developed in the field of computer graphics and imaging process [17, 18, 19, 20], and the (dis)similarity of multiple packing images may also be compared in a quantitative manner. Principal component analysis (PCA) [21, 22] is one of the most popular linear transform based statistical techniques, and has been extensively used in a wide variety of applications. It has proved to be a powerful tool that is often employed for data analysis in general, and dimension reduction in multi-variance analysis, and pattern recognition in signal and imaging processing in particular.

Following this idea, a novel characterising method is developed in our previous work [23] using the principal component of a packing image as the signature of the packing. Furthermore, a dissimilarity coefficient is defined which can provide a single valued metric to quantitatively compare two packings. This has been made possible by the definition of the principal variance function that maps the ordinal numbers of individual principal variances into a non-dimensional unit interval. The current work extends this novel system from 2D particle packings to 3D cases. Gaussian Quadrature is adopted to obtain the volume matrix of the particle packing. Then the digitalisation of packing is obtained by converting cross-sectional images along one direction into column vectors. The signature of the particle packing can be revealed by the subsequent principal component analysis.

The current PCA method is a microscopic based technique which focuses on the void ratio distribution of a particulate packing. Existing methods that belong to the same category include density profiles, the distribution of void ratio or the correlation functions. For the density profile or the distribution of void sizes, Huang et al.[14] evaluate the homogeneity of the samples by dividing the packing into eight horizontal layers in the vertical direction and four rings in the horizontal direction. The density profiles of the packings are analysed by comparing the layers or the rings directly and the results are inevitably verbose to a certain degree. The density profile is, however, not unique for different packings and is strongly affected by the division resolution, as will be showed later. Alshibli et al.[15] study the spatial variation of the void shear band thickness based on the CT images. Some qualitative conclusions are obtained that a cross-sectional variation in void ratio is detected and the void ratio within the shear band is relatively high. For correlation functions, also called the pair distribution function[24] and radial distribution function[25], the results are usually illustrated by the relation between the pair correlation function and the distance to the

centre of the reference sphere. The corresponding curves have complex shapes and can only be explained qualitatively. The current PCA method exploits the full scale spatial correlation in the void distribution of a packing and utilises the Principal Variance function (PV) and the Dissimilarity Coefficient (DC) to qualitatively compare different packings, or different regions/orientations of a single packing.

The paper is organised as follows. The whole analysis procedure of applying PCA to a spherical particle packing is introduced in Section 2, including the formation of packing image, the subsequent numerical treatment of the image matrix and the resulted characterisation of particle packing revealed by the principal variance function and the dissimilarity coefficient. In Section 3, this packing characterising system is applied to several sets of purposefully generated random packings. Then comprehensive investigations are conducted to illustrate that the principal component analysis can reveal the different effects on packings caused by packing randomness, particle distribution, packing density and particle size as well as the uniformity and isotropy of the particle packing. Concluding remarks are made in Section 4.

2 Principal component analysis

This section is devoted to the full description of the numerical procedure that is involved in principal component analysis of a packing and how such analysis can be applied to characterise particle packings. The main principals and terminologies adopted in the current work is derived from our previous work [23] which has some minor differences with the traditional PCA, as explained in [23].

2.1 Packing digitalisation and formation of packing image

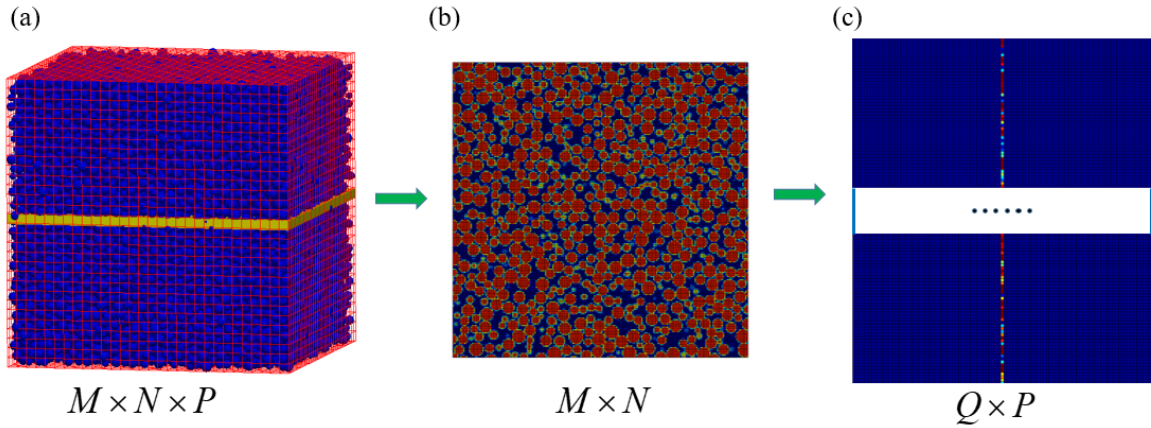


Figure 1: A random particle packing and the corresponding digital representation

First consider a particle assembly $\Omega_p = \bigcup_i \Omega_i$ where Ω_i denotes the domain of the i -th particle, and arbitrarily choose a cuboid solid region \mathcal{V} of dimensions $L_x \times L_y \times L_z$, termed the *analysis region*. The region can be divided into a regular grid of $M \times N \times P$ cubic cells with spacing $h = L_x/M = L_y/N = L_z/P$ (Figure 1(a)). For a cubic cell at (i, j, k) with the volume denoted as \mathcal{V}_{ijk} , compute its average volume covered by particles, or grey-scale as

$$v_{ijk} = \frac{|\Omega_g \cap \mathcal{V}_{ijk}|}{|\mathcal{V}_{ijk}|} \quad (1)$$

where $|\Omega|$ denotes the measure or volume of a domain Ω ; $|\mathcal{V}| = L_x L_y L_z$; and $|\mathcal{V}_{ijk}| = h^3$. An empty cell with no overlapping with any particle has $v_{ijk} = 0$; while a cell fully covered by a particle has $v_{ijk} = 1$. A cell partially covered by particles has $v_{ijk} < 1$. So in general $v_{ijk} \in [0, 1]$. The computation of v_{ijk} is however not straightforward. Different numerical schemes used to evaluate v_{ijk} are discussed in Subsection 2.5.1.

The collection of all the average volumes v_{ijp} at the layer p in the z -direction (highlighted in Figure 1(a)) forms an $M \times N$ matrix, which can be viewed as a digitalised grey-scale representation of the origin 3D packing at the height of $z(p)$ (as shown in Figure 1(b)). The matrix is then re-shaped into a vector simply by concatenating the rows into a single vector with $Q = MN$ elements, and placed as the p -th column vector in a larger $Q \times P$ matrix \mathbf{V}_h . By processing all the P layers in the z -direction, the resulting matrix \mathbf{V}_h can be viewed as a digitalised grey-scale representation of the original packing Ω_g , thus is termed as the packing matrix or image, as shown in Figure 1(3). Despite converting a layer matrix along the z -direction to a column vector, the resulting matrix \mathbf{V}_h can also be created by converting a layer along the x -direction or y -direction to a column vector. The image matrices generated by different directions can reveal the uniformity of the packings as shown in Section 3.3.1.

2.2 Formulations and numerical procedures

The mean value of the packing matrix \mathbf{V}_h , i.e. the packing density of the region \mathcal{V} , can be computed

$$\rho_v \equiv \frac{|\Omega_g \cap \mathcal{V}|}{|\mathcal{V}|} = \frac{1}{QP} \sum_{i=1}^Q \sum_{j=1}^P v_{ij} \quad (2)$$

Let $v(\mathbf{x})$ be the material distribution function with v taking the value of 1 for a point located within a particle, and 0 otherwise. It is not difficult to derive that the total variance of a packing in the region \mathcal{V} is related to the packing density by

$$\sigma_v = \frac{1}{|\Omega|} \int_{\Omega} (v - \rho_v)^2 d\Omega = \rho_v (1 - \rho_v) \quad (3)$$

The total variance of the matrix is defined as

$$\sigma_h = \frac{1}{QP} \sum_{i=1}^Q \sum_{j=1}^P (v_{ij} - \rho_v)^2 \leq \sigma_v \quad (4)$$

i.e. the total variance of the packing σ_v is the upper bound of any packing image.

Let q_j be the mean value of the j -th column of the packing matrix \mathbf{V}_h

$$q_j = \frac{1}{Q} \sum_{i=1}^Q v_{ij} \quad (5)$$

By subtraction of its mean from each column vector of \mathbf{V}_h , the column centralised matrix $\bar{\mathbf{V}}_P$ of \mathbf{V}_h is obtained as:

$$\bar{\mathbf{V}}_P = \mathbf{V}_h - \mathbf{e}_Q \mathbf{q}_P \quad (6)$$

where \mathbf{e}_Q is an $Q \times 1$ column vector with all its elements being 1's; and \mathbf{q}_P is the $1 \times P$ mean value vector $\mathbf{q}_P = \{q_j\}$.

Define the covariance matrix of $\bar{\mathbf{V}}_P$ as

$$\mathbf{S}_P = \frac{1}{Q} \bar{\mathbf{V}}_P^T \bar{\mathbf{V}}_P \quad (7)$$

where \mathbf{S}_P is a $P \times P$ square matrix. Notice in the above that Q instead of $Q - 1$ is used. Further define the column-wise total variance as

$$\sigma_P^c = \frac{1}{P} \text{Tr}(\mathbf{S}_P) = \frac{1}{P} \sum_{i=1}^P (\mathbf{S}_P)_{ii} \quad (8)$$

which may be (slightly) different from the total variance σ_h in general.

By solving the eigenvalue problem of \mathbf{S}_P , it yields the following matrix decomposition

$$\mathbf{S}_P \mathbf{U}_P = \mathbf{U}_P \mathbf{D}_P \quad (9)$$

with

$$\mathbf{D}_P = \mathbf{U}_P^T \mathbf{S}_P \mathbf{U}_P, \quad \mathbf{U}_P^T \mathbf{U}_P = \mathbf{I}_P$$

where the diagonal matrix $\mathbf{D}_P = \text{diag}\{d_i\}$ contains all the eigenvalues d_i in descending order, which are termed the *principal variances* (PVs); and $\mathbf{U}_P = \{\mathbf{u}_i\}$ are the orthonormal vectors, termed the *principal modes*. As $\bar{\mathbf{V}}_P$ is column centralised, \mathbf{S}_P is a semi-positive definite matrix with at least one zero principal variance. It is also well known that the sum of the PVs and the total column-wise variance is related by

$$\frac{1}{P} \sum_{i=1}^P d_i = \sigma_P^c \quad (10)$$

\mathbf{S}_P can be recovered from the principal variances and modes as

$$\mathbf{S}_P = \mathbf{U}_P \mathbf{D}_P \mathbf{U}_P^T = \sum_{i=1}^{P-1} d_i \mathbf{u}_i \mathbf{u}_i^T \quad (11)$$

The column-wise total variance σ_P^c , the mean value vector \mathbf{q}_P , the principal variance matrix \mathbf{D}_P and the corresponding modes \mathbf{U}_P form a unique set \mathcal{C}_P , termed the column-wise *characteristic set*, that fully determines the packing image.

$$\mathcal{C}_P = \{\sigma_P^c, \mathbf{q}_P, \mathbf{D}_P, \mathbf{U}_P\} \quad (12)$$

As the PVs and the column-wise total variance is related by (10), and the total variance (and also the column-wise variance) is related to the density (see (3)), the PVs play a dominant role to characterise a packing image and therefore can be viewed as the *signature* of the packing.

2.3 Principal variance function

To facilitate the comparison between different sets of principal variances, particularly when they are obtained from different resolutions P , the ordinal number i of a principal variance d_i is mapped from 1 to P to a non-dimensional 'position' variable $x \in [0, 1]$ by

$$x(i) = \frac{1}{P} \left(i - \frac{1}{2} \right) \quad (13)$$

Then a continuous function $\mathbf{d}(x)$, $x \in [0, 1]$, termed the *principal variance function*, can be constructed to interpolate the discrete variances d_i using piecewise linear or higher order interpolation functions such that

$$\mathbf{d}(x_i) = d_i, \quad x_i = x(i), \quad i = 1, \dots, P \quad (14)$$

2.4 Packing image similarity

Consider two packing images with their principal variance functions $d_1(x)$ and $d_2(x)$ obtained, define a so-called *dissimilarity coefficient* (DC) between these two images as

$$\mathcal{D}_c = \left[\frac{1}{\Sigma_1 + \Sigma_2} \int_0^1 [d_1(x) - d_2(x)]^2 dx \right]^{1/2} \in [0, 1] \quad (15)$$

where Σ_1 and Σ_2 are defined as

$$\Sigma_i = \int_0^1 d_i^2(x) dx \quad (i = 1, 2)$$

Consequently, the degree of similarity of these two packing images can be quantified by the *similarity index* $\in [0, 100]$ defined as

$$\mathcal{S}_{\mathcal{I}} = (1 - \mathcal{D}_c) \times 100 \quad (16)$$

2.5 Additional Issues

This subsection discusses some additional issues related to the principal component analysis of 3D cases.

2.5.1 Different Schemes to compute the volume average of each cubic cell

The most time consuming part of the above proposed procedure is the evaluation of the density, or the particle covered volume of a cubic cell (1). In 2D cases, the covered area of a square cell by circular particle(s) may be computed analytically depending on their overlapping situations which can be classified by the number of grid nodes contained in the circle(s), while it is difficult to follow a similar procedure to obtain the exact overlap volume between spherical particle(s) and a cubic cell due to the complexity of the contact situations in 3D cases. Two different computational schemes are therefore adopted to numerically compute the volume average of each cubic cell.

The first scheme, termed the *area-based*, is derived from the method used in the 2D case which can determine the overlap area between a circle and a square cell. The 3D problem can be reduced to the 2D one by integrating the overlap area in the x - y plane along the z -direction. Gaussian Quadrature is implemented to numerically evaluate the integration. The volume average of a cubic cell of side h and centred in a local coordinate system is obtained by

$$v_{ijk} = \int_{-h/2}^{h/2} a_{ij}(z) dz = \sum_{l=1}^n w_l a_{ij}(z_l) \quad (17)$$

where $a_{ij}(z)$ is the average area covered by circles at z , and z_l and w_l are the position and corresponding weight of a Gaussian point.

The second scheme, termed the *point-based*, simplifies the problem by checking whether a point is inside a particle or not. A certain number of points are selected in each cubic cell. Two values 0 (point outside the particle) or 1 (inside the particle) can be assigned to each point. Gaussian Quadrature is again applied in this scheme to select the position and weight of each point in the cubic cell. The volume average of each cubic cell is the integration of the point values. Using n Gaussian points in each directions gives rise to the formula

$$v_{ijk} = \sum_{l=1}^n \sum_{s=1}^n \sum_{t=1}^n w_{lst} v(x_l, y_s, z_t) \quad (18)$$

Table 1: The overall packing density of three regular packings computed by different computational schemes

N	R1			R8			R64		
	1-Point	3-Point	Area	1-Point	3-Point	Area	1-Point	3-Point	Area
25	0.5259	0.5231	0.5236	0.5212	0.5230	0.5232	0.5212	0.5232	0.5235
50	0.5260	0.5238	0.5236	0.5259	0.5231	0.5236	0.5212	0.5230	0.5232
100	0.5240	0.5236	0.5236	0.5260	0.5238	0.5236	0.5259	0.5231	0.5236

where w_{lst} , (x_l, y_s, z_t) and $v(x_l, y_s, z_t)$ are the weight, coordinates and material distribution value of each Gaussian point.

Several numerical tests are conducted to compare the above two schemes. Three regular packings are generated in a unit cubic region where there are respectively 1, 8 and 64 equal-sized particles in the packings that are labelled as R1, R8 and R64 respectively. It is obviously that the three packings have the same overall packing density $\rho = 3/4\pi/8 = 0.5236$. The region is then divided into a regular grid of $N \times N \times N$ cubic cells, and three resolutions $N = 25, 50, 100$ are considered. The packing density can also be computed by (2). In the first scheme (*area-based*), the number of sections along the z -direction in each cubic for integrating is 3 which is able to obtain sufficiently accurate results. In the second scheme, two different numbers of Gaussian integration points $n = 1$ and 3 are tested. Table 1 provides the overall packing densities of the three regular packings for different grid resolutions computed by the two computational schemes, while the first principal variances of the three packings with different resolutions are listed in Table 2. These results indicate that the two computational schemes can have a similar accuracy when the number of Gaussian points used is reasonably large.

2.5.2 Properties of repetitive packings

In our previous work [23] for 2D cases, the relationship between packings with repetitive, periodic and symmetric natures has been established, and it is theoretically derived that for a m -repetition packing, $1/m$ of the principal variances will be m times those of the basic structure. This conclusion is also valid for 3D cases. Take the previous three regular packings as an example: Packing R1 can be regarded as the basic structure, R8 is a 2-repetition packing of R1, and R64 is a 2-repetition packing of R8 and a 4-repetition packing of R1. From the first principal variances of the three regular packings R1, R8 and R64 given in Table 2, it is evident that, regardless of the integration scheme used, the first PV of R64 with $N=100$ is about twice of that of R8 with $N=50$ which itself is also about twice of that of R1 with $N=25$. Note that different resolutions N are used for different packings to ensure that a constant relative resolution is maintained for the comparison. It can also be observed that the accuracy of the computational scheme used does not influence the relative relationship of two packings.

2.5.3 Analysis Region and Formation of Packing Matrix

The analysis region of a packing is not necessarily a cuboid as is assumed earlier. In fact the analysis region can be, in general, a prism which has the same cross-section along one direction (defined as the z -direction). Then each section image can be reshaped to a vector

Table 2: The first principal variances of three regular packings computed by different computational schemes

N	R1			R8			R64		
	1-Point	3-Point	Area	1-Point	3-Point	Area	1-Point	3-Point	Area
25	11.715	11.623	11.587	12.043	11.187	11.140	12.043	9.817	9.781
50	23.291	23.434	23.450	23.401	23.217	23.146	24.056	22.347	22.253
100	47.026	47.053	47.054	46.567	46.853	46.885	46.788	46.420	46.279

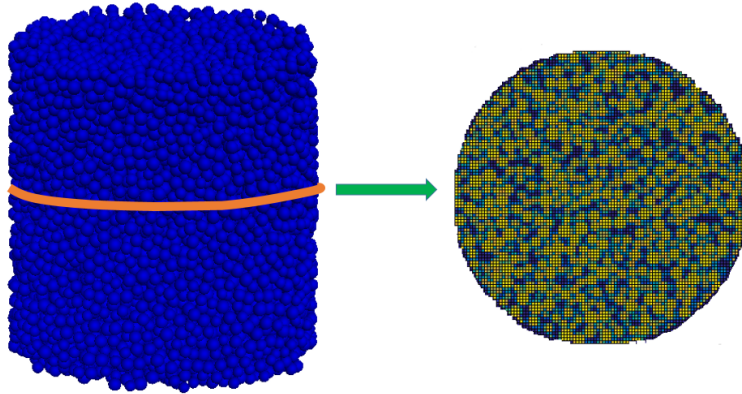


Figure 2: A cylindrical analysis region with a circular cross section image

with the same number of elements which forms one column of the final analysis matrix \mathbf{V}_h . For instance, the analysis region could be a cylinder and the cross section along the z -direction is a disc as shown in Figure 2. Only the elements in the circular area are selected to construct the column vector of the packing image matrix.

Alternative to representing each section image as a column vector in the final image matrix, all the pixels with the same (x, y) coordinates along the z -direction can instead be represented by a column vector, resulting in an different packing matrix which is nothing but \mathbf{V}_h^T , the transpose of the original image matrix \mathbf{V}_h . \mathbf{V}_h and \mathbf{V}_h^T may have different principal variances in general, but both can be used to characterise particle packings. In the following discussions, \mathbf{V}_h is assumed.

2.5.4 Evaluation of main principal variances

The computational cost associated with the solving eigen-value problem (9) may be high particularly when the resolution P (in the z -direction) is large. One option to reduce this cost is to evaluate only a specified number of main principal variances, as will be demonstrated below that the principal variances with largest values play a more dominant role in characterising packings and qualitatively comparing different packings using the (dis)similarity coefficient/index defined earlier.

Note from the definition of the principal variance function that compares images with different resolutions (in the z -direction), the same proportion of the selected number of principal variances over the resolutions should be specified.

Table 3: Properties of random packings: U-set and G-set

Set	Uniform Distribution (U-set)			Gaussian Distribution (G-set)		
	U1	U2	UL2	G1	G2	GL2
Group Name	U1	U2	UL2	G1	G2	GL2
Mean Density	0.6038	0.6062	0.5664	0.6026	0.6053	0.5673
Particle No.	16804	2094	1939	14770	1841	1701
r_{min}	0.015	0.030	0.030	0.015	0.030	0.030
r_{max}	0.025	0.050	0.050	0.025	0.050	0.050
r	0.020	0.040	0.040	0.020	0.040	0.040

Table 4: Packing Properties of the L-set

Layer	1	2	3	4	5	6
r_{min}	0.010	0.020	0.030	0.040	0.050	0.060
r_{max}	0.015	0.030	0.045	0.060	0.075	0.090
r	0.0125	0.025	0.0375	0.050	0.0625	0.075

3 Packing Characterisation

This section is devoted to illustrating how the principal variance function and dissimilarity coefficient defined in the previous section can be applied to quantitatively characterise the features of different packings. As the same relative relation of two packings can be obtained by different volume average computational schemes, only the results obtained by the *area-based* integration scheme are shown below. Several packing samples to be used are given in the next subsection, while a number of characterisation issues will be discussed in later subsections.

3.1 Numerical samples

Two sets of random particle packings are generated within a unit cubic domain $[-0.5, -0.5, -0.5] \times [0.5, 0.5, 0.5]$ with the periodic condition applied to all three directions. The first set, termed U-set, has particle sizes uniformly distributed within a range; the second set, termed G-set, has particle sizes obeying a Gaussian distribution with limited minimum and maximum values. Each set has three groups of packing with different particle sizes or densities, while within each group, 10 random packing samples with the same packing properties are generated. The packing properties used are listed in Table 3, and one generated packing sample for each group is illustrated in Figure 3. In the U-set, groups U1 and U2 have the same density of 0.6 but the particle size distribution range of U2 is double of U1; groups U2 and UL2 have the same particle size range but different densities with 0.60 for U2 and 0.57 for UL2. The same goes for the groups in the G-set.

The third set, or L-set, is a group of layered particle packings randomly generated within the same domain. Each packing is divided into six layers along the z -direction, having particle sizes obeying the uniform distribution within each layer. The particle size distribution ranges of the layers form an arithmetic sequence. The packing density is 0.6 which is the same as those of U1, U2, G1 and G2. The packing properties are given in Table 4 and one packing sample is displaced in Figure 4.

The principal component analysis introduced in Section 2 is applied to each sample of the

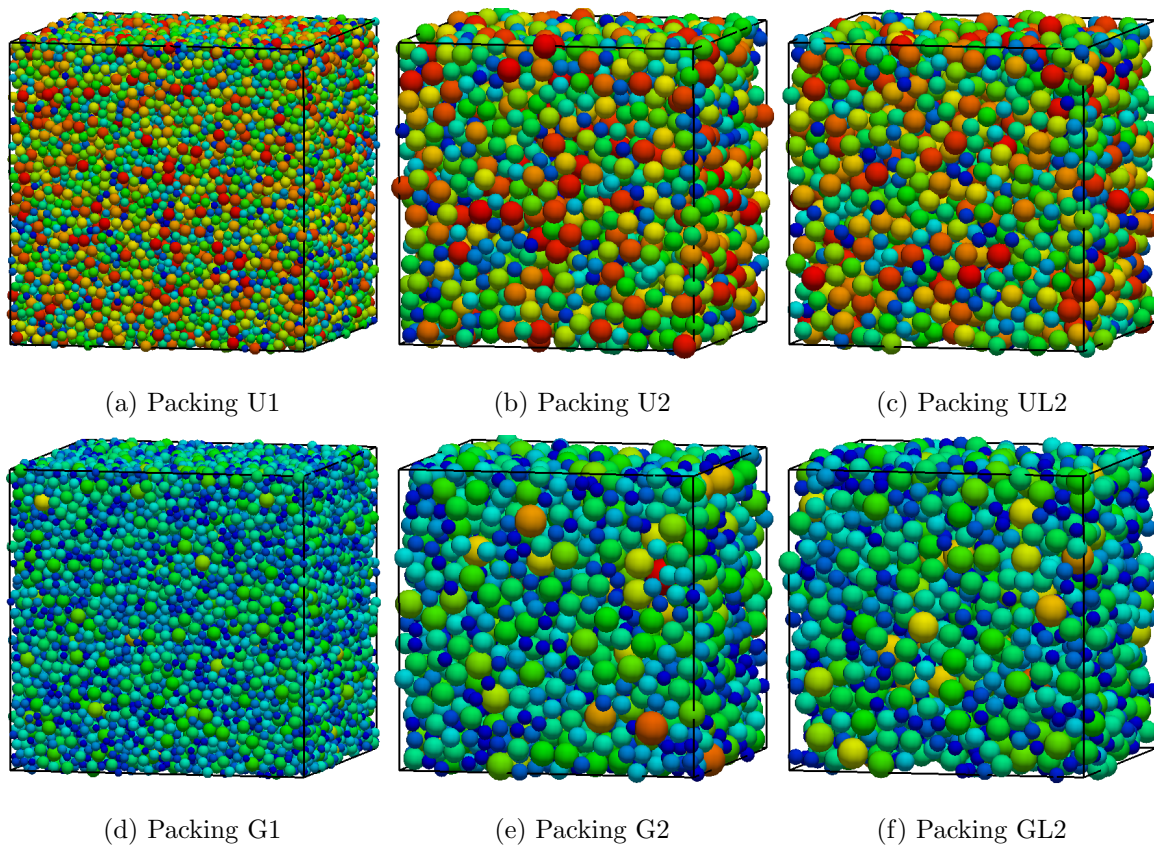


Figure 3: Sample Packings of both U-set and G-set

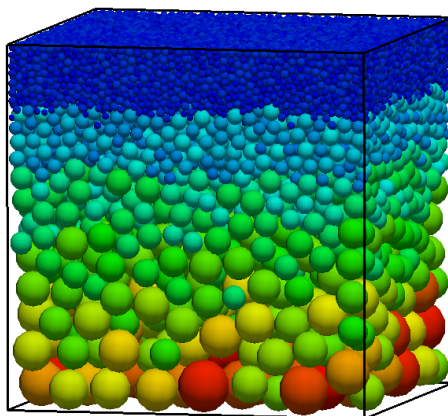


Figure 4: A sample packing of the L-set

U-set, G-set and L-set.

3.2 Different effects on packing samples

3.2.1 Packing randomness effects

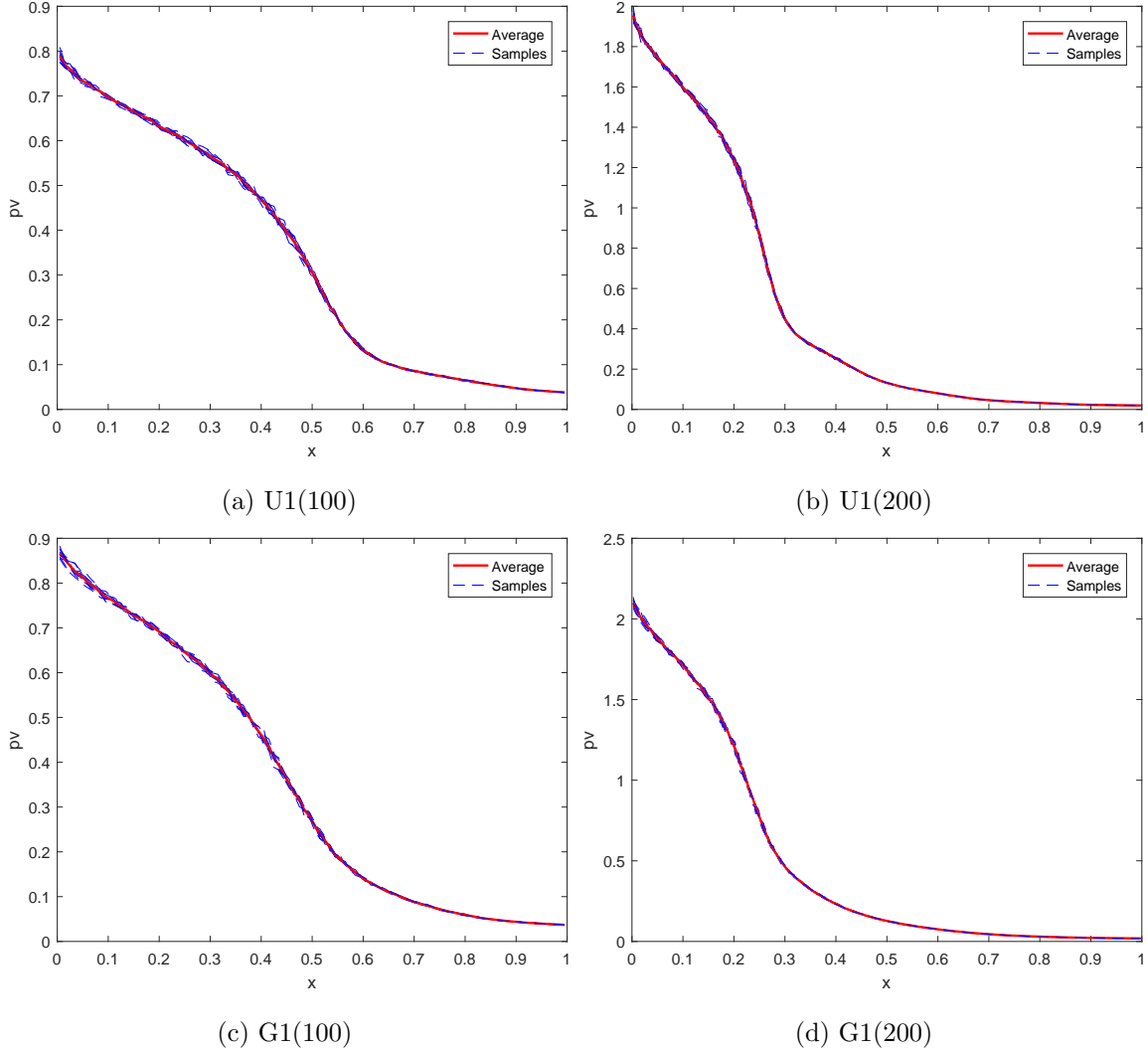
For each group of U1, U2, G1 and G2, 10 random packing samples are generated with the same packing properties. It is obvious that the packing configurations of these samples in each group are statistically different, but the difference can hardly be evaluated by existing

Table 5: Average dissimilarity coefficients of each group in U-set and G-set

P	U1	U2	G1	G2
100	0.0064	0.0119	0.0070	0.0123
200	0.0061	0.0114	0.0062	0.0117

conventional methodologies. As will be shown below, the currently proposed novel PCA based approach has the capability of measuring the effects caused by the randomness of packing configurations.

For each packing group, the PVs of all 10 samples are computed and their averages are taken to be the PVs of the group. The principal variance functions of the 10 samples for groups U1 and G1 at two grid resolutions $M(= N = P) = 100, 200$ are displayed in Figure 5. Clearly the PV functions of the 10 samples in each group are located within a narrow band around the group mean value. This indicates that these randomly generated samples from the same particle size distribution indeed have very similar statistical features.

Figure 5: Average principal variance functions of groups U1 and G1 for $N=100, 200$

To quantify the difference, the dissimilarity coefficient of 10 samples in each group are cal-

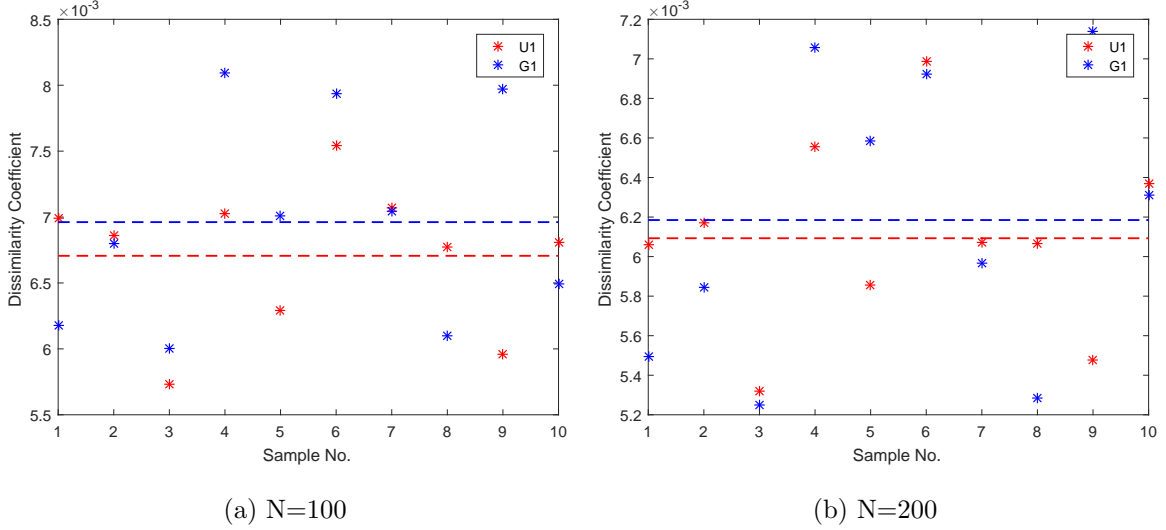


Figure 6: Dissimilarity coefficients of groups U1 and G1 for $N=100$, 200

culated based on the formula (15) against their group average for two resolutions: 100 and 200. The dissimilarity coefficients of groups U1 and G1 are displayed in Figure 6 and the group average of the 10 samples are provided in Table 5. As Figure 5 visually shows a minor difference among PVs functions of samples in each group, which is confirmed by very small dissimilarity coefficients in Table 5, it can be concluded that the effects of the particle distribution randomness is indeed insignificant. It has also been found that the difference decreases when the number of particles in the packing increases.

3.2.2 Particle distribution effects

The effects of different particle distributions on packings can be illustrated by the comparison between the corresponding groups in both U-set and G-set. For each packing group, the average PVs are taken to be the PVs of the group. The principal variance functions of groups U1, U2, G1 and G2 at two grid resolutions $P = 100$ and 200 are displayed in Figure 7. To quantify the difference caused by different particle distributions, the dissimilarity coefficients of U1-G1 and U2-G2 are provided in Table 6. The dissimilarity between the corresponding groups of the two distributions are larger than that of the samples within each group. In other words, the particle distribution has a more significant influence on the packing features than the packing configuration randomness.

The density profiles of the U1 packing for different resolutions are also presented in Figure 8, clearly indicating that the profile is strongly dependent of the resolution N , while the general shapes of the PV distribution remain very similar for different N , as shown in Figure 7.

Table 6: Dissimilarity coefficients between different groups

P	Particle distribution		Packing density		Particle size	
	U1-G1	U2-G2	U2-UL2	G2-GL2	U1-U2	G1-G2
100	0.0592	0.0467	0.1397	0.1392	0.4798	0.4664
200	0.0463	0.0423	0.1353	0.1346	0.4111	0.4041

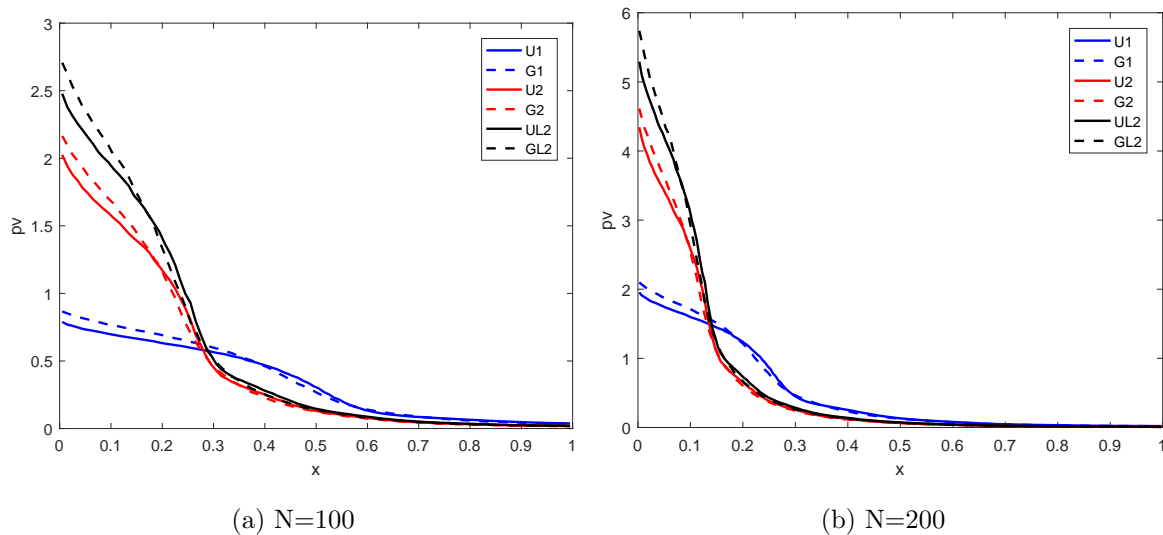


Figure 7: Average principal variance functions of different groups

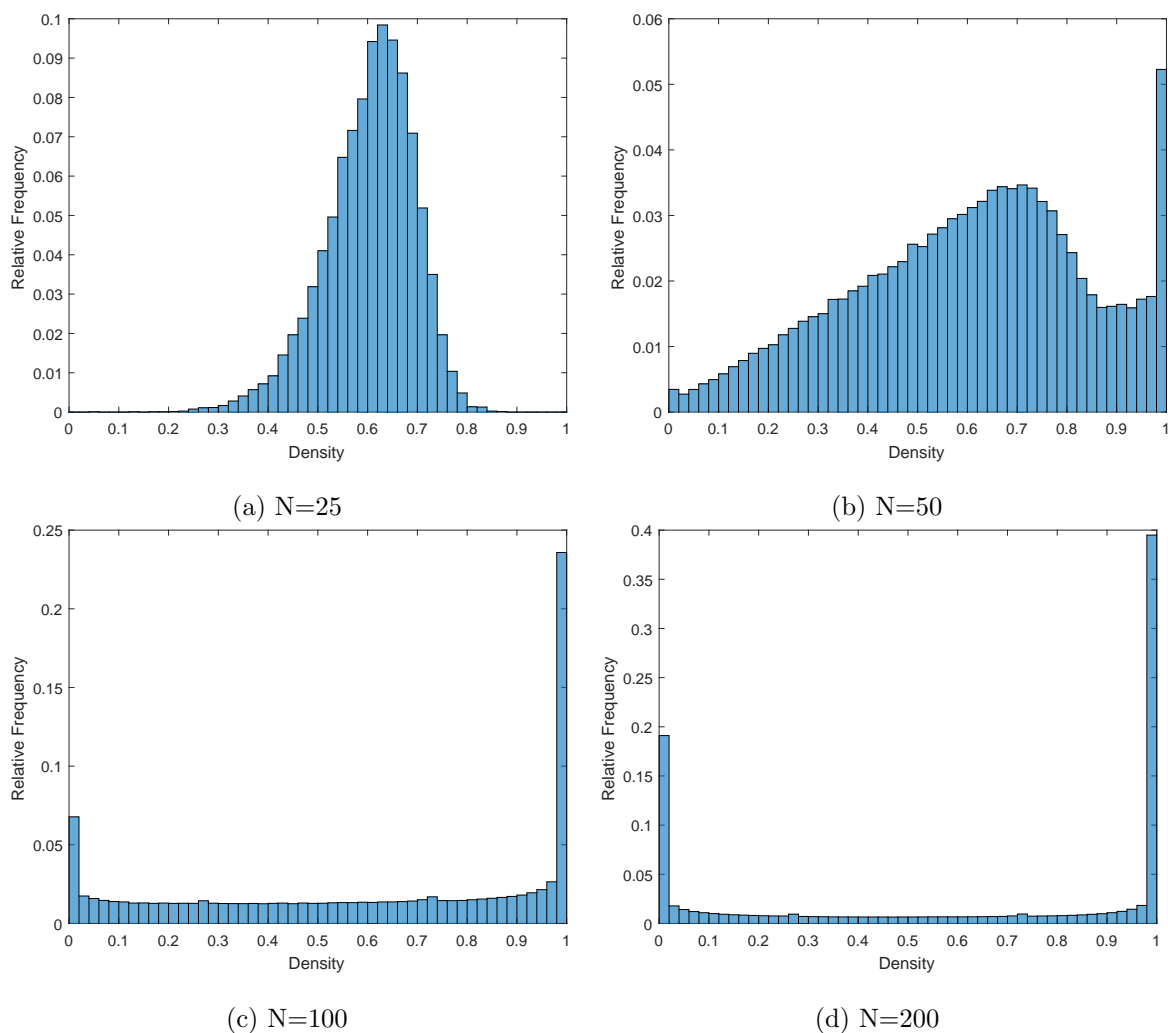


Figure 8: Density profiles of the U1 packing with different resolutions

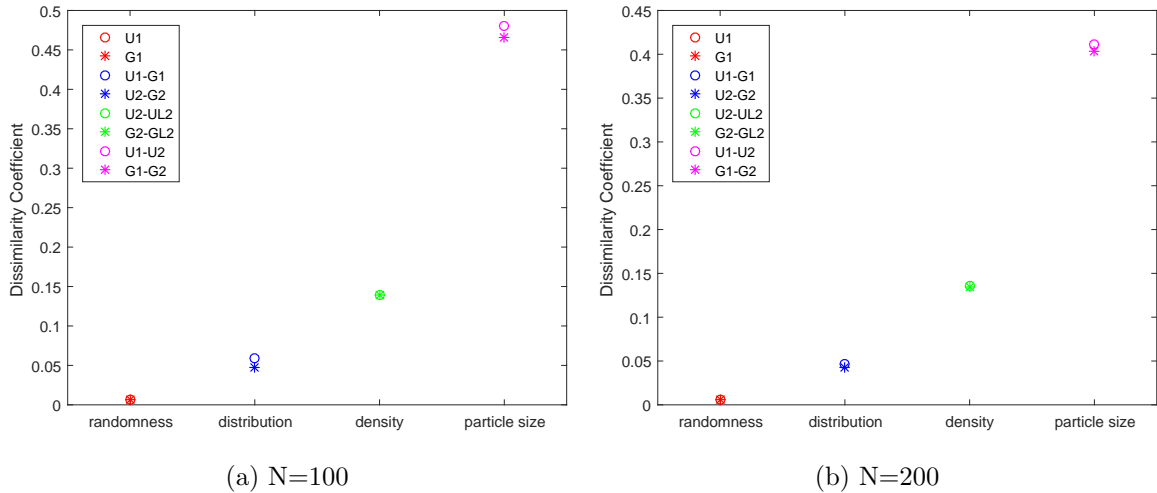


Figure 9: Dissimilarity coefficients of different samples and groups

3.2.3 Packing density effects

The effects of different packing densities are observed by comparing U2 against UL2, and G2 against GL2. The PVs of UL2 and GL2 are displayed in Figure 7 and the corresponding dissimilarity coefficients are listed in Table 6. Clearly, the dissimilarity caused by the packing density is larger than that of the particle distribution.

3.2.4 Particle size effects

Finally, the effects of particle sizes with the same distribution nature (uniform or Gaussian, for instance) are investigated by comparing U1 against U2, and G1 against G2. The dissimilarity coefficients of U1-U2 and G1-G2 are also listed in Table 6 which are significantly larger than the others.

3.2.5 Summary

To visualise the effects caused by different packing properties, all DCs are plotted in Figure 9. It can be concluded that the randomness of the packings has a minor influence on the packing features compared to the other three factors. The influence extent of both packing density and particle size is related to the value range of each factor. To compare these two effect factors, more simulation results are conducted. The base sample has a density of 0.5 and a particle size range [0.025,0.035]. Both the packing density and the particle size have a variation of 10%. The dissimilarity between the base sample and the other samples are listed in Table 7. It can be seen that packing density has the most significant influence.

3.3 Packing uniformity and isotropy

Besides having capability of comparing the features of two particle packings, this PCA based approach can also be employed to exploit the properties of a packing in more detail, such as its spatial uniformity and isotropy.

Table 7: Effect extent of density and particle size

Density	0.51	0.52	0.53	0.54	0.55
DC	0.0318	0.0666	0.1007	0.1219	0.1623
Particle size	[0.0255,0.0357]	[0.0260,0.0364]	[0.0265,0.0371]	[0.0270,0.0378]	[0.0275,0.0385]
DC	0.0258	0.0381	0.0516	0.0672	0.0847

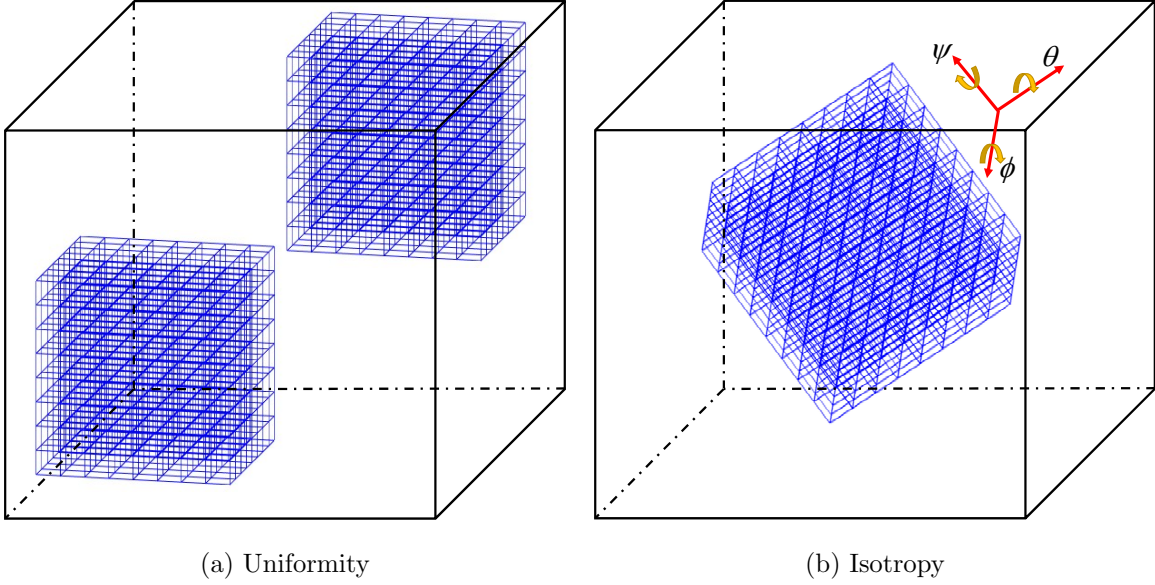


Figure 10: Uniformity and isotropy checks of a packing

3.3.1 Uniformity

For one packing, its uniformity in space can be checked by applying a moving analysis region and obtaining the PVs at some selected locations (as shown in Figure 10), and then compute the DCs of these PVs by (15). A uniformly small DC indicates that the packing may be statistically homogeneous. Alternatively, if the volume matrix \mathbf{V}_h of a packing within a large analysis region is obtained, by selecting smaller analysis regions with the same grid spacing h , or by simply selecting sub-matrices from \mathbf{V}_h , the spatial uniformity of the packing can also be investigated at a smaller scale.

To illustrate this approach, U1(200) and L(200) are used as two examples. 10 of $[100 \times 100 \times 100]$ sub-matrix blocks are randomly extracted from the two packing volume matrices and principal component analysis is applied to every sub-matrix to obtain its PVs. Figure 11(a) shows PVs of the sub-matrices and the averages of different packings. Clearly the PV functions of 10 sub-matrices of Set U1 are located within a narrow band around the mean value while those of Set L show notable variations. The dissimilarity coefficients of the sub-matrices against the averaged value are depicted in Figure 11(b). A small level of spatial in-homogeneity exists in Packing U1 while a large degree of in-homogeneity is observed in Packing L, as expected.

As mentioned in Section 2.1, the image matrices generated along different directions can also be used to check the uniformity of packings. Two image matrices by converting the matrix in a XY layer or a YZ layer into a column vector are created and the corresponding PVs are

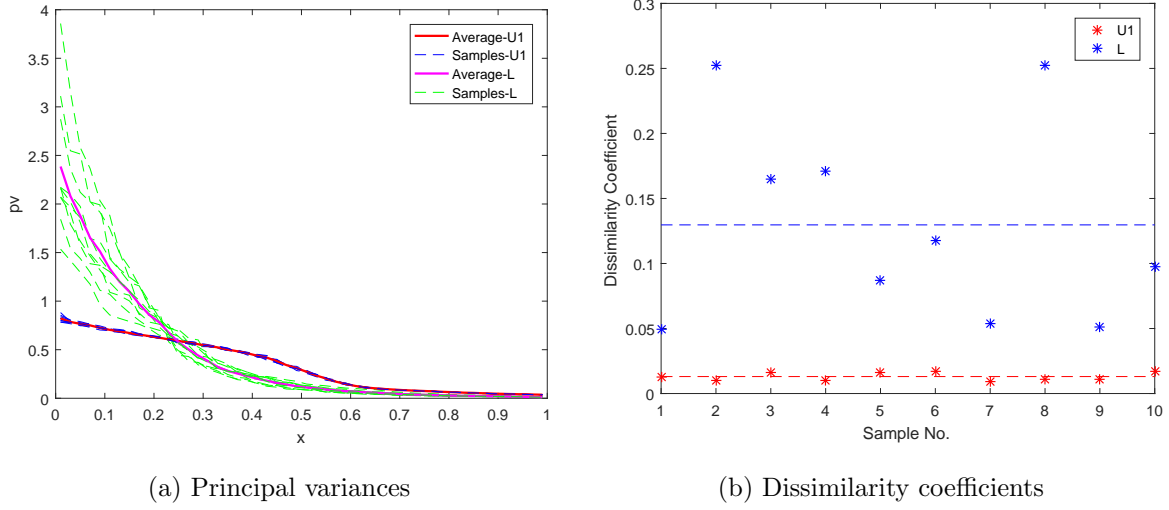


Figure 11: Packing uniformity checkings for U1 and L

shown in Figure 12. It can be seen clearly that the two PV functions of Set U1 are almost the same while the two PVs of Set L show apparent difference. It is easily concluded that a large degree of in-homogeneity exists in Packing L especially along the z -direction.

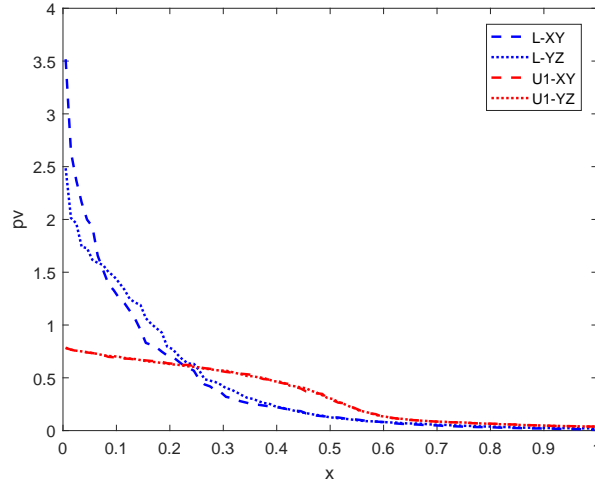


Figure 12: Principal variances of image matrices generated along different directions

3.3.2 Isotropy

The isotropy of a packing can be examined by rotating an analysis region in three directions, and applying principal component analysis to each analysis region, as displayed in Figure 10(b). Then the DCs of the PVs of these rotated analysis regions against their averaged PVs will reveal if the given packing is generally isotropic or not.

For illustrative purpose, this approach is applied to Packings U1 and L. The analysis region is dimensioned as $[0.5 \times 0.5 \times 0.5]$ with the resolution $[50 \times 50 \times 50]$. Seven values of rotate angles θ, ϕ, ψ are chosen from 0° to 180° as $[0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ, 180^\circ]$. Then the total number of rotated analysis regions is 343. Figure 13 shows PVs of the rotated regions and the averages of Packings U1 and G. Obviously, the PV function of Packing U1 is smaller and stabler than that of Packing L. The dissimilarity coefficients of the rotated analysis regions

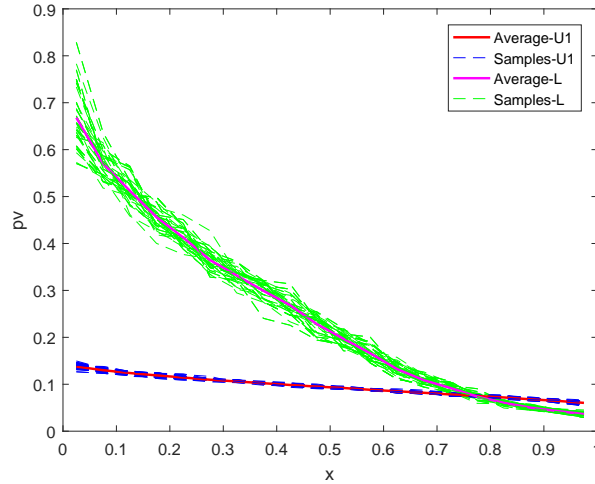
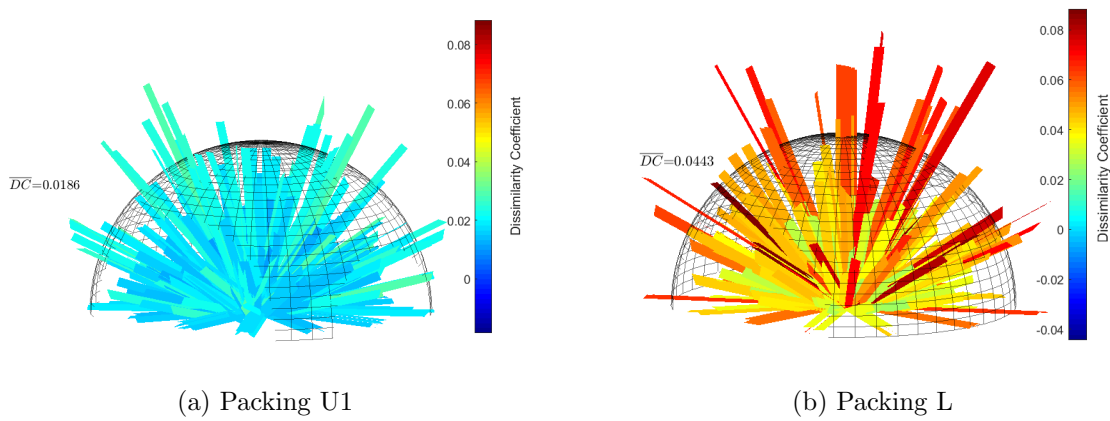


Figure 13: Principal variances for isotropic checking



(a) Packing U1

(b) Packing L

Figure 14: Dissimilarity coefficient for isotropic checking

against the averaged value are plotted in Figure 14. The average DC of Packing U1 is 0.0186 but that of Packing L is 0.0443, which indicates that a weak isotropy exists in Packing U1 while a notable an-isotropy exists in Packing L.

4 Conclusion

The present work has extended the methodology which characterises the particle packing based on Principal Component Analysis to 3D cases. Gaussian Quadrature is adopted to obtain the volume matrix of a particle packing. Then the digitalised image of the packing is obtained by converting the image matrix along the z -direction into a column vector of the image matrix of the packing. Both principal variance function and dissimilarity coefficient are proposed to compare different packing images.

This approach has been applied to several purposefully generated packings. Numerical investigations have shown that differences between two packings with different features can be characterised by the principal variance and dissimilarity coefficient. The values of PVs and DC can indicate different levels of effects on packing caused by configuration randomness, particle distribution, packing density and particle size. The uniformity and isotropy of a

packing can also be investigated by this PCA based approach. Although only spherical particle packings are considered, other non-spherical packings can equally be treated within the proposed characterisation methodology.

In addition, as noted in [23] the packing characterisation through the illustrative examples has been mainly focused on the quantitative comparison between different packings or different regions/orientations of a single packing using their principal variances or the dissimilarity coefficient. More work is being pursued to directly relate PVs and other values in the characteristic sets to packing features of a packing.

Acknowledgment

The work was partially supported by the National Natural Science Foundation of China (Project No. 11772135) and the special fund from the Institute of Manufacturing Engineering of Huaqiao University, China. This support is greatly acknowledged.

References

- [1] C.H. Liu, S.R. Nagel, D.A.Schecter and S.N Coppersmith. Force fluctuations in bead packs. *Science*. 269(5223): 513, 1995.
- [2] J. Bruji, S.F. Edwards, I. Hopkinson and H.A. Makse. Measuring the distribution of interdroplet forces in a compressed emulsion system. *Physica A: Statistical Mechanics and its Applications*. 327(3): 201-212, 2003.
- [3] M.E. Cates, J.P. Wittmer, J.P. Bouchaud and P. Claudin, Jamming, force chains, and fragile matter. *Physical review letters*. 81(9): 1841, 1998.
- [4] M. Kramr, A. Goulet, L. Kondic and K. Mischaikow Evolution of force networks in dense particulate media. *Physical Review E*.90(5):052203, 2014.
- [5] T.S. Majmudar and R.P. Behringer Contact force measurements and stress-induced anisotropy in granular materials. *Nature*. 435(7045): 1079-1082, 2005.
- [6] P. A. Cundall, and O. D. L. Strack. A discrete numerical model for granular assemblies. *Geotechnique*, 29(1):47-65, 1979.
- [7] Y. T. Feng, K. Han, D. R. J. Owen. Filling domains with disks: An advancing front approach. *International Journal For Numerical Methods in Engineering*. 56: 699-713, 2003.
- [8] K. Han, Y. T. Feng, D. R. J. Owen. Sphere packing with a geometric based compression algorithm. *Powder Technology*, 155(1):33-41, 2005.
- [9] K. Lochmann, L. Oger and D. Stoyan. Statistical analysis of random sphere packings with variable radius distribution. *Solid State Sciences*, 8(12):1397-1413, 2006.
- [10] B. Cambou , P. Dubujet and C. Noguier-Lehon Anisotropy in granular materials at different scales. *Mechanics of materials*, 36(12):1185-1194, 2004.
- [11] B. A. Klumov, S. A. Khrapak, and G. E. Morfill. Structural properties of dense hard sphere packings. *Phys. Rev. B*, 83, 184105, 2011.

- [12] Li, X.S. and Dafalias, Y.F. Dilatancy for cohesionless soils. *Geotechnique*. 50(4): 449-460, 2000.
- [13] W. Zhou, J. Liu, G. Ma, and X. Chang. Three-dimensional DEM investigation of critical state and dilatancy behaviors of granular materials. *Acta Geotechnica*. 12(3): 527-540, 2017.
- [14] X. Huang, K.J. Hanley, C. O'Sullivan and F.C. Kwok Effect of sample size on the response of DEM samples with a realistic grading. *Particuology*. 15: 107-115, 2014.
- [15] K.A. Alshibli and A. Hasan. Spatial variation of void ratio and shear band thickness in sand using X-ray computed tomography. *Geotechnique*. 58(4):249-257, 2008.
- [16] L. Papadopoulos , M.A. Porter, K.E. Daniels, and D.S. Bassett. Network analysis of particles and grains. *Journal of Complex Networks*. 6(4): 485-565, 2018.
- [17] H. H. Barret. Foundations of Image Science. 3rd Edition, John Wiley & Sons, New Jersey, UK, 2004.
- [18] R. C. Gonzales and R. E. Woods. Digital Image Processing. 2nd Edition, Prentice Hall, 2002.
- [19] S. Haykin. Neural Networks, A Comprehensive Foundation. IEEE Press, USA, 1984.
- [20] M. Petrou and P. Bosddogianni. Image Processing: The Fundamentals. John Wiley & Sons, UK, 2000.
- [21] I. T. Jolliffe. Principal Component Analysis, Springer Series in Statistics, 2nd ed., Springer, NY, 2002
- [22] H. Abdi and L. J. Williams. Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*. 2 (4): 433-459, 2010.
- [23] Y. T. Feng, T. Zhao, M. Wang and D. R. J. Owen. Characterising particle packing by principal component analysis. *Computer Methods in Applied Mechanics and Engineering*. 340: 70-89, 2018.
- [24] N. Cusack The physics of structurally disordered matter: an introduction. *Adam Hilger*. 1987.
- [25] J.M. Ziman Models of disorder: the theoretical physics of homogeneously disordered systems. *CUP Archive*. 1979.