A Hierarchical Approach to Simulate the Packing Density of Particle Mixtures on a Computer

Michael Kolonko · **Steffen Raschdorf** · **Dominic Wäsch**

Abstract In many fields of materials science it is important to know how densely a particle mixture can be packed. The "packing density" is the ratio of the particle volume and the volume of the surrounding container needed for a random close packing of the particles.

We present a method for estimating the packing density for spherical particles based on computer simulations only, i.e. without the need for additional experiments. Our method is particularly suited for particle mixtures with an extremely wide range of particle diameters as they occur e.g. in modern concrete mixtures. A single representative sample from such mixtures would be much larger than can be handled on present standard computers.

In our hierarchical approach the diameter range is therefore divided into smaller intervals. Samples from these limited diameter intervals are drawn and their packing density is estimated from a simulated packing. The results are used to "fill" the interstices in the sample from the next larger particle interval. To account for the interaction between particles of different sizes we include larger particles into the sample of smaller ones. The larger ones act as part of the boundary during the packing. Thus we obtain more realistic estimates of how dense a fraction of particles can be packed within the whole mixture.

The focus of this paper is on the divide-and-conquer approach and on how the simulation results from the fractions can be collected into an overall estimate of the packing density. We do not go into details of the simulation technique for the single packing.

We compare our results to some experimental data to show that our method works at least as good as the classical analytical models like CPM without the need for any experiments.

1 Introduction

Packing of particles and estimating the properties of a random dense packing are problems that appear in many fields of materials science, see, e.g. [15] for a comprehensive overview. A basic property of a packing is its *packing density*, or *space filling*, i.e. the ratio of the net volume of the particles and the volume of the packing.

Our work is particularly motivated by problems from concrete research. Advanced technologies like Ultra-High Performance Concrete (UHPC) require a specific granulometric composition that reduces the void fraction in the dry mixture [11] to obtain a harder and less porous concrete. Typically, the particles in a dry concrete mixture have sizes that vary over an extremely wide range, e.g. from less than 0.1 µm (filler, e.g. fly ash or microsilica) up to 200 µm (cement) with even larger particles at 2 mm and above if aggregates like gravel or sand are added.

To find a particle mixture with a high packing density in a systematic way, it is advantageous to have a tool that allows to estimate the packing density of a given mixture with as little effort as possible. In particular, one would like to avoid the time consuming experiments with the real mixture and its components. As there is no purely analytical way available, a solution that uses only computer simulations would be useful. In particular, one could then estimate the packing density even before the mixture exists.

To this end, we first have to characterize the mixture in a suitable way. In this paper, we restrict ourselves to (hard)

Corresponding author, E-mail: kolonko@math.tu-clausthal.de

Institute of Mathematics, Clausthal University of Technology, Erzstraße 1, D-38678 Clausthal-Zellerfeld, Germany

three-dimensional spherical particles and neglect any other (e.g. chemical) properties or environments (like water in concrete). This simplification is somewhat justified as we are only interested in the static geometric result of the packing. It also allows to deal with huge sample sizes in acceptable computing time. Hence the only relevant characteristic for the packing density is the frequency distribution of the particle diameters in the mixture, the so called *particle size distribution* (PSD) which we call f.

There is a rich literature on the search for PSDs that yield a good space filling starting with the classical work of Fuller [5, 1] more than a century ago. Although this work is sometimes still regarded as a standard, results on "gap gradings" [13] indicate that PSDs with some intermediate particle sizes missing may result in higher packing densities.

Generally, there seem to be two approaches for the estimation of packing densities in the more recent literature. The so-called *analytical approach* uses a mathematical model of how particles of different size interact geometrically and derives an approximate formula for the packing density of a mixture. As input parameters, these methods need empirical data for each (more or less monodisperse) component in the mixture, such as its mean diameter, eigenpacking density or even densities of all binary mixtures of components. Examples of this analytical approach include the Toufar model [17], the linear packing model by Stovall [12] and the mixture packing model by Yu and Standish [20] who also developed a unification of the last two models, the linear-mixture packing model (LMPM) [21, 22]. Another prevalent approach is the compressible packing model (CPM) by de Larrard [3, 4]. Numerous experiments, described, e.g. in [8], show that these approaches allow to estimate the packing density at least of certain simple mixtures with high accuracy.

The second approach to estimate the packing density takes a (theoretical) sample of particles from the distribution and simulates the random close packing of these particles on a computer. There are a number of different algorithms to perform the simulation, see, e.g. [7, 2] and [10] for an overview. Though this approach seems to be "model-free" in contrast, e.g. to the linear packing model, it turns out that to some extent the results depend on the algorithms used. As Torquato [16] (see also [9]) pointed out, the term "random close packing" is not well defined. The randomness as well as the closeness depend on the packing algorithm and on the measures used for randomness.

Both approaches do not work well with wide range size distributions. For the analytical approach too many components have to be examined experimentally to obtain the necessary input data and for the simulation approach the sample size required to obtain a *representative* sample is much too large to be simulated on present computers. The necessary

size of samples will be considered in greater detail in Section 3 below.

In this paper, we present a new approach to estimate the packing density that is capable to handle size distributions with a wide range. It benefits from simulation to determine the packing density of small restricted samples of particles and it uses a mathematical model to obtain an estimator for the overall packing density from these simulations. We divide the particle size range into several smaller intervals (called *fractions*) from which samples can be drawn that allow to simulate the packing density of these restricted ranges. We then use a simple recursive formula to calculate the packing density of the whole mixture from the separate simulation results. In contrast to, e.g. the aforementioned packing models, we do not use an elaborate mathematical approximation formula with elementary data but a simple mathematical formula with data obtained from simulating a complex sample that includes the interaction of different particle sizes, see below. Besides being less costly, the purely software based approach has the additional advantage that it does not depend on any technical restrictions like sieve sizes or limited precision of measurements. The only limitations we know of are the number of spheres that can be packed within reasonable time and—with our present packing software—the inability to handle agglomerations of particles (see below). Note that the present paper is mainly concerned with the divide-and-conquer aspect of the approach and not with the exact details on how the packing of a single fraction can be simulated.

Hierarchical approaches to deal with extreme polydisperse mixtures have been used before. Webb and Davies [19] address this problem with a reduced-dimension algorithm where spheres are sequentially deposited into nested cylinders. Packing characteristics can then be estimated from the innermost cylinder which is the only one that contains spheres of all radii. Though particle radius ratios up to 80:1 were simulated, the number of different sphere sizes is limited because each single radius needs to be represented by enough spheres to fill its own cylinder. Another notable method to reduce the number of particles is the hierarchical model developed by Wackenhut et al. [18]. They collect smaller particles into artificial "soft particles" of larger diameter. Starting with the smallest particles, on each level of diameters all true and soft particles are condensed into soft particles on the next level until the number of remaining (true and soft) particles is small enough to be packed in a simulation. This approach reduces the complexity and polydispersity of the packing to a well-manageable size, but it excludes any interaction with particles hidden in the soft particles. In particular, the small particles can no longer fill the interstices between the larger ones and the packing density will generally be underestimated.
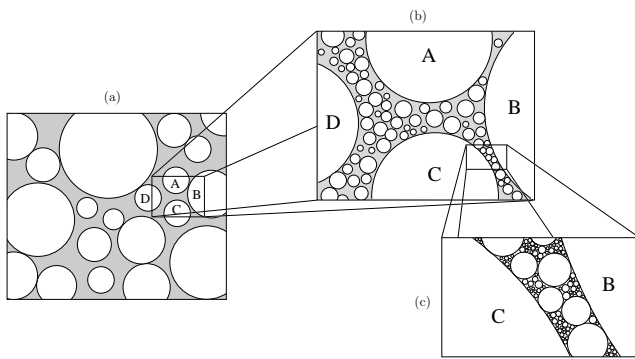
Fig. 1 The hierarchical approach treats smaller spheres as homogeneous substance between larger ones.

The main idea of our approach is explained in Figure 1. When simulating the packing of particles sampled from the interval of largest radii as, e.g., in Fig. 1 (a), we consider the empty space between the spheres as filled by some homogeneous substance with a particular density. This density is the packing density that is determined by simulating the packing of spheres from the next smaller interval as indicated by Fig. 1 (b). This again requires first to simulate the packing of smaller spheres as in Fig. 1 (c) and so on. Technically, we start by simulating the packing density of the smallest spheres and then insert this value when packing the next larger spheres.

Treating the empty space in a packing as if it was evenly filled by spheres from smaller fractions neglects the limited arrangement of small particles in the interstices between larger ones. This error is reduced by the way the space filling of the smaller fraction is simulated: we enlarge the sample to be packed by additional larger particles. After a dense packing has been obtained, the packing density is calculated with respect to the smaller particles only. The additional larger particles are considered to be "holes" in the container. Thus, the packing density of the smaller particles is simulated in the presence of larger ones yielding a more realistic result that includes the interaction of particles across the fractions. This is indicated in Fig. 1 (b) where the spheres A–D from the larger fraction are present, simulating the boundary between the fractions.

We discuss in detail how the sampling mechanism should be adapted to the choice of the intervals and to the size distribution. To obtain representativity of our samples we use a particular stratified sampling method also known as systematic sampling (see, e.g. [14]) that guarantees to always include some of the (typically) rare large particles in our sample. As will become clear, the two aspects, systematic sampling and recursive estimation, are closely related to each other.

We also present some experimental results showing that our hierarchical approach yields excellent estimates of the packing density (within 2% of experimental results). The

problem with real mixtures of, e.g., concrete is that particles of size below a certain threshold tend to form rigid agglomerates of irregular shape that may be far from spherical even if the particles themselves are spheres. These agglomerates cannot be packed densely and therefore usually decrease the packing density when compared to the packing of the corresponding non-agglomerated spheres. The simulation algorithms we used for the packing cannot at present simulate agglomeration. We therefore first use six different mixtures of glass spheres in which no agglomeration occurs to compare our simulation results with the experimental data (and with results from other analytical tools as they were mentioned above). In a second step we used realistic concrete mixtures. Here, we compared our (agglomeration-free) results with the results from the analytical tools only (they can be made to neglect agglomeration also). It turns out that only de Larrard's compressible packing model [4] yields results of comparable quality at the price of a slightly higher effort.

Let us stress once more that the focus in the present paper is *not* on the packing of the sample itself but on how the results from separate packings may be collected for the overall packing density and how the interaction between particles of different fractions can be taken into account. Since our method is basically independent of a particular sphere packing method, we shall not discuss in detail the actual algorithm we used in our experiments. It is an algorithm of the "collective rearrangement" type that was adopted from [7]. It places the spheres into a container that is too small initially, so that overlaps between the spheres must occur. These overlaps are reduced iteratively by simulating a local repulsion of the spheres and by enlarging and shrinking the container until practically no overlaps remain while the container is kept as small as possible.

Also it might be possible that our approach as summarized in Section 9 works well with soft and/or non-spherical particles if an appropriate packing algorithm for these particles would be available. One requirement is that such an algorithm should be able to pack at least, say 50 000 particles in reasonable time as this simulation has to be repeated more than 30 times (in our examples). The model will most likely be applicable to hard disks, i.e. to two-dimensional spheres. At present, our experience is restricted to hard spheres in 3D which is also the adequate model for our main application in concrete research.

The present paper is organized as follows: we start with a more formal definition of the packing density in Section 2. The shortcomings of a conventional sampling and simulation method for extreme polydisperse mixtures are shown in Section 3. In the next three Sections we present the main ideas of estimating the packing density from subdivided diameter ranges that constitute the backbone of our approach. In Section 7 and 8 we discuss how to choose the fractions and sample sizes and how to draw the samples. The com-

plete algorithm RESOS is summarized in Section 9. In Section 10, we present some empirical results of our approach, in particular we compare it to other standard models. Finally, a conclusion is given followed by an Appendix on statistical properties of our sampling technique.

2 Estimating the Space Filling

As mentioned before, we restrict ourselves to mixtures of ideal hard spherical particles as any more complicated model would prevent us from using samples as large as they are needed in our setting. The mixtures are completely characterized by the relative frequencies of the particle diameters appearing in the PSD f (or, equivalently, by the volume share of spheres with a certain diameter). We assume that no agglomeration occurs.

Fig. 2 shows the PSD of a real concrete mixture that is used for our simulations in Section 10 below. Both the frequency as well as the volume share are shown as step functions with steps determined by the limited resolution of the measuring device (e.g., mesh width of sieves) through which the components of the mixture pass. The values of f are a result of the recipe of the mixture, e.g. in this case it is "12.5% fly ash, 37.5% cement (10-63 μm) and 62.5% granulated cinder (125-250 μm)" where, as it is usual in materials science, the percentage is given with respect to mass. Volume share and relative frequency have to be calculated from that based on the specific weight of the different materials. In Fig. 3 the same functions are given with a logarithmic y-scale showing that the mixture has a certain logarithmic homogeneity with respect to volume: the logarithm of the volume share of particles is almost constant in the size of the particle, whereas its frequency decreases with growing size.

We assume that the PSD f is normalized such that the area under the graph sums to one, then f is a probability density function. We also use the cumulative density function (cdf) F with

$$F(t) := \int_0^t f(s)\,ds, \quad t \geq 0.$$

Fig. 4 shows the cdf for the two densities of Fig. 2.

Selecting a particle (a diameter) at random from the mixture can be simulated by sampling from f. As was mentioned above, the steps of f correspond to intervals $(s_{v-1}, s_v]$, $v = 1, 2, \ldots, n$ of diameters which cannot be distinguished in practice. When sampling from f, one would first select an interval $(s_{v-1}, s_v]$ with probability

$$p(s_v) := F(s_v) - F(s_{v-1}) \tag{1}$$

and then select a value from $(s_{v-1}, s_v]$. The most natural way would probably be to select it according to a normal distribution concentrated on the interval as this would in some sense

simulate the behavior of particles in a measuring device. Our experiments showed, however, that it makes no difference if we select just one fixed value from each interval, say its right endpoint. This is probably due to the fact that intervals with high density values are extremely small. In this way, we have actually turned our continuous density function f into a discrete distribution p concentrated on the right endpoints $s_v, v = 1, \ldots, n$, of the intervals. We shall therefore use $p(\tau)$ to denote the probability of the step (interval) that has τ as its right endpoint.

The above example shows two things which are typical for the distributions we consider in this paper: first, the range of possible diameters covers several magnitudes and secondly, larger particles are relatively rare but their volume cannot be neglected. Note, that in the example frequencies of particles larger than 1 μm can only be distinguished in the logarithmic scale of Fig. 3.
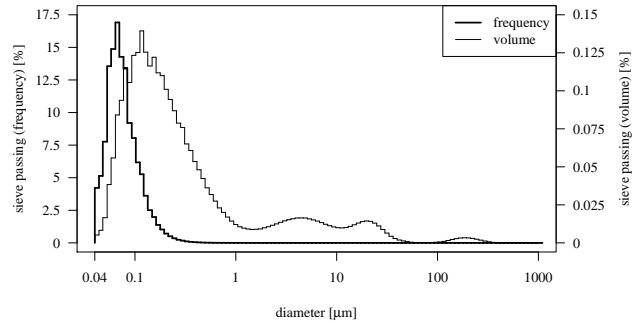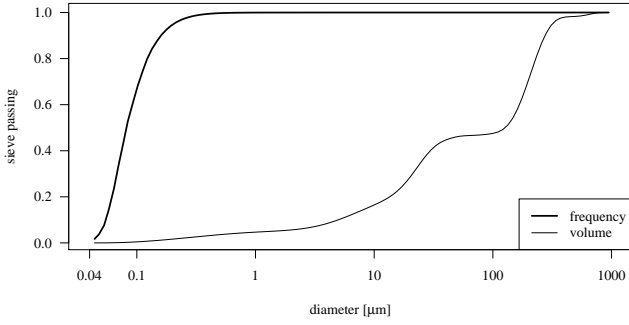

Fig. 2 The relative frequencies of diameters of particles and their relative share in the volume of a mixture given in detail in Section 10.2.
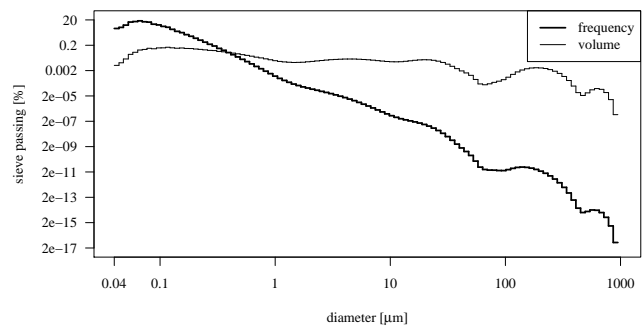
Fig. 3 The relative frequencies of diameters of particles and their relative share in the volume with logarithmic y-axis.

In the sequel, we assume that we are given a PSD f with its cumulative density function F. We start with a definition of the packing density corresponding to f. Let

$$x := (x_1, \ldots, x_M)$$

Fig. 4 The cumulative density function for frequencies of diameters and their relative share in the volume.

denote a large sample of diameters drawn randomly from the mixture. We denote by

$$V(x) = V(x_1, \ldots, x_M) := \sum_{j=1}^{M} \frac{\pi}{6} x_j^3 \qquad (2)$$

the *volume of the spheres* from this sample. In spite of the problems connected with defining and measuring random close packings as they were mentioned in the Introduction, we assume here that we can associate a number $C(x) = C(x_1, \ldots, x_M)$ with the sample x that gives the *space requirement*, i.e. the volume of the smallest container holding a random close packing of the spheres with diameters $x_1, \ldots, x_M$.

Then we define the *packing density of sample x* by

$$\varphi(x) = \frac{V(x_1, \ldots, x_M)}{C(x_1, \ldots, x_M)}. \qquad (3)$$

If the sample is large enough, we may use $\varphi(x)$ as an estimate of the true packing density $\Phi(f)$ of the whole mixture. In the literature, the symbols V_V or ρ are sometimes used instead of $\Phi(f)$. A theoretical packing density of $\Phi(f) = 0.8$, e.g., means that in a random close packing of an arbitrarily large sample 80 % of the space is covered by spheres and 20 % remains empty on the average.

We could improve the estimator $\varphi(x)$ by repeatedly drawing independent samples $x^{(1)}, \ldots, x^{(n)}$, each of size M and use the average of $\varphi(x^{(1)}), \ldots, \varphi(x^{(n)})$.

3 Naive Sampling for Space Filling

To get an impression of suitable sample sizes M, note that our packing program mentioned in the Introduction presently has the capacity to simulate the dense random packing of $M = 1\,000\,000$ spheres in one run of about 38 h on a Intel Xeon computer running at 2.33 GHz.

If the 1 000 000 diameters are sampled randomly from a cumulative density function F as in Fig. 4, the probability that the sample contains at least one diameter *larger* than t is

$$\alpha(t) := 1 - F(t)^{1\,000\,000} \qquad (4)$$

If we take, e.g., $t = 213.2\ \mu\text{m}$, we obtain for the function F depicted in Fig. 4

$$F(213.2\ \mu\text{m}) \approx 0.999999999168772 \quad \text{hence}$$
$$\alpha(t) \approx 1 - 0.999999999168772^{1\,000\,000}$$
$$\approx 0.000830882490483.$$

Therefore, random samples even as large as 1 000 000 will hardly ever pick any of the diameters larger than 213.2 μm. However, the relative share in the volume of particles larger than 213.2 μm is about 25 %.

To measure the *representativity* of a sample x from a given PSD f, we look at the rarest size τ of the PSD, i.e. τ is the right endpoint of that step for which f attains its lowest non-zero value. We assume throughout this paper that this rarest size τ is also the largest size. In the PSD of Fig. 2 f attains its minimum on the last interval from 863.9 μm to $\tau := 948.3\ \mu\text{m}$ with probability $p(948.3) \approx 10^{-15}$, p as defined in (1). As a rough measure of how representative a sample x of size M is, we use the mean number $\beta := M \cdot p(\tau)$ of particles from this interval appearing in the sample. In the example above we obtain $\beta = M \cdot 10^{-15}$. If we want to make sure that on the average at least one of the largest spheres is contained in the sample, i.e. $\beta \geq 1$, we would have to take a sample of size $M \geq 10^{15}$ and simulate its packing. This is clearly impossible on todays standard computers and may also be intractable on larger clusters or grids as packing algorithms are in general not well suited for parallelization and grid computers due to the complex interaction between the particles. It is an additional advantage of our hierarchical approach that the simulation of the fractions (see below) can be performed in parallel.

Thus it seems impossible to include the smallest spheres and spheres of diameters larger than, e.g. 213.2 μm (in our example) in one representative sample that could be packed with a standard algorithm. Still this interaction is important for the estimation of the packing density as the smaller spheres might fill the interstices between the larger ones more or less densely.

We solved this problem by (a) splitting the estimation of the overall packing density into a set of recursive estimations of smaller size over parts of the PSD and (b) applying an appropriate sampling scheme that includes the boundary effects between these parts.

4 Packing Density over a Subdivided Diameter Range

We first describe how the packing density of the mixture as a whole can be obtained from results based on restricted diameter ranges only.

Let $(0, T]$ be the interval of possible diameters of spheres. We partition this interval into m smaller intervals

$$D_1 := (t_0, t_1], D_2 := (t_1, t_2], \ldots, D_m := (t_{m-1}, t_m], \qquad (5)$$

with $t_0 = 0$ and $t_m = T$. We call these intervals *fractions*. The simplest way to determine the D_i would be to choose fractions of equal length, or to use the sieve sizes from which the PSD was built. A more appropriate partitioning is discussed in Section 7.

We now consider a single random sample $x = (x_1, \dots, x_M)$ drawn from $(0, T]$ from which we want to estimate the packing density $\varphi(x) = V(x)/C(x)$ as defined in (3). Our aim is to derive an approximation of φ that is based on separate simulations over the fractions $D_i, i = 1, \dots, m$. We start by studying how the sample $x = (x_1, \dots, x_M)$ from $(0, T]$ behaves on the fractions.

Let $x_{(i)} := (x_{i,1}, \dots, x_{i,\kappa_i})$ be the part of x that falls into fraction D_i, i.e. $t_{i-1} < x_{i,j} \le t_i, j = 1, \dots, \kappa_i$.

Then

$$V(x_{(i)}) = V(x_{i,1}, \dots, x_{i,\kappa_i}) = \sum_{j=1}^{\kappa_i} \frac{1}{6} \pi\, x_{i,j}^3. \qquad (6)$$

is the net volume of the spheres from sample x that fall into fraction D_i. Again, let us assume that the space needed for a dense packing of these spheres is well defined and can be calculated exactly. Denote this space requirement by $C(x_{(i)})$. Then $C(x_{(i)}) - V(x_{(i)})$ is the empty space between the densely packed spheres from fraction D_i.

Similarly, let $C(x_{(1)}, \dots, x_{(i)})$ denote the space requirement for a dense packing of all spheres from x that fall into one of the first i fractions $D_1, D_2, \dots, D_i$, i.e. have diameters $\le t_i$. Then, $C(x_{(1)}, \dots, x_{(m)}) = C(x)$ is the space requirement of the whole sample.

The following recursive relation between $C(x_{(1)}, \dots, x_{(i)})$ and $C(x_{(i)})$ plays a crucial role in our algorithm (note that this is only an approximation details of which are explained below):

$$
\begin{aligned}
&C(x_{(1)}, \dots, x_{(i)}) = C(x_{(i)}) \\
&\quad \text{if } C(x_{(1)}, \dots, x_{(i-1)}) \le C(x_{(i)}) - V(x_{(i)}) \quad \text{and} \\
&C(x_{(1)}, \dots, x_{(i)}) = C(x_{(1)}, \dots, x_{(i-1)}) + V(x_{(i)}) \\
&\quad \text{if } C(x_{(1)}, \dots, x_{(i-1)}) > C(x_{(i)}) - V(x_{(i)})
\end{aligned} \qquad (7)
$$

or, equivalently,

$$
\begin{aligned}
&C(x_{(1)}, \dots, x_{(i)}) \\
&\quad = \max\left\{ C(x_{(i)}),\, C(x_{(1)}, \dots, x_{(i-1)}) + V(x_{(i)}) \right\}
\end{aligned} \qquad (8)
$$

for $i = 2, \dots, m$. The first if-case "$C(x_{(1)}, \dots, x_{(i-1)}) \le C(x_{(i)}) - V(x_{(i)})$" in (7) holds if the empty space $C(x_{(i)}) - V(x_{(i)})$ of the i-th fraction is large enough to pack all the spheres from the preceding smaller fractions $D_1, \dots, D_{i-1}$ into it. Consequently, the space needed for all fractions up to (and including) the i-th is just the space $C(x_{(i)})$ needed for the i-th fraction alone. See Fig. 5 for a rough sketch of this situation. Note that in the terminology of Yu and Standish [21], D_i is the "controlling component" of the sub-mixture $D_1, \dots, D_i$.

In the other case, the empty space of the i-th fraction is not large enough and hence the volume $C(x_{(1)}, \dots, x_{(i)})$ needed for fractions $D_1, \dots, D_i$ is the space for the fractions $D_1, \dots, D_{i-1}$ plus the net volume $V(x_{(i)})$ of the spheres of fraction D_i. In this case, the large spheres from fraction D_i will not be close together but rather float in the mass of the smaller spheres, see Fig. 6.
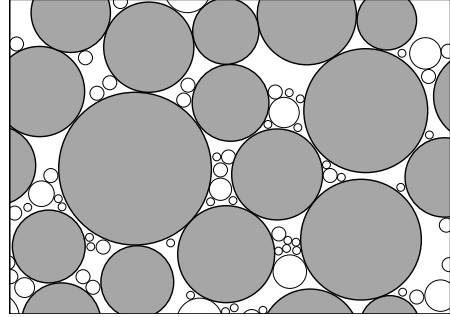
Fig. 5 The dark spheres from fraction D_i dominate the space needed for fractions $D_1, \dots, D_i$.
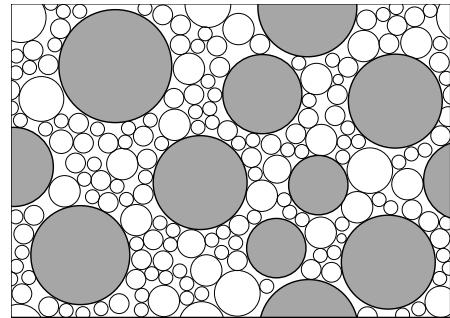
Fig. 6 The dark spheres from fraction D_i are dominated by the smaller fractions.

Of course, this expression is only an approximation to the real situation. In particular, condition $C(x_{(1)}, \dots, x_{(i-1)}) \le C(x_{(i)}) - V(x_{(i)})$ treats the space $C(x_{(1)}, \dots, x_{(i-1)})$ needed by the preceding smaller fractions as a continuum that will fit into the empty space $C(x_{(i)}) - V(x_{(i)})$ no matter what shape the latter has. In reality, there may be small gaps between spheres of the i-th fraction into which none of the smaller spheres from fractions $D_1, \dots, D_{i-1}$ will fit. A similar error occurs if $C(x_{(1)}, \dots, x_{(i-1)}) > C(x_{(i)}) - V(x_{(i)})$. In Section 6, it is shown how this error can be reduced by a more sophisticated setup of the simulation of $C(\cdot)$.

Eq. (8) allows to determine the value $C(x_{(1)}, \dots, x_{(i)})$ when $C(x_{(1)}, \dots, x_{(i-1)})$ is known, by evaluating $C(x_{(i)})$ and $V(x_{(i)})$ which can be done by looking only at sample values from D_i alone. In the next Section, we want to perform this step by a separate simulation with a sample drawn from D_i only.

5 Local Simulation of the Packing Density

When the sample $x = (x_1, \ldots, x_M)$ is partitioned into the subsamples $x_{(1)}, \ldots, x_{(m)}$ on the fractions as in the last section, then the actual size of the i-th subsample $x_{(i)}$ is a random variable. Its mean value is Mp_i, where

$$p_i := F(t_i) - F(t_{i-1}), \quad i = 1, \ldots, m \tag{9}$$

is the probability to pick a value from fraction $D_i = (t_{i-1}, t_i]$. As shown above, Mp_i might become too small for some fractions D_i even for large values of M.

To overcome this problem, we take samples $x_{(i)} := (x_{i,1}, \ldots, x_{i,K_i})$ of *fixed* (large) size K_i from each fraction D_i with $i = 1, \ldots, m$. Then $M := K_1 + \cdots + K_m$ is the total sample size. How to choose $K_1, \ldots, K_m$ is discussed below. We then determine the volume of the spheres and the enclosing container for each sample $x_{(i)}$.

Before we can use this expression in the recursion as (8) we have to take into account that by choosing fixed sample sizes K_i we have given the i-th fraction a relative weight that might not correspond to its importance Mp_i within the original mixture. To come up with the same weight at least on the average we therefore have to multiply the volumes based on the fixed sample sizes by a correction factor

$$\zeta_i := \frac{Mp_i}{K_i} \tag{10}$$

which is the ratio of the expected no. of values from D_i in a random sample of size M and the fixed sample size K_i used now. In this way, the importance of the samples $x_{(i)}$ are balanced among each other according to the whole distribution.

Now let

$$V_i := \zeta_i \sum_{j=1}^{K_i} \frac{1}{6} \pi x_{i,j}^3 \tag{11}$$

be the volume of the spheres with the correct relative weight and let

$$C_i := \zeta_i C(x_{(i)}) \tag{12}$$

be the corresponding weighted container volume. In analogy to (8) we define

$$C_{1\ldots i} := \max\{C_i, C_{1\ldots i-1} + V_i\}, \quad i = 2, \ldots, m, \tag{13}$$

with $C_{1\ldots 1} := C_1$. Hence $C_{1\ldots i}$ describes the space requirement of spheres from fractions $D_1, \ldots, D_i$, estimated from collection of "local" samples $x_{(1)}, \ldots, x_{(i)}$ with fixed sizes $K_1, \ldots, K_i$. This way, we may enforce representative samples from fractions that would otherwise appear rarely in a random sample from the whole range.

6 Including Boundary Effects

As was noted above, when evaluating $C_{1\ldots i}$ in (12) we neglect the behavior of spheres from smaller fractions at the possibly ragged "boundary" formed by spheres from larger fractions. We can improve the estimators if we take this special interior boundary effect into account while simulating the expected container size for $x_{(i)}$ with $1 \le i < m$.

To do so, we first draw a sample $x_{(i)} = (x_{i,1}, \ldots, x_{i,K_i})$ from D_i for each $1 \le i < m$ as before. We call the corresponding spheres the *active* spheres. Then we draw additional spheres (i.e. diameters) $y_{(i)} = (y_{i,1}, \ldots, y_{i,k_i})$ of size k_i from the remaining *larger* fractions $D_{i+1}, \ldots, D_m$. These spheres are called the *boundary* spheres. How to choose the numbers K_i and k_i is discussed below. For the enlarged sample

$$z_{(i)} := (x_{i,1}, \ldots, x_{i,K_i}, y_{i,1}, \ldots, y_{i,k_i}) \tag{14}$$

a dense random placement is simulated to obtain a minimal container volume $C(z_{(i)})$. Now the volume of the boundary spheres is subtracted from the container volume, resulting in the space requirement of the active spheres from D_i in the presence of a sample of larger spheres. Again correcting the relative weight of the container we obtain

$$C_i^+ := \zeta_i \left(C(z_{(i)}) - \sum_{j=1}^{k_i} \frac{1}{6} \pi y_{i,j}^3 \right), \tag{15}$$

for $i = 1, \ldots, m-1$. This means that after the packing of all spheres from $z_{(i)}$, the boundary spheres are treated as if not belonging to the sample, they form holes in the container creating a container boundary to the active spheres that is similar to the conditions in a huge sample from the whole diameter range. Note that for the last fraction $D_m = (t_{m-1}, T]$ no boundary spheres are possible, hence $k_m = 0$.

This improved container volume together with the volume of the corresponding active spheres is now inserted into (12):

$$C_{1\ldots 1}^+ := C_1^+ \tag{16}$$
$$C_{1\ldots i}^+ := \max\{C_i^+, C_{1\ldots i-1}^+ + V_i\}$$

The overall packing density is now estimated by

$$\hat{\Phi}(z_{(1)}, \ldots, z_{(m)}) := \frac{V_1 + \cdots + V_m}{C_{1\ldots m}^+}.$$

In Section 10.1 below it is shown that with this refinement, the estimation of the packing density becomes much more realistic.

Note that $\zeta_1, \ldots, \zeta_m$ are relative weights only, we may use $Q \cdot \zeta_i$ for some constant Q in the above formula without changing the final estimator $\hat{\Phi}$ of the packing density as the common factor Q may be reduced from the fraction. This fact is used below in Section 9 to simplify matters.

7 Determining Fractions and Sample Sizes

Determining the fractions D_i, the sample sizes K_i, k_i, $i = 1, \ldots, m$ and the way samples are drawn are tasks that are closely related to each other. One may start fixing the fractions and then derive suitable sample sizes but in our experiments it turned out to be more adequate to start with the sample sizes.

Let N be the maximal number of spheres that can be packed in a single run of our packing program within reasonable time. Hence, the total sample size of the enlarged samples $(x_{(i)}, y_{(i)})$ (see (14)) should be N. We now fix the sizes of active and boundary part of these samples to be

$$K_i \equiv K \ \text{ and } \ k_i \equiv N - K, \ i = 0, \ldots, m-1, \ K_m := N,$$

where K is a given number. Note, that we then have the total sample size $M = (m-1)K + N$ and the correction factor ζ_i from (10) becomes $(m-1+N/K)p_i$ which may be replaced by $\zeta_i := p_i$, see the remark at the end of Section 6.

We can now derive fractions $D_i = (t_{i-1}, t_i]$ by the following argument. The i-th enlarged sample $(x_{(i)}, y_{(i)})$ of size N is drawn from $D_i \cup D_{i+1} \cup \cdots \cup D_m = (t_{i-1}, T]$. The active spheres $x_{(i)}$ are drawn from $D_i = (t_{i-1}, t_i]$ alone. Hence K/N, the relative size of $x_{(i)}$ within $z_{(i)}$ should reflect the relative weight of D_i within $(t_{i-1}, T]$. If this weight is measured by its relative frequencies or probabilities as given by the cdf F, we have the requirement

$$\frac{F(t_i) - F(t_{i-1})}{1 - F(t_{i-1})} \ = \ \frac{K}{N} \tag{17}$$

or equivalently

$$F(t_i) \ = \ 1 - (1 - K/N)(1 - F(t_{i-1}))$$

for $i = 1, \ldots, m-1$ and $F(t_0) = F(0) = 0$, $F(t_m) = F(T) = 1$. Starting with $t_0 = 0$, we see that (17) implies

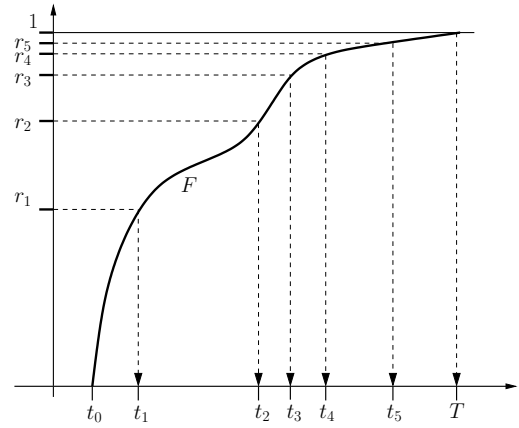$$F(t_i) \ = \ 1 - (1 - K/N)^i$$

or

$$t_i \ := \ F^{-1}\left(1 - (1 - K/N)^i\right), \ i = 0, \ldots, m-1, \ t_m := T, \tag{18}$$

where F^{-1} denotes the inverse[1] of the cdf F. Figure 7 shows an example.

Note that the only parameters that have to be fixed in advance are the number of fractions m, the sample size N and the number of active spheres K. As (18) shows, it is even enough to fix m and the *active share* K/N.

The question remains how to choose K and K/N. In Section 3 we have used β, the mean number of spheres of the rarest diameter τ in a random sample as a measure for the

---

[1] In case F does not allow an inverse function in the strict sense we may use the *generalized inverse* defined by $F^{-1}(r) := \min\{t \geq 0 \mid F(t) \geq r\}$, $r \in [0, 1]$.

Fig. 7 The interval $[0, 1]$ on the y-axis is partitioned by $r_i := 1 - (1 - K/N)^i$, $i = 0, \ldots, m-1$, the dashed lines indicate the operation $F^{-1}(\cdot)$. Here $K/N = 0.5$ and $m = 6$.

representativity of the sample. We assume here as throughout the paper that τ is also the largest diameter and therefore $\tau \in D_m = (t_{m-1}, T]$. As was pointed out before, we cannot expect the value τ to appear in our samples in general. However, in the last sample restricted to $D_m = (t_{m-1}, T]$ the probability that τ is sampled is $p(\tau)/(1 - F(t_{m-1}))$. Hence, the mean number of values τ in this sample (with sample size $K_m = N$) is

$$\beta \ := \ N \cdot \frac{p(\tau)}{1 - F(t_{m-1})} \ = \ N \cdot \frac{p(\tau)}{(1 - K/N)^{m-1}}. \tag{19}$$

Solving for K/N this yields

$$K/N = 1 - \left(\frac{N p(\tau)}{\beta}\right)^{1/(m-1)}. \tag{20}$$

Hence, we may express the active share K/N as a function of the number m of fractions, the maximal sample size N and the relative frequency $p(\tau)$ of the largest diameter together with its required appearance β.

As was indicated above, we may also start fixing the fractions, e.g. as fractions of equal weight $F(t_i) - F(t_{i-1}) \equiv 1/m$. Then

$$t_i := F^{-1}\left(\frac{i}{m}\right), i = 0, \ldots, m$$

and to maintain the ratio (17) we have to choose

$$K_i := \frac{N}{m - i + 1}, \quad k_i := N - K_i, \quad i = 0, \ldots, m-1.$$

8 Systematic Sampling

So far we have assumed that the sample points are generated by drawing randomly (with replacement) from the fraction according to the PSD f. As was pointed out before, if F

has regions where its value does not change much as, e.g. in Fig. 4 for $t > 100\,\mu\text{m}$, then the rare values might be difficult to catch even when sampling is restricted to fractions, see (19).

To overcome this problem at least partly, we here suggest a sampling method known as systematic sampling (see [14] and the references therein). It guarantees a more evenly spread sample and fits well into the fractioned sample space we use. We assume that the maximal sample size N is fixed and that the fractions $D_i = (t_{i-1}, t_i]$ are already determined. Instead of drawing active and boundary spheres separately, we now draw one enlarged sample $z_{(i)} = (z_{i,1}, \ldots, z_{i,N})$ of size N from the interval

$$(t_{i-1}, T] = D_i \cup D_{i+1} \cup \ldots \cup D_m$$

of all sphere diameters contained in D_i or larger. The sample points $z_{i,j}$ that fall into D_i form the active spheres, the rest are the boundary spheres. Systematic sampling as explained below guarantees that the sample sizes K_i, k_i resulting from this procedure are as required.

To draw a standard random sample from $(t_{i-1}, T]$ one would use the "inversion principle". That means that N points $\theta_1, \ldots, \theta_N$ are sampled uniformly from the interval $(F(t_{i-1}), 1]$ (on the y-axis), then the sample points $F^{-1}(\theta_1), \ldots, F^{-1}(\theta_N)$ are used. In contrast to that, systematic sampling from $(t_{i-1}, T]$ starts by dividing the interval $(F(t_{i-1}), 1]$ into N subintervals of equal length

$$\varepsilon_i := \frac{1 - F(t_{i-1})}{N}. \tag{21}$$

From each of these subintervals one point is chosen, the first, say ξ_i, is chosen randomly (uniformly distributed) from the first subinterval $(F(t_{i-1}), F(t_{i-1}) + \varepsilon_i]$, then the next ones are chosen exactly ε_i apart, i.e.

$$\xi_i, \ \xi_i + \varepsilon_i, \ \xi_i + 2\varepsilon_i, \ldots, \xi_i + (N-1)\varepsilon_i, \tag{22}$$

see Fig. 11 in the Appendix for an example. Then the sample points are:

$$z_{i,j} = F^{-1}(\xi_i + j\varepsilon_i), \quad j = 0, \ldots, N-1, \tag{23}$$

see Fig. 8 below. Note that in contrast to the formulation in Section 5 where we assumed fixed sample sizes K_i, k_i of spheres for each fraction, here the numbers of active and boundary spheres have become random variables depending on the sample. However, as it is shown in the Appendix, due to the systematic sampling, these random numbers deviate from the prescribed K_i and k_i by at most 1.

Fig. 8 shows an example of systematic sampling applied to $m = 6$ fractions with sample size $N = 5$. The fraction boundaries t_i are chosen as in Fig. 7 with $K/N = 0.5$. The sampling for the first fraction is shown in detail: three of the selected diameters fall into $D_1 = (0, t_1]$ and are active

spheres indicated by black circles, whereas the white circles are boundary spheres as described in Section 6. For the second fraction D_2, only the equidistant points on the y-axis are given. The ones marked with black arrows will lead to the active spheres, the white ones to boundary spheres. Note that here we have only two active spheres. From $K/N = 0.5$ we have $K = 2.5$.
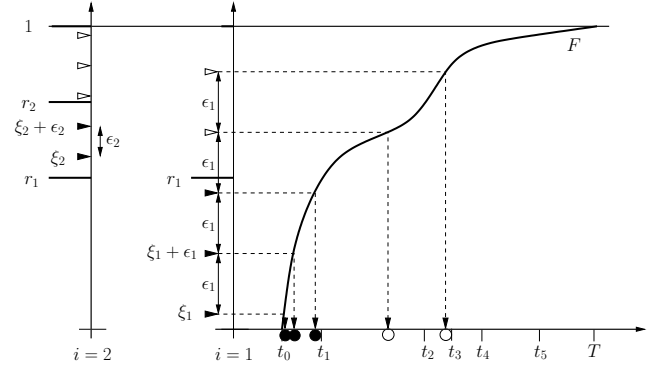
Fig. 8 For the first sample the equidistant points $\xi_1 + j\varepsilon_1, j = 0, \ldots, 4$ are shown on the y-axis together with the sample points. For the second sample different values of ξ_2 and ε_2 are used as described in the text.

9 Summarizing the Complete Algorithm

Before we give some experimental results we summarize the complete algorithm RESOS (recursive estimation with systematic overlapping sampling) with both its parts, the recursive estimation as given in Section 5 and the particular way of fixing parameters and taking samples as described in the last two Sections:

Input: a PSD f on the range $(0, T]$ with cdf F, τ the largest value $t \in (0, T]$ with $f(t) > 0$; m the number of fractions; N the maximal allowable number of spheres per run of the packing program and β, the minimal average number of largest spheres with diameter τ to appear in the last subsample (see (19)).

Output: an estimate of the packing density $\Phi(f)$.

Steps of the algorithm :

I.Setup Determine the active share K/N as

$$K/N := 1 - (Np(\tau)/\beta)^{1/(m-1)}.$$

Partition the interval $(0, T]$ into fractions $D_i := (t_{i-1}, t_i]$, $i = 1, \ldots, m$ with

$$t_i := F^{-1}\left(1 - (1 - K/N)^i\right), \quad i = 0, \ldots, m-1,$$

$t_m := T$.

Put $\zeta_i := F(t_i) - F(t_{i-1}), i = 1, \ldots, m$.

II. For each $i = 1, \ldots, m$ do

II.1 Sampling Put $\varepsilon_i := (1 - F(t_{i-1}))/N$, draw ξ_i uniformly distributed from the interval $(F(t_{i-1}), F(t_{i-1}) + \varepsilon_i]$ and let

$$z = (z_1, \ldots, z_N)$$
$$:= \left(F^{-1}(\xi_i), F^{-1}(\xi_i + \varepsilon_i), \ldots, F^{-1}(\xi_i + (N-1)\varepsilon_i) \right).$$

Define $L_i := \max\{l \mid z_l \leq t_i\}$, then $z_1, \ldots, z_{L_i}$ are in D_i and form the active spheres.

II.2 Simulation Put

$$V_i := \zeta_i \frac{1}{6}\pi \sum_{j=1}^{L_i} z_j^3$$

and simulate the container volume $C(z)$ for a random close packing of z. Let

$$C_i^+ := \zeta_i \left(C(z) - \frac{1}{6}\pi \sum_{j=L_i+1}^{N} z_j^3 \right).$$

III. Recursion Define $C_{1\ldots i}^+$ recursively as

$$C_{1\ldots 1}^+ := C_1^+,$$
$$C_{1\ldots i}^+ := \max\left\{ C_i^+, C_{1\ldots i-1}^+ + V_i \right\}, \quad i = 2, \ldots, m.$$

IV. Estimation Obtain as estimator for the packing density $\Phi(f)$

$$\hat{\Phi}(f) := \frac{V_1 + \cdots + V_m}{C_{1\ldots m}^+} \tag{24}$$

V. Repetition If necessary, repeat steps II - IV with independent samples and take the average to obtain an estimator with a smaller variance.

To get an impression of the time complexity of this algorithm we compare it to a packing algorithm without fractions. Let us assume that a typical packing algorithm (e.g. of collective rearrangement type) needs time $O(M \log(M))$ to pack a sample of size M. We choose M such that the largest value τ has a prescribed expected frequency $\beta = M \cdot p(\tau)$ as described in Section 3. For a fair comparison, the same value of β should hold for the fractioned sampling, i.e. from (19)

$$M \cdot p(\tau) = \frac{Np(\tau)}{(1 - K/N)^{m-1}} \quad \text{or} \quad N = M \cdot (1 - K/N)^{m-1}.$$

As we need m simulations of samples of size N in RESOS it has time complexity

$$O(m \cdot N \log(N))$$
$$= O(m \cdot M \cdot (1 - K/N)^{m-1} \log(M \cdot (1 - K/N)^{m-1})). \tag{25}$$

The relative share K/N of active spheres is always less than 1, therefore the time complexity of RESOS decreases exponentially fast for increasing number of fractions. For the particular value $K/N \approx 0.5$ as in our experiments below, the run

time bound for RESOS in (25) is smaller than $M \log(M)$ for all $m > 2$. As the simulations in RESOS may be performed one after another (or in parallel), the space complexity is $O(N) = O(M \cdot (1 - K/N)^{m-1})$ for RESOS and $O(M)$ for the unfractioned simulation.

In RESOS by far the most time is needed for the simulation of C_i^+ in step "II.2 Simulation". As we will show now, some of these packing simulations may be skipped or stopped if we assume that the packing is built up iteratively with an increasing packing density (as it is the case in the rearrangement algorithms). To see this, note that in the main recursion (16) the exact value of C_i^+ is not needed if it is smaller than $C_{1\ldots i-1}^+ + V_i$ as we have

$$C_{1\ldots i}^+ = C_{1\ldots i-1}^+ + V_i$$
$$\iff C_i^+ < C_{1\ldots i-1}^+ + V_i. \tag{26}$$

As $C_{1\ldots i-1}^+ + V_i$ may be calculated before the simulation of C_i^+ starts, we can abort the simulation as soon as the container used for the simulation of C_i^+ embodies a valid sphere packing and has a volume less than $C_{1\ldots i-1}^+ + V_i$. No further densification of the packing is then necessary, we have $C_{1,\ldots,i}^+ = C_{1\ldots i-1}^+ + V_i$ and we are in a situation as indicated in Fig. 6.

If we are able to give a lower bound b on the "local packing density" V_i/C_i^+ we may even skip the simulation of C_i^+ completely if $C_{1\ldots i-1}^+ + V_i \geq V_i/b$. As V_i/b is an upper bound on C_i^+ we then have $C_{1\ldots i-1}^+ + V_i \geq C_i^+$. E.g., in a polydisperse mixture we can be sure that $V_i/C_i^+ \geq 0.4$ for all fractions. If $C_{1\ldots i-1}^+ + V_i \geq V_i/0.4$ for some i, we have $C_i^+ \leq V_i/0.4 \leq C_{1\ldots i-1}^+ + V_i$ and we may skip the simulation of C_i^+ and obtain $C_{1,\ldots,i}^+ := C_{1\ldots i-1}^+ + V_i$. In this sense, (26) may be used as a stopping criterion for unnecessary simulations.

10 Simulation Results

In this Section we present some simulation results to show the validity of our approach and to compare it to the analytical tools as sketched in the Introduction.

10.1 Mixture of glass spheres

The most convincing test would be to compare results from our algorithm RESOS to experimental results for real mixtures with a broad size distribution. As was mentioned above, real packings of, e.g. concrete mixtures contain ill-shaped agglomerations of small particles that usually decrease the density of the packing. Simulating the effect of these agglomerations on the packing density is quite difficult. With our present packing tool, we have to restrict ourselves to

spherical, non-agglomerating particles for a direct comparison with experimental data.

We use six different mixtures of glass spheres that do not agglomerate. Three mixtures follow the classical Fuller curve (Fu1 - Fu3), the rest are simpler bi- and tridisperse mixtures (Bi1, Bi2, Tri), see Table 1 for details. We compared our RESOS algorithm to the analytical model of Toufar [17], the LMPM [22] and the CPM [4]. These models need as input data the empirical packing densities of the monodisperse packings for each sphere diameter, the so-called eigen-densities, which are given in Table 2.

diameter	mass fraction [%]					
[mm]	Fu1	Fu2	Fu2	Bi1	Bi2	Tri
0.444	37.95	26.59	20.56	34.41	—	34.05
0.940	18.06	12.65	9.78	—	40.36	50.94
1.552	16.13	11.30	8.74	65.59	—	—
2.963	27.86	19.52	15.09	—	—	15.01
6.003	—	29.94	23.18	—	59.64	—
10.022	—	—	22.66	—	—	—

Table 1 The mixtures of glass spheres used in the experiments.

diameter [mm]	eigen-density
0.44	0.64476
0.940	0.65943
1.552	0.64879
2.963	0.64315
6.003	0.62562
10.022	0.59147

Table 2 The empirical eigen-densities of monodisperse packings of the glass spheres.

The packing densities resulting from the different methods are shown in Figure 9. The experimental results were obtained using a pycnometer, they are averages from seven measurements each. For our RESOS algorithm we used the (moderate) sample size $N = 50\,000$ with $m = 6$ fractions and $K = 25\,541$ active spheres resulting in a ratio $K/N = 0.510821$. The results are also averages of seven runs. As can be seen, the estimated packing densities are very close to the experimental ones, the difference being less than 2 % on the average.

The model of Toufar tends to underestimate the density, a similar observation was made with the modified Toufar model [6]. The linear mixture packing model LMPM [22] constantly overestimates the density. Only the compressible packing model CPM yielded results comparable to our hierarchical approach. Here we used the compaction index 9 (as suggested by de Larrard [4] for vibration and compression).
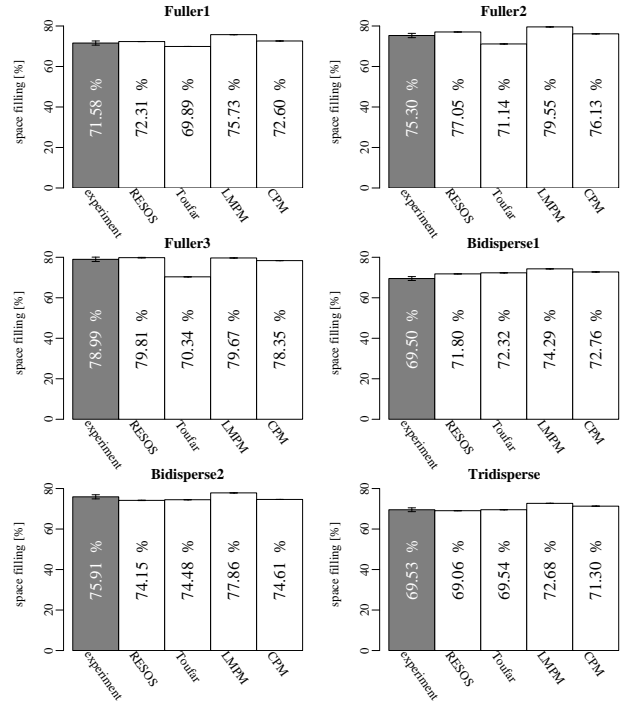
Fig. 9 The packing densities of the glass mixtures from Table 1 as they were obtained by five different methods. For the experimental and RESOS results, standard deviations from seven runs are indicated.

We may also use the data of Tables 1 and 2 to explicitly show the effect of the additional boundary spheres we have used in the simulation of the single fractions. If we use no boundary spheres with the above $m = 6$ fractions we obtain monodisperse packings and the local packing densities $C_i/V_i = C_i^+/V_i$ are just the eigen-densities as given in Table 2. These values may be inserted into the recursion (16) resp. (13) for the six mixtures of Table 1. Table 3 shows the results. Here, the second and third column repeat the values from Figure 9, the last column shows the new estimates obtained without boundary spheres. It is obvious, that these values are far from the true experimental values and far from the estimates of RESOS with boundary spheres.

Mixture	Empirical	RESOS with boundary spheres	RESOS without boundary spheres
Fuller1	71.58	72.3130	82.70675
Fuller2	75.30	77.0482	87.22186
Fuller3	78.99	79.8149	89.82571
Bidisperse1	69.50	71.8008	84.06279
Bidisperse2	75.91	74.154	82.75106
Tridisperse	69.53	69.0561	84.20319

Table 3 Comparison of packing densities (in %) of RESOS with and without boundary spheres

10.2 Concrete Mixtures

To show that our hierarchical approach in RESOS can cope with broader size ranges, we applied it to two full-range PSDs. The first is the one depicted in Figures 2–4, it consists of 12.5% fly ash, 37.5% cement (10-63 μm) and 62.5% granulated cinder (125-250 μm), where all % are % by mass. The second mixture consists of 50% cement (CEM I 32,5R) and 50% fly ash, its PSD is shown in Figure 10. In both mixtures, a high percentage of the small particles form agglomerates. As our RESOS simulation does not contain agglomerations, we only compared it to results obtained from analytical tools, namely LMPM and CPM.
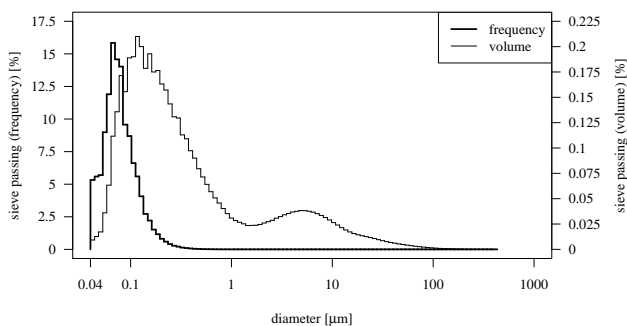
Fig. 10 The relative shares in volume and frequency of the second mixture used for our comparison.

Both LMPM and CPM use the eigen-densities of the components of the mixture, i.e. small intervals of the PSD that could be treated as more or less monodisperse fractions. If these densities are determined experimentally, they will contain the effects of agglomeration. To make results comparable, we used eigen-densities as simulated by our packing tool, without any agglomerations. This also had the side effect to make efforts for the different approaches comparable: it is the total sample size of simulated packings.

To determine the parameters for RESOS as described in Section 7 we identify the largest diameter in the PSD of Figure 2 as $\tau = 948.3$ with

$$p(\tau) = 4.551914 \cdot 10^{-15}.$$

We therefore obtain from (20)

$$\frac{K}{N} = 1 - (N \cdot 4.551914 \cdot 10^{-15})^{1/(m-1)}. \tag{27}$$

Applying the same sample size $N = 50\,000$ and a similar share of active spheres K/N as with the glass spheres, we obtain from (27) $m = 33$ as necessary number of fractions, leading to $K/N \approx 0.5003$. In a similar way, we obtained for the second mixture with the largest diameter $\tau = 373.1$: $p(\tau) \approx 1.506527 \cdot 10^{-15}$, $m = 34$ fractions with $N = 50\,000$ and $K/N \approx 0.5066$.

The most natural components for LMPM and CPM are the diameter intervals determined by the sieve sizes which results in 108 components for the first and 98 components for the second mixture. The eigen-densities of these components were obtained by simulation, each with 50 000 spheres as in RESOS. For a fair comparison, we also applied LMPM and CPM with 33 and 34 components only, i.e. with the same total of sample points as in RESOS. The results are given in Table 4.

method	mixture of Fig. 2		mixture of Fig. 10	
	no. of fractions or components	packing density [%]	no. of fractions or components	packing density [%]
	m		m	
RESOS	33	89.2460	34	86.4908
LMPM	33	88.7871	34	86.5861
LMPM	108	89.5217	98	86.3390
CPM	33	89.6559	34	86.3665
CPM	108	89.5493	98	86.3012

Table 4 Results for the two concrete mixtures. Each component or fraction was simulated with $N = 50\,000$ spheres.

For the first quite complex mixture it turned out that results from LMPM with 33 components, i.e. with the same effort as RESOS, were not satisfactory. Only with a much higher effort, with $m = 108$ components, LMPM was able to obtain results of similar quality as RESOS. For the second mixture, differences seem insignificant.

The CPM needs the so-called compaction index K. As the simulation uses ideal conditions (exact spheres and periodic container walls), we used the value $K = \infty$ which can be derived from de Larrard's formulas [4]. With this parameter, CPM produced about the same results as RESOS even with just $m = 33$ or $m = 34$ components.

From these results one may conclude that RESOS yields slightly better results than LMPM and is comparable to CPM if this has a suitably adjusted compaction index. Still we claim that our method is more flexible than the empirical models as it definitely needs no experiments. Moreover, CPM has the disadvantage that its performance depends crucially on the right compaction index which for general particles has to be determined by experiments, even if the component data should be obtained from simulation.

11 Conclusion

In this paper, a hierarchical method RESOS for the estimation of the space filling of polydisperse sphere packings has been developed. It divides the particle size distribution of the mixture into a number of smaller fractions. Although

the fractions are treated separately the method takes into account the filler effect of small particles between bigger ones by including spheres from other larger fractions as the container's inner boundary. This has been combined with a sampling method which ensures representativity of the sample even for rare particles as they typically appear in mixtures like cement, concrete or mortar. The method can be used with different packing techniques, present simulation results are based on collective rearrangement.

RESOS yields fast and reliable estimates for the packing density of mixtures with a broad size distribution. It opens the possibility to systematically search for particle size distributions with high packing density, as they are needed in materials science. However, before one can think about designing high density mixtures on the computer, at least two problems have to be solved: the packing of non-spherical particles and the simulation of agglomerates within packings. These will be the topics of our future work.

Appendix

We assume that the fractions $D_1, \ldots, D_m$ and the maximal sample size N are fixed and that systematic sampling is applied as described in Section 8. Now fix i and let L_i be the random number of sample points in the i-th sample that fall into D_i and lead to active spheres. In (17) the ideal value for the share K_i/N of active spheres is given as

$$R_i := \frac{F(t_i) - F(t_{i-1})}{1 - F(t_{i-1})}, \quad i = 0, \ldots, m-1. \tag{28}$$

We want to show that L_i/N is as close to R_i as possible, more precisely

$$|L_i - NR_i| \leq 1 \quad \text{and} \quad \mathbb{E}L_i = NR_i$$

for $i = 1, \ldots, m-1$, where $\mathbb{E}$ indicates expectation. For a proof first note that L_i is the number of values from $(z_{i,1}, \ldots, z_{i,N}) = (F^{-1}(\xi_i), \ldots, F^{-1}(\xi_i + (N-1)\varepsilon_i))$ that fall into $D_i = (t_{i-1}, t_i]$. It must be equal to the number of values $\xi_i + j\varepsilon_i, j = 0, \ldots, N-1$ that fall into $(F(t_{i-1}), F(t_i)] = (r_{i-1}, r_i]$ where we put $r_i := F(t_i)$. As the $\xi_i + j\varepsilon_i$ have constant distance ε_i we see that

$$L_i = \max\{l \in \mathbb{N} \mid \xi_i + l\varepsilon_i \leq r_i\} + 1 = \left\lfloor \frac{r_i - \xi_i}{\varepsilon_i} \right\rfloor + 1$$

where $\lfloor x \rfloor$ is the largest integer less or equal to x. Note that we have "+1" to account for $\xi_i + 0\varepsilon_i$. As ξ_i takes on values in $(r_{i-1}, r_{i-1} + \varepsilon_i]$ only (see Section 8), we obtain

$$\left\lfloor \frac{r_i - r_{i-1}}{\varepsilon_i} \right\rfloor \leq L_i \leq \left\lfloor \frac{r_i - r_{i-1}}{\varepsilon_i} \right\rfloor + 1.$$

From (21) and (28) we have that $NR_i = (r_i - r_{i-1})/\varepsilon_i$, hence the random variable L_i can take on only the two values $\lfloor NR_i \rfloor$

or $\lfloor NR_i \rfloor + 1$ and $|L_i - NR_i| \leq 1$ follows. Fig. 11 shows an example for two different values of the starting point ξ_i. The black arrows indicate values $\xi_i + j\varepsilon_i$ that lead to active spheres whereas the white arrows belong to boundary spheres.
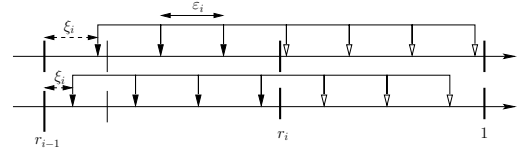
Fig. 11 Depending on the first point ξ_i, the number of points $\xi_i + j\varepsilon_i$ in $(r_{i-1}, r_i]$ may take on two different values.

From straightforward but tedious calculations it follows that L_i takes on the value $\lfloor NR_i \rfloor + 1$ with probability $NR_i - \lfloor NR_i \rfloor$ from where $\mathbb{E}L_i = NR_i$ follows.

More generally, it can be shown in a similar way that with systematic sampling the number of sample points that fall into an arbitrary interval I deviates from the expected number of points in I under *random sampling* only by at most 1 and has the same expectation.

References

1. A. H. M. Andreasen and J. Andersen. Über die Beziehung zwischen Kornabstufung und Zwischenraum in Produkten aus losen Körnern (mit einigen Experimenten). *Colloid & Polymer Science*, 50(3):217–228, March 1930.
2. Katalin Bagi. An algorithm to generate random dense arrangements for discrete element simulations of granular assemblies. *Granular Matter*, 7(1):31–43, apr 2005.
3. F. de Larrard and T. Sedran. Optimization of ultra-high-performance concrete by the use of a packing model. *Cement and Concrete Research*, 24(6):997–1009, 1994.
4. François de Larrard and Thierry Sedran. Mixture-proportioning of high-performance concrete. *Cement and Concrete Research*, 32(11):1699–1704, November 2002.
5. W. B. Fuller and S. E. Thompson. The laws of proportioning concrete. *Trans. Am. Soc. Civ. Eng.*, 59:67–172, 1907.
6. Per Goltermann, Vagn Johansen, and Lars Palbøl. Packing of Aggregates: An Alternative Tool to Determine the Optimal Aggregate Mix. *Materials Journal*, 94(5):435–443, September 1997.

7. D. He, N. N. Ekere, and L. Cai. Computer simulation of random packing of unequal particles. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 60(6):7098–7104, December 1999.

8. M. Jones, L. Zheng, and M. Newlands. Comparison of particle packing models for proportioning concrete constitutents for minimum voids ratio. *Materials and Structures*, 35(5):301–309, June 2002.

9. K. Lochmann, A. Anikeenko, A. Elsner, N. Medvedev, and D. Stoyan. Statistical verification of crystallization in hard sphere packings under densification. *The European Physical Journal B – Condensed Matter and Complex Systems*, 53(1):67–76, 2006.

10. G. T. Nolan and P. E. Kavanagh. Computer simulation of random packing of hard spheres. *Powder Technology*, 72(2):149–155, 1992.

11. M. Schmidt, E. Fehling, and C. Geisenhanslüke, editors. *Ultra High Performance Concrete (UHPC) – Proceedings of the International Symposium on UHPC*, volume 3 of *Schriftenreihe Baustoffe und Massivbau – Structural Materials and Engineering Series*, Kassel, 2004. kassel university press.

12. T. Stovall, F. De Larrard, and M. Buil. Linear packing density model of grain mixtures. *Powder Technology*, 48(1):1–12, 1986.

13. P. Stroeven, M. Stroeven, and D. D. Bui. On optimum packing density. In Ying shu Yuan, Surendra P. Shah, and Heng lin Lü, editors, *Proceedings of International Conference on Advances in Concrete and Structures*, pages 793–800. Rilem Publications SARL, 2003.

14. Yves Tillé. *Sampling Algorithms*. Springer Series in Statistics. Springer-Verlag, New York, 2006.

15. Salvatore Torquato. *Random Heterogenous Materials: Microstructure and Macroscopic Properties*, volume 16 of *Interdisciplinary Applied Mathematics*. Springer, 2nd edition, 2006.

16. Salvatore Torquato, Thomas M. Truskett, and Pablo G. Debenedetti. Is random close packing of spheres well defined? *Physical Review Letters*, 84(10):2064–2067, mar 2000.

17. W. Toufar, M. Born, and E. Klose. Beitrag zur Optimierung der Packungsdichte polydisperser körniger Systeme. *Freiberger Forschungsheft*, 558:29–44, 1976.

18. M. Wackenhut, S. McNamara, and H. J. Herrmann. A Hierarchical Model for simulating very polydisperse Granular Media. In R. García-Rojo, H. J. Herrmann, and S. McNamara, editors, *Powders and Grains 2005*, pages 1005–1008, Rotterdam, 2005. A. A. Balkema.

19. M. D. Webb and I. L. Davis. Random particle packing with large particle size variations using reduced-dimension algorithms. *Powder Technology*, 167(1):10–19, 2006.

20. A. B. Yu and N. Standish. An analytical–parametric theory of the random packing of particles. *Powder Technology*, 55(3):171–186, July 1988.

21. A. B. Yu and N. Standish. Estimation of the porosity of particle mixtures by a linear-mixture packing model. *Industrial & Engineering Chemistry Research*, 30(6):1372–1385, 1991.

22. A.B. Yu, R.P. Zou, and N. Standish. Modifying the Linear Packing Model for Predicting the Porosity of Non-spherical Particle Mixtures. *Industrial & Engineering Chemistry Research*, 35(10):3730–3741, 1996.