Multiscale Shannon's Entropy Modelling of Orientation and Distance in Steel Fiber Micro-Tomography Data

John P. Chiverton*, Olubisi Ige, Stephanie J. Barnett, Tony Parry

Abstract

This work is concerned with the modelling and analysis of the orientation and distance between steel fibers in X-ray Micro-Tomography (XCT) data. The advantage of combining both orientation and separation in a model is that it helps provide a detailed understanding of how the steel fibers are arranged, which is easy to compare. The developed models are designed to summarise the randomness of the orientation distribution of the steel fibers both locally and across an entire volume based on multiscale entropy. Theoretical modelling, simulation and application to real imaging data are shown here. The theoretical modelling of multiscale entropy for orientation includes a proof showing the final form of the multiscale taken over a linear range of scales. A series of image processing operations are also included to overcome interslice connectivity issues to help derive the statistical descriptions of the orientation distributions of the steel fibers. The results demonstrate that multiscale entropy provides unique insights into both simulated and real imaging data of steel fiber reinforced concrete.

I. Introduction

Steel Fiber Reinforced Concrete (SFRC) is a common engineering material in the construction industry. The steel fibers help to improve the mechanical properties of the material. The orientation of the fibers is important because they help to strengthen the material, particularly in the direction orthogonal to the fibers. In the extreme case they can help to bridge a gap if a crack forms. If the fibers are not oriented perpendicular to the direction of the crack then they will not be able to bridge the gap according to Abrishambaf et al. [1]. Thus, the aim of the work here, is to provide a quantitative insight into, how randomly distributed steel fibers or any other randomly oriented structures are located and oriented in volumetric imaging data. Multiscale Entropy modeling is investigated here to describe the randomness of steel fibers in micro-tomography data of SFRC. This enables the orientation distribution and distancing of the fibers to be summarized in an effective way.

A number of techniques exist for estimating the orientation distribution of the fibers. The orientation distribution is a distribution function which can help in understanding how well distributed the fibers are within the volume. The approach taken in Vicente et al. [2] and Abrishambaf et al. [3] was to look at the distributions of the fibers against the 3 axes in combination with fiber distribution efficiency factors, one for each axis (e_x, e_y, e_z). For each of the efficiency factors, a variety of scenarios may be considered, such as complete alignment with a given axis (e.g. $e_x = 2$), perpendicular to that axis (e.g. $e_x = 0$). These are useful as they summarize, for each of the axes, the extensive information in the orientation distribution for a volume. A software system for interactive visualization and investigation of fiber based materials that combined similar axis dependent distribution information was described by Weissenböck et al. [4].

Another approach taken by Axelsson and Svensson in [5], [6] is via a structure tensor that describes the orientations of the fibers through out the volume. The tensor can be summarized via a combination of three relative anisotropy terms (c_1, c_2, c_3) calculated from combinations of the three eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$ of the structure tensor. For the case of when the fibers are well distributed then c_2 will be small, c_1 small and c_3 large. This measurement of anisotropy can be understood in terms of how the eigenvalues are inter-related (since the relative anisotropy terms are derived from the eigenvalues).

Alternatively we may prefer to somehow summarize the distribution of the fibers in terms of how randomly they are distributed. One way of doing this is with the use of a Scalar Order Parameter (SOP). Hermann et al. [7], [8] used SOP as a scalar value, $S \in \left[-\frac{1}{2}, 1\right]$ to quantify the amount of variation in the distribution of fibers. The SOP can be calculated via the second order orientation tensor $\mathbf{O}$ which is given by the sum of the outer products of the orientation vectors. The SOP S can then be calculated from the largest eigenvalue of $\mathbf{O}$ (but scaled by 3/2) or the average of the second Legendre polynomial, (see e.g. Jankun-Kelly and Mehta, [9] and Mottram and Newton [10]) dependent upon the angle α which is the angle between each of the fibers' vectors and the director's vector that describes the mean orientation.

Each of the above techniques rely on the processing of knowledge about the locations of the steel fibers, as might be possible from image processing operations performed on microtomography data. However fiber orientation estimation has also been

Manuscript received October 15 2016; revised XXX. *Asterisk indicates corresponding author.*

J.P. Chiverton is with the School of Engineering, University of Portsmouth, john.chiverton@port.ac.uk

O. Ige and S.J. Barnett are with the School of Civil Engineering and Surveying, University of Portsmouth, olubisi.ige@port.ac.uk & stephanie.barnnett@port.ac.uk

T. Parry is with the Faculty of Engineering, University of Nottingham, tony.parry@nottingham.ac.uk

 DOI: 10.1109/TIP.2017.2722234

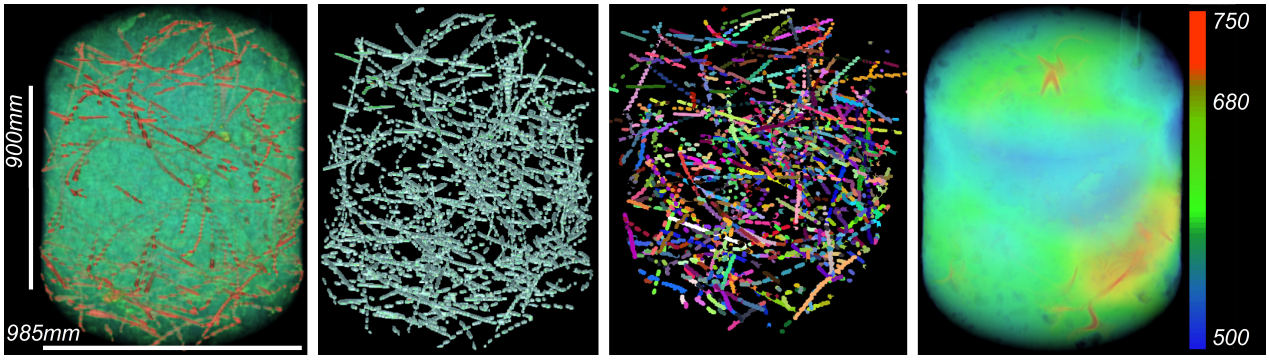


Fig. 1. Example of a volume rendered real Steel Fiber Reinforced Concrete (SFRC) core (left), result of thresholding (middle left), individual steel fibers (middle right) and volume rendered result of a corresponding multiscale Entropy estimation process, as described here. The individual steel fibers are obtained using an iteratively applied Random Sample Consensus process applied to a morphologically thinned version of the segmented steel fibers.

performed using electromagnetic measurement techniques, relying on the magnetic properties of the steel fibers, as shown in Juan-García et al. [11]. These techniques are interesting but a high resolution 3D microtomography data set would likely be considered as a potential source of a gold standard in this type of scenario.

Many of the aforementioned techniques attempt to describe the amount of concentrated directionality in the distribution of the fibers. This is often referred to as *anisotropy*. Anisotropy is a popular topic in the medical imaging community to help understand the connectivity of the brain. Westin et al. [12] derived an eigenvalue based description of the shape of a diffusion tensor to describe the directionality of water diffusion in Diffusion Tensor Magnetic Resonance Imaging (DTMRI) data. Similar techniques were applied to more general computer vision problems by Westin and Knuttson, [13], [14] for symmetry detection and motion vector field estimation. Anisotropy is also useful for fingerprint authentication. For example Jiang, [15] used the eigenvectors of the *gradient covariance matrix* for fingerprint authentication, similar to the structure tensor approach used previously in computer vision by Bigun and Granlund in 1987 [16] and Knutsson in 1989 [17]. In particular the magnitude of the image gradient was used as a measure of anisotropy, which is not so applicable to steel fiber like data as a measure of anisotropy. Another fingerprint indexing approach by Liu and Yap in 2012 [18] looked at polar complex moments at different scales. The use of different scales was interesting but the inherent assumption with fingerprints with their technique is that there must be a direction to infer.

The majority of the techniques described above derive some form of measurement from the eigenvalues of a second order moment matrix, such as the orientation tensor as used by Hermann et al. [7], [8], a structure tensor as used by Axelsson and Svensson [5], [6] or even a covariance like matrix as in Principal Component Analysis (PCA), also used by Axelsson and Svensson [6]. As such the eigenvalues describe the variability of a system, with the greatest eigenvalue in the direction of maximum variability. These approaches implicitly assume a uni-modal distribution. An alternative approach to describing the statistical variability which does not assume a uni-modal distribution can be based on Shannon's Entropy H, first proposed in 1948 [19] for application in telecommunications. A comparative illustration of the different properties of standard deviation versus Shannon's Entropy can be seen in Fig. 14.

Shannon's Entropy is sometimes referred to as useful as a spectral Entropy described by Sharma et al. [20]; similarly for Renyi's Entropy [21]. In the same context, there are other forms and uses of Entropy such as approximate Entropy as proposed by Pincus [22] and sample Entropy described by Richman and Moorman [23]. Both these techniques have often been used in the context of analyzing the complexity of time series data. They are often used as a regularity statistic, e.g. by Ho et al. [24]. Both approximate Entropy and sample Entropy have also been used within a multiscale Entropy framework, see Costa et al. [25] where the multiple scales provide an advantage enabling the randomness or complexity of a signal to be quantified across a range of scales so that signals can be compared more easily. For example, Ramdani et al. [26] and Fino et al. [27] looked at posture fluctuations to compare falling in elderly people; artefact detection in electroencephalographic signals is also another popular topic e.g. Mariani et al. [28]; analysis of various engineering type problems such as for two phase flow e.g. Gao et al. [29]; and quantifying human heartbeat complexity and similar has also been another popular topic for the application of multiscale Entropy e.g. Valencia et al. [30] and Costa and Goldberger [31]. Shannon's Entropy has also been used in multiscale formulations by Zhang [32] and also Fogedby [33], but not for the description and summarisation of orientation information which is the problem considered here.

The steel fibers are included in steel fiber reinforced concrete to provide improved mechanical properties to the concrete. The mechanical modeling of composite materials can be done in a number of ways, such as with Finite Element Analysis (FEA) (e.g. Abbas et al. [34]) or statistically via the Weibull distribution, often used to model tensile strength via weakest link theory, see e.g. Zhang et al. [35]. Both types of techniques have advantages, but they do not incorporate the orientation distribution information, an often used descriptor of fiber enhanced materials. FEA provides insight into the mechanical properties of a material at various levels of detail but statistical techniques are often limited to providing a global summary or at best limited

to sub-volumes. Therefore it is of interest to investigate statistical models that are able to provide a global summary of the orientation information but also to focus at the imaging point measurement level, i.e. for individual voxels.

The work described here is based on a multiscale Entropy approach. It is used to quantify the randomness of the distribution of steel fibers in steel fiber reinforced concrete. The randomness is quantified in 3D both in terms of a 3D angle (ϕ, θ) in section II-A and the 3D angle is combined with distance between fibers in section II-B. An example multiscale Entropy estimate can be seen in Fig. 1 for a real SFRC core sample. Shannon's Entropy rather than approximate Entropy or sample Entropy is used here as Shannon's Entropy quantifies overall randomness rather than irregularity (as for approximate Entropy or sample Entropy). Irregularity is not appropriate here because regularly spaced fibers could still provide improved mechanical properties to a material. Rather a measure of uniformity, whether regular or irregular is appropriate here. Also remembering that the use of multiscale ensures that the properties of the orientation distribution are quantified consistently across scale.

Section III then follows with a series of experiments and corresponding results concluding with a discussion and conclusions in section IV.

II. Methodology

A. Multiple Scale Discrete Orientation Distribution

Each individual steel fiber can be represented as a vector $\boldsymbol{p}$ in 3D cartesian space i.e. $\boldsymbol{p} = (x, y, z)^{\mathrm{T}}$. Alternatively it can be represented in a spherical coordinate system with $\boldsymbol{v} = (\phi, \theta, \rho)^{\mathrm{T}}$, where ϕ and θ are the in-plane and inter-plane angles respectively and r is the Euclidean length of the steel fiber vector. These terms can be calculated with $\phi = \tan^{-1}\left(\frac{y}{x}\right)$, $\theta = \cos^{-1}(z)$, $\rho = \sqrt{x^2 + y^2 + z^2}$. For each fiber i in a volume, discretized angle indices n_i and m_i are calculated with $n_i \leq N$ and $m_i \leq M$ are whole number indices in the range provided by the selected scale given by N and M. The indices n_i and m_i are calculated from the continuous underlying angles θ and ϕ respectively.

A histogram can be created for a volume containing fibers binned at discretized angle indices $f(n, m)$; containing the counts for the fibers at a particular combination of angle ranges. The histogram can be used to estimate a probability mass function for the orientations of the fibers, with $\hat{P}(n, m) = f(n, m) / \sum_n \sum_m f(n, m)$ so that $\sum_n \sum_m \hat{P}(n, m) = 1$. We will assume an infinite sample size so that $P(n, m) = \hat{P}(n, m)$. This assumption is considered in more detail in section II-F. The discretized orientation distribution at a particular scale (N, M) can then be summarized with Shannon's Entropy, i.e.

$$H_{NM} = \sum_n \sum_m P(n, m) h(n, m); \tag{1}$$

where the self information is given by $h(n, m) = -\log_2(P(n, m))$. Extending across multiple scales gives,

$$H_{\mathrm{ms}} = \sum_N \sum_M H_{NM} = \sum_N \sum_M \sum_n \sum_m P(n, m) h(n, m). \tag{2}$$

The question we might ask, is, *what value will H_{ms} take for a volume consisting of fibers distributed over a particular range of angles, e.g. r, where $\theta \in [0, q] \times 180$ and $\phi \in [0, r] \times 180$?* Expressing this as $H_{\mathrm{ms}}(q, r)$; it is shown in appendix A that $H_{\mathrm{ms}}(q, r)$ is approximately logarithmically proportional to q and r, where the Entropy is calculated across a set of scales, with maximum $\mathfrak{M} \times \mathfrak{N}$ i.e. $M \in [1, \mathfrak{M}]$ and $N \in [1, \mathfrak{N}]$. An example of this can be seen in Fig. 2. Also shown in Fig. 2 are data points obtained from simulation (see III). These data points closely follow the theoretical line for the majority of the curve. This is interesting because it demonstrates the result of a multiscale operation is an easily understood function and also one which can be used to indicate the range of angles over which the fibers are distributed. Using similar steps and assumptions, it is also shown in appendix A that the variance of this estimator is given by

$$\sigma^2_{\mathrm{ms}} = 4 \log_2^2(\mathfrak{N}\mathfrak{M}) - \log_2(\mathfrak{N}\mathfrak{M})(4 + 2/\ln(2)) + 4/\ln^2(2). \tag{3}$$

Through numerical simulation it can be seen that mean H_{ms} and the sample standard deviation σ_{ms} are of similar value where $H_{\mathrm{ms}}/\sigma_{\mathrm{ms}} \in [0.8, 1.8]$. The standard error $\mathcal{S}_{\mathrm{ms}} = \frac{\sigma_{\mathrm{ms}}}{\sqrt{\mathcal{C}}}$ for the estimation of H_{ms} will however be significantly reduced because of the numerous fibers in every volume *e.g.* $\mathcal{C} > 70$. A further useful observation regarding (3) is the independence from the range of angles $([0, q], [0, r])$ that a particular volume might have fibers distributed across.

The use of angle alone is interesting because of the original premise regarding highly randomized fibers contributing to a material with good mechanical properties (see e.g. ACI Committee [36]). But the microtomography data provides more detailed information, i.e. local information that can be relied upon to help localize the quantification of the randomness, through another type of Entropy calculation, as can be seen in the following section.

B. Probability of Angle and Distance Combined

The mechanical properties of a material should improve when fibers are distributed evenly throughout a volume in a highly randomized way (see e.g. ACI Committee, [36]). The orientation distribution only describes the angular distribution, not the spatial distribution. This spatial information is not captured in the preceding model. Furthermore, as noted above, the microtomography data provides highly detailed, localized information regarding the location of fibers. This information can be

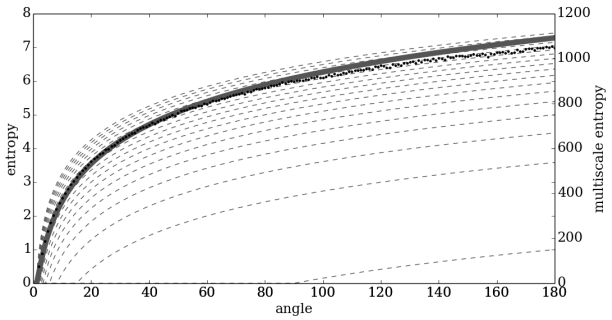


Fig. 2. Illustration for 2D of the effect of including multiple scales (dashed lines), with a solid gray line for the resulting combination of the multiple scale lines. Also shown are dots corresponding to simulations which closely follow the theoretically derived expression, (in Appendix A).

used to develop a new, more sophisticated model of fiber distribution; both at the voxel level and also capturing information about the distribution of the fibers throughout the volume.

The probability of angle and distance combined can be considered. For a particular point $\boldsymbol{x}_j$ in the volume and a particular direction (n, m), at angular scale (N, M) and distance scale σ, we have

$$P(\boldsymbol{x}_j|n, m, \sigma) \propto \sum_{\substack{i:n_i=n,\\ m_i=m}} \exp\left(-\frac{d_{\text{ls}}(j|i)^2}{2\sigma^2}\right) \tag{4}$$

where $d_{\text{ls}}(j|i)$ is the distance from point $\boldsymbol{x}_j$ to a steel fiber with end points $\boldsymbol{x}_{1,i}$ and $\boldsymbol{x}_{2,i}$. Details on the computation of this distance can be found in section II-D.

The probabilities are normalized for each point across all discretized angles. Using Bayes theorem it can then be shown

$$P(n_i, m_i, \sigma|\boldsymbol{x}_j) = \frac{P(\boldsymbol{x}_j|n_i, m_i, \sigma)P(n_i, m_i, \sigma)}{\sum_{n_k}\sum_{m_k} P(\boldsymbol{x}_j|n_k, m_k, \sigma)P(n_k, m_k, \sigma)}, \tag{5}$$

which describes the probability of fibers oriented in direction (n_i, m_i) at distance scale σ influencing the volume at point $\boldsymbol{x}_j$. Uniform priors are assumed because fibers could potentially point in any direction i.e. $P(n_k, m_k, \sigma) = 1$ and the denominator is a marginal distribution $P(\boldsymbol{x}_j)$ so that

$$P(n_i, m_i, \sigma|\boldsymbol{x}_j) = \frac{P(\boldsymbol{x}_j|n_i, m_i, \sigma)}{P(\boldsymbol{x}_j)}, \tag{6}$$

where $P(\boldsymbol{x}_j) = \sum_{n_k}\sum_{m_k} P(\boldsymbol{x}_j|n_k, m_k, \sigma)$. This marginal is important in terms of the conventional notion of conditional probabilities. The normalizing sum or marginal $P(\boldsymbol{x}_j)$ also carries important details regarding the proximity and consequently influence of any fibers, which is preserved to scale the resulting Entropy calculations as will be seen shortly.

C. Conditional Multiscale Entropy

The influence of fibers pointing in direction (n, m) at varying distances d_{ls} to a point $\boldsymbol{x}_j$ are important in understanding whether an individual point is sufficiently influenced in terms of proximity, as well as the complexity of the distribution of orientations. Therefore the marginal term $p(\boldsymbol{x}_j)$ from (6) is introduced to weight the Entropy calculation in terms of the distance of influencing fibers d_{ls} for a particular orientation (n, m). The result of this is a conditional Entropy of a point $\boldsymbol{x}_j$ being influenced by fibers pointing in directions (n, m) for a particular scale (N, M, σ)

$$\begin{aligned} H(N, M, \sigma|\boldsymbol{x}_j) &= E\left[-\log_2 P(n, m, \sigma|\boldsymbol{x}_j)|\boldsymbol{x}_j\right] \\ &= -\sum_n\sum_m P(n, m, \sigma, \boldsymbol{x}_j)\log_2 P(n, m, \sigma|\boldsymbol{x}_j) \\ &= -\sum_n\sum_m P(n, m, \sigma|\boldsymbol{x}_j)P(\boldsymbol{x}_j)\log_2 P(n, m, \sigma|\boldsymbol{x}_j) \end{aligned} \tag{7}$$

The conditional Entropy at point $\boldsymbol{x}_j$ across all scales is then

$$H_{\text{vol}}(\boldsymbol{x}_j) = \sum_\sigma\sum_N\sum_M H(N, M, \sigma|\boldsymbol{x}_j) \tag{8}$$

Each $H_{\text{vol}}(\boldsymbol{x}_j)$ describes the Entropy across multiple scales of orientations (N, M) and distance (σ). This conditional Entropy will be used shortly (in III) to quantify the amount of randomness in the orientation distribution of steel fibers.

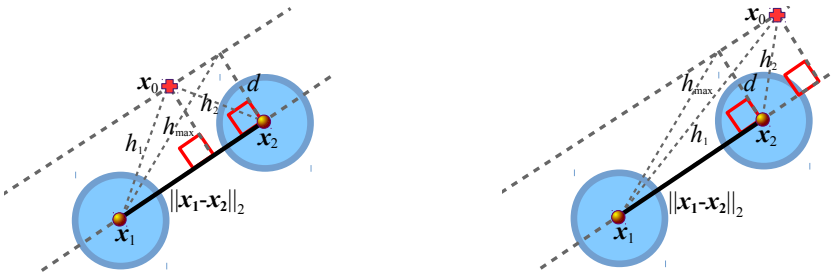


Fig. 3. Illustration of the geometry involved in the calculation of the distance from the line segment, d_{ls}. When $h_1, h_2 \leq h_{\max}$ (left) then $d_{\text{ls}} = d$. When h_1 or $h_2 \geq h_{\max}$ (right) then $d_{\text{ls}} = \min(h_1, h_2)$.

D. Fiber Modeling

Steel fibers typically have hooked ends to help improve anchorage in the concrete, see e.g. Abdallah et al. [37]. Sometimes they may also have a zig-zag form. Furthermore many of the steel fibers will also bend when combined with aggregate (small stones) in the concrete mix.

The steel fibers are modeled here as line segments, a straight line but terminated at two end points. This provides a convenient approach for the purposes of modeling. Furthermore, the fiber detection process (see see section II-E) does not assume a constant fiber length. This means that a bent fiber can be approximated with a series of shorter straight line segments.

If a point $\boldsymbol{x}_0$ is perpendicular to the line segment with end points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ then the perpendicular straight line distance can be used. Otherwise the distance will need to be calculated in another way, as shown in the following lemma.

Lemma. *The minimum distance of a point $\boldsymbol{x}_0$ to a line segment with end points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ is defined as*

$$d_{\text{ls}} = \begin{cases} d & \text{if } h_1 \leq h_{\max} \text{ and } h_2 \leq h_{\max}; \\ \min(h_1, h_2) & \text{otherwise;} \end{cases} \tag{9}$$

where $d = \frac{\|(\boldsymbol{x}_0 - \boldsymbol{x}_1) \times (\boldsymbol{x}_0 - \boldsymbol{x}_2)\|_2}{\|\boldsymbol{x}_2 - \boldsymbol{x}_1\|_2}$ is the perpendicular distance of $\boldsymbol{x}_0$ to a straight line with points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ as described by Weisstein, [38], $h_1 = \|\boldsymbol{x}_0 - \boldsymbol{x}_1\|_2$, $h_2 = \|\boldsymbol{x}_0 - \boldsymbol{x}_2\|_2$ and $h_{\max} = \sqrt{d^2 + \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_2^2}$.

Proof. The distance to a line segment from a point $\boldsymbol{x}_0$ can be calculated by observing that two right angle triangles are formed between $\boldsymbol{x}_0$ and the closest point on the straight line and the two end points of the line segment $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ (forming two different right angle triangles). The hypotenuses h_1 and h_2 for the two triangles are then the distances between two end points, i.e. $h_1 = \|\boldsymbol{x}_0 - \boldsymbol{x}_1\|_2$ and $h_2 = \|\boldsymbol{x}_0 - \boldsymbol{x}_2\|_2$. A third right angled triangle is also formed, defined as having the points $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ and a point on a line running through $\boldsymbol{x}_0$ parallel to the line with $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, perpendicular at $\boldsymbol{x}_1$ (or equivalently $\boldsymbol{x}_2$) with hypotenuse $h_{\max} = \sqrt{d^2 + \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_2^2}$. If both hypotenuses h_1 and h_2 are less than $h_{\max}$ then there will be a perpendicular point within the line segment. If either h_1 or h_2 are greater than $h_{\max}$ then the right angle triangle formed will have a perpendicular point that lies outside of the line segment (although still on the straight line). In this case, the perpendicular distance (to the straight line) will be less than the distance to the closest point on the line segment. This is illustrated in Fig. 3. The minimum distance from $\boldsymbol{x}_0$ to the line segment can therefore be calculated with (9) as given. □

E. Fiber Detection

The directionality of steel fibers in volumetric data may be corrupted because of the thickness of the steel fiber in relation to the inter-plane thickness in the acquired XCT data (inter-slice gap). This can be observed in the 3D rendering and cross section shown in Fig. 4. Techniques such as the one proposed by Fritz et al. [39] used region growing to segment the background concrete from isolated steel fibers then assumed that individual fibers could be individually isolated. Another technique proposed by Vicente et al. [2] involved taking a relatively high threshold and then determining fibers based on their connectivity in a particular direction within a given tolerance by performing a regression like process of the resulting point cloud for the entire set of segmented points. Other techniques by Eberhardt and Clarke [40], Salaberger et al. [41] and Tausif et al. [42] rely on finding short straight lines in the 3D image data and then connecting them together, again similar to Vicente et al. [2] in a particular direction within a given tolerance. The resulting potential ambiguities such as the directionality means simple techniques based on connectivity are not sufficient for accurate detection of fibers. Random sample consensus (RANSAC), first proposed by Fischler and Bolles [43] is used here as a robust clustering model based detection method to robustly and iteratively fit models, (straight lines in this case) corresponding to steel fibers. The volumetric data is thresholded, morphologically thinned (see e.g. Gonzalez and Woods [44]) and then RANSAC is applied to find the set of points that best match a straight line. Those points are then deleted and then the process is repeated until no data points remain. This approach is reasonable and is often used in computer vision research, see e.g. Davies [45]. An example output can be seen in Fig. 1.

If the imaging data is acquired with a sufficiently high resolution, including little or no inter-slice gap then other techniques could instead be used such as the one proposed by Salaberger et al. [41].

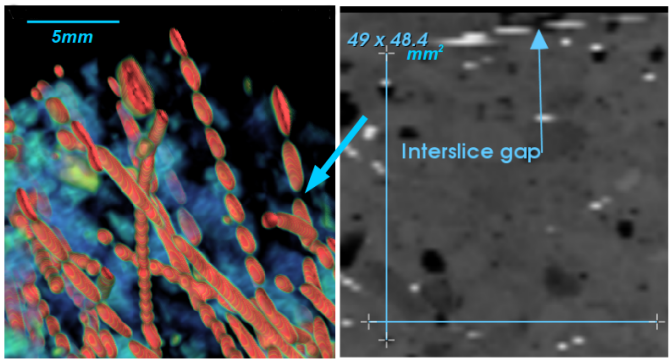


Fig. 4. Illustration of effect of inter-plane finite resolution when the dimensions of the steel fibers are small compared to the inter-plane resolution. Left: cropped image of 3D rendering of isosurfaces (red) together with volume rendering of semi-transparent aggregate material (blue/ green). Right: 2D gray-scale cropped cross section of same volume also illustrating fibers (0.75mm diameter) with poor inter-connectivity and a scale is included (mm). Voxel dimensions given by $0.01652^2 \times 1\text{mm}^3$ for a core size of $\pi 50^2 \times 100\text{mm}^3$ with 197000 fibers/m^3. Data provided by a user study, see section III-B, particularly table II.

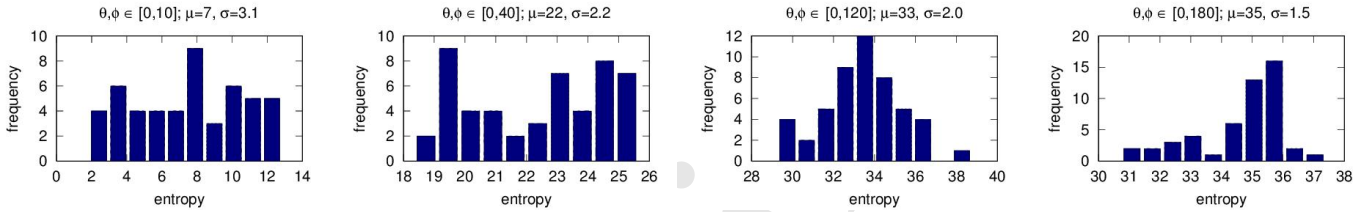


Fig. 5. Histograms for repeated estimation of the mean Entropy for simulated fiber volumes (similar to bootstrapping with replacement sampling). These results appear to show clustering around centralized mean like values. Standard deviations range from 1.5 to 3.1, all reflecting relatively low deviation from the mean. These results help to show the repeatability of obtaining a particular value of Entropy using the described processes. Also note the estimation process includes the fiber detection and Entropy estimation processes.

F. Entropy Estimation

The steel fibers are likely to be quite numerous, thus complicating the fiber detection process. Furthermore accurate estimation of Entropy $\hat{H}$ is dependent on accurate estimation of the underlying probability distribution(s), e.g. using "*plug-in*" Entropy estimation, see Antos and Kontoyiannis [46]

$$\hat{H}_{\text{MLE}} \equiv -\sum_k \hat{P}_k \log_2 \hat{P}_k. \tag{10}$$

which requires a large sample size to accurately estimate the probabilities $\hat{P}_k$. Obviously, this is impractical from a computational point of view. However the stochastic nature of the fiber detection process provides an opportunity to repeat it over a number of iterations, which can provide somewhat dependent re-sampling of the fiber volume.

The "*plug-in*" formula in (10) for Entropy estimation is also a negatively biased estimator according to Antos and Kontoyiannis [46] and Paninski [47], i.e. $\hat{H}_{\text{MLE}} - H \leq 0$. Thus making robust inference of the Entropy from a quite densely populated volume and or low contrast volume potentially erroneous. The good news is that the variance on the Entropy estimation part of the process is bounded $\text{Var}(\hat{H}_{\text{MLE}}) \leq (\log \mathcal{C})^2/\mathcal{C}$ due to central limit properties [47].

A further potential source of variation is from the fiber detection stage involving the RANSAC algorithm, as described in the preceding section. The RANSAC algorithm does not always provide a perfect solution and the results returned can be somewhat validated by re-running the sampling process. Therefore, not only are there multiple fibers from which to estimate but also the estimation process is repeated multiple times ($\mathcal{C} = \mathcal{X} \times 50$ for some of the experiments that follow shortly where $\mathcal{X}$ is the number of fibers in a volume). This reduces the resulting weighting of a potentially poor fiber detection process and also to take advantage of the central limit theorem. Some results of this bootstrap with replacement like process can be seen in Fig. 5.

III. Experiments and Results

The experiments that follow have utilized both simulated data and real XCT data of steel fiber concrete cores. The benefit of simulating XCT data is because it enables parameters to be precisely controlled such as the orientation distribution and distancing of the simulated steel fibers, the inter slice thickness, noise levels and it is also possible to have precise knowledge about the location of the simulated fibers. The fibers were randomly located in the imaging data using

$$\begin{cases} \theta = 2\pi u \\ \phi = \cos^{-1}(2v - 1). \end{cases} \tag{11}$$

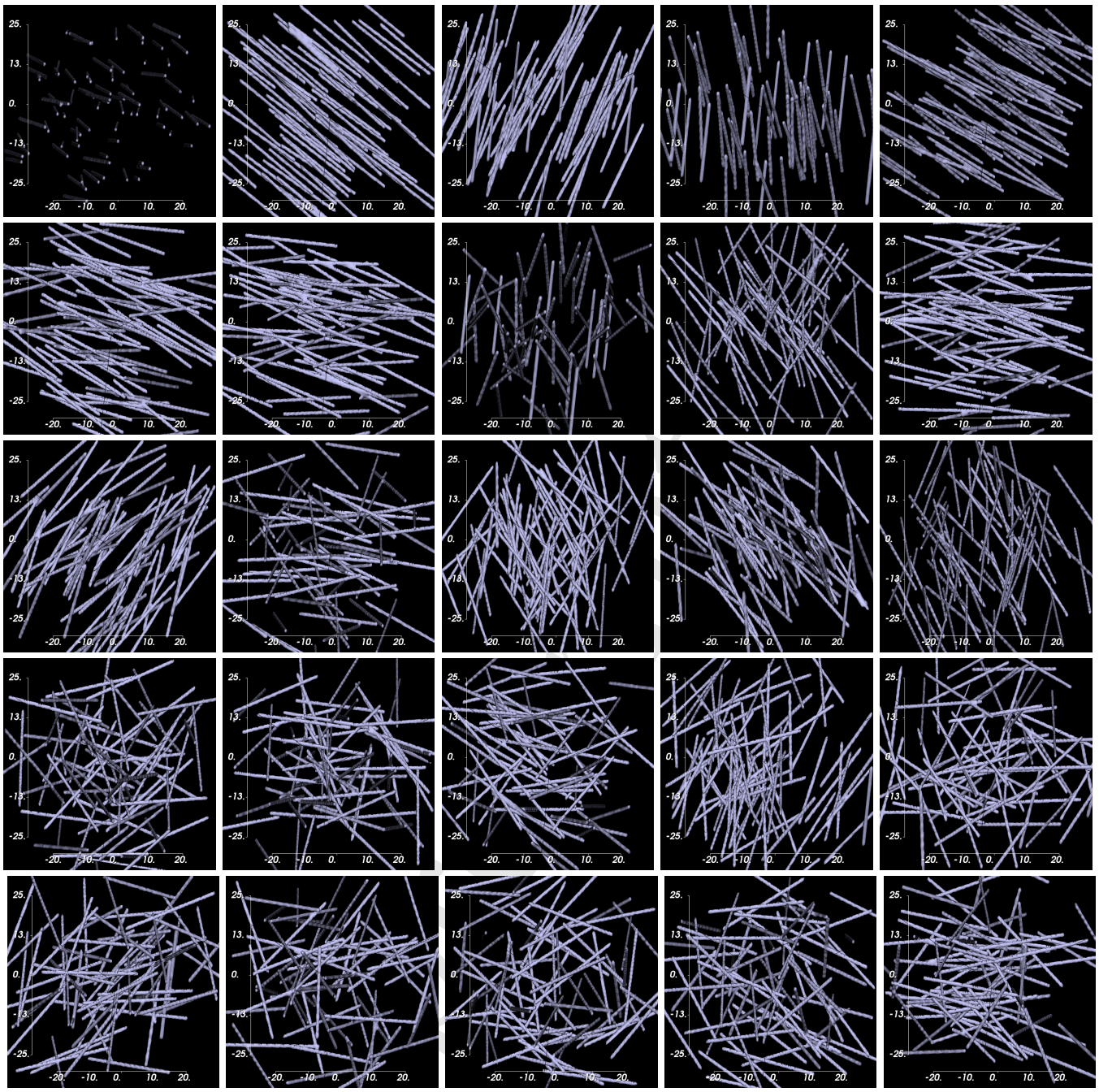


Fig. 6. Surface renderings of simulated volumes. Each row represents increasing ranges of angles over which fibers were distributed, where $\theta, \phi \in (0, 10)$, $(0, 40)$, $(0, 60)$, $(0, 120)$, or $(0, 180)$ where the first row has $\theta, \phi \in (0, 10)$, second row $\theta, \phi \in (0, 40)$ etc. Scale axes are shown with dimensions in mm.

where u and v are sampled from uniform distributions in (0,1). Rejection sampling was used to constrain θ and ϕ within a given range, e.g. to generate fibers in the range $\theta, \phi \in [0, 60]$ then u and v were sampled from the uniform distribution and the corresponding θ and ϕ values were kept unless $\theta > 60$ or $\phi > 60$. Fibers were generated in a high resolution volume as straight lines with 30mm lengths and 1mm diameters. Straight lines were generated and then dilated to create the specified fiber dimension. Additive white Gaussian noise was added and convolution performed with a simulated point spread function, assumed here to be Gaussian in form. Downsampling was then performed to a specified low resolution number of voxels, see Table I. Volume renderings of a range of simulated volumes can be seen in Fig. 6.

The real XCT data are scans of 100mm core samples of steel fiber reinforced concrete square panels that have undergone various mechanical tests such as flexural strength testing as performed by Ige et al. [48]. The cores were imaged with a 350kV X-ray source with a Venlo H350 manufactured by Shaw Inspection Systems. The technical details include a 350kV 2-D mini-focus, fan-shaped beam with a tungsten target, linear diode array using a $1\times$ 3672 pixel area in conjunction with a

TABLE I
DETAILS OF SIMULATED FIBERS VOLUMES, ILLUSTRATED IN FIG. 6 WITH RESULTS SHOWN IN FIGS. 5, 8 AND 7.

θ, ϕ	Downsampled # Voxels	High Resolution # Voxels	# volumes
$[0, 10]$	128×128×128	768×768×768	10
$[0, 20]$	128×128×128	768×768×768	10
$[0, 40]$	128×128×128	768×768×768	10
$[0, 60]$	128×128×128	768×768×768	10
$[0, 90]$	128×128×128	768×768×768	10
$[0, 120]$	128×128×128	768×768×768	10
$[0, 150]$	128×128×128	768×768×768	10
$[0, 180]$	128×128×128	768×768×768	10

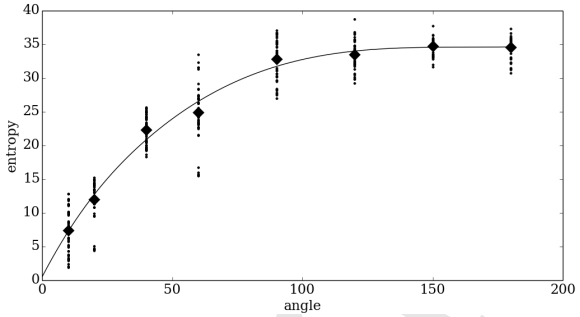


Fig. 7. Multiscale Entropy calculated for simulated volumes (#400) with simulated fibers (#75) constrained to an upper limit for both the in-plane and inter-plane angles (θ, ϕ). Rejection sampling was used to randomly generate within a given interval. The point values are the multiscale Entropy estimates (#50 for each angle). The means, conditioned on the angle are diamonds and an inverse decay exponential curve of the form $H_{\max}(1 - \exp(-\text{angle}/\tau))$ was found to have a very good fit to these mean values.

12-bit digital output and a Gadolinium Oxysulphide scintillator. The focal spot is 0.9mm. The scan parameters included 296kV, current 2.1mA, exposure time of 260ms and each scan was rotated 360^o in 450 seconds.

Unless otherwise stated, the angle histograms for n and m are calculated with $(N, M) \in (3, 4, 5, 6)^2$ so that the orientation distribution is discretized into 16 different combinations of scale, i.e. for $N = 3$, $M = 3, 4, 5, 6$; for $N = 4$, $M = 3, 4, 5, 6$, etc. The multiple scale variance σ^2 in (4) was given values $\sigma^2 \in \{1, 100, 10000, 1000000\}$.

The simulation software was implemented using a number of software tools including C++ with the Insight ToolKit (ITK) for the volumetric processing framework, the Point Cloud Library (PCL), GNU Octave and for visualization purposes Python and the Visualization ToolKit (VTK) were used. The developed software will be made available from Chiverton [49] for the purposes of reproducible research.

A. Fiber Volume Simulations Entropy Estimation Results

The estimation of the Entropy is a random process. It is therefore desirable to look at how the Entropy estimates vary depending on a given orientation angle. Simulated fiber volumes were generated with the distribution of the fibers for particular simulations constrained within a particular range of angles, details of which can be seen in Table I.

Some histograms of the results of Entropy calculations estimated on the simulated volumes can be seen in Fig. 5. This is an interesting comparison where the simulated maximum angle of deviation orientations of the fibers in the simulated fiber volumes are plotted against the resulting multiscale Entropy values calculated for those volumes. Clusters can be seen for the majority of the estimated Entropy values for each fiber orientation distribution. The relatively low standard deviations range from 1.5 to 3.1 help to show the repeatability of a particular Entropy estimate for a fiber orientation distribution.

The conditional Entropy values for the simulated data was then plotted as discrete data points as a function of angle which can be seen in Fig. 7. Also shown are the means of those data points for each angle simulation. A spline curve was fitted to the mean values but an inverse decay exponential curve of the form $H_{\max}(1 - \exp(-\text{angle}/\tau))$ was also found to have a very good fit. Exemplar volume renderings of the Entropy estimation volumes can be seen in Fig. 8.

Asymmetric Fiber Distributions: Further simulations were performed to determine the ability of the conditional Entropy to quantify asymmetry of the distributed fibers. Fibers were simulated with angles $\theta \in [1, 36] \times 5$ and $\phi \in [1, 36] \times 5$ so that θ and ϕ took values between angles 5 and 180 degrees in steps of 5 degrees resulting in $36 \times 36 = 1296$ sets of simulated fibers from which the multiscale Entropy was calculated creating a total of 1296 Entropy data points. Results of calculating the conditional Entropy for these simulated volumes can be seen in Fig. 9. The results of the Entropy estimates on these simulated volumes containing asymmetric fiber distributions show symmetric estimates of Entropy along the line $\theta = \phi$. This demonstrates the Entropy estimation process is consistent whether, e.g. $\theta = 30$ and $\phi = 45$ or $\theta = 45$ and $\phi = 30$. Such asymmetric ranges

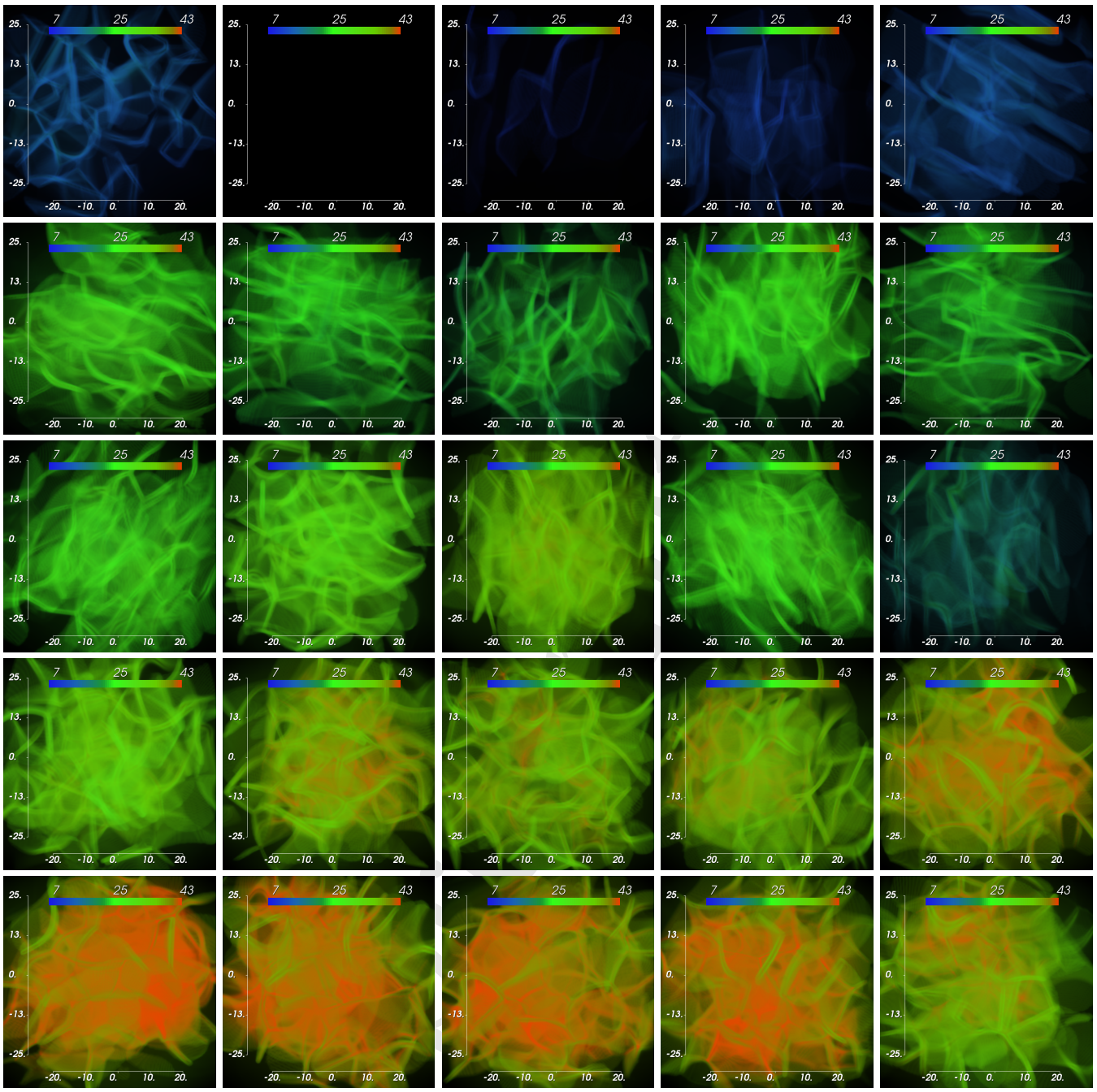


Fig. 8. Volume renderings of single runs of the conditional Entropy estimation process for corresponding data volumes shown in Fig. 6. In total five estimations were performed for each simulated volume which included fiber detection and Entropy estimation processes. Entropy color scales are shown along with scale axes with dimensions in mm. Upper rows show generally lower estimated conditional Entropy values in comparison to lower rows, consistent with the randomness of the fibers in each of the volumes.

of angles in fibers could potentially create ambiguities for techniques that are dependent on a particular reference direction or similar such as Scalar Order Parameter (SOP). Indeed, there is no reference direction, director or similar required in the calculation of the Entropy as described here.

Varying numbers of fibers: It is interesting to observe the performance of the fiber detection process. The number of fibers in a volume was varied from 190730 fibers/m^3 to 1907300/m^3 and fibers were detected using the techniques described here. The results for this process can be seen in Fig. 10. It is interesting to note that the number of fibers in the real SFRC described shortly have a maximum number of 229200 fibers/m^3, which is well under the 1×10^6 fibers/m^3 point at which the number of detected fibers starts to be less than the actual number.

Computation times: The computation times for the RANSAC fiber detection stage and multiscale Entropy estimation stage were computed for simulated volumes consisting of a range of different numbers of fibers. These can be seen in Fig. 11. The

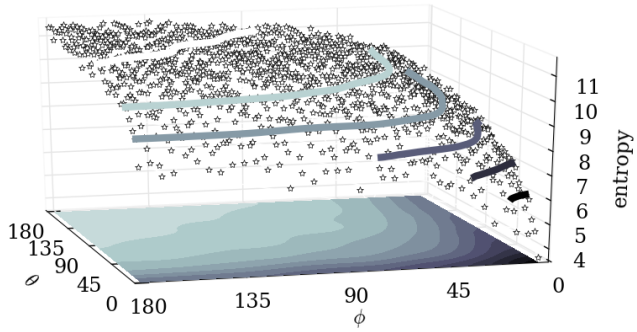


Fig. 9. Results of calculating the conditional Entropy for simulated data volumes with anisotropic variations, where the ranges of possible angles that the fibers were distributed in are $5 \leq \theta \leq 180$ and $5 \leq \phi \leq 180$. The calculated conditional entropies for these anisotropic ranges of angles show a symmetric distribution of Entropy values around the line $\theta = \phi$, indicating a consistent approach to the quantitative summarisation of the orientations of fiber distributions.

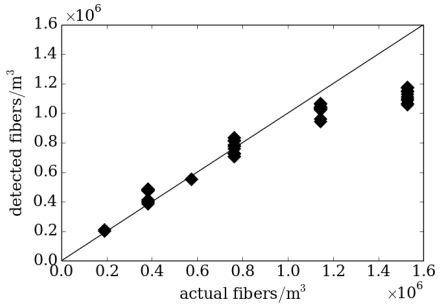


Fig. 10. Results of applying the fiber detection process to 64^3mm^3 data volumes containing a range of different fiber counts (from 190730 fibers/m^3 to 1907300 fibers/m^3). Each fiber is approximately 1mm in diameter and each voxel is 0.5mm^3 in size. The solid line represents the ideal, illustrating the number of fibers are correctly detected up to 1×10^6 fibers/m^3.

computation times for the Entropy estimation stage appear to be somewhat constant whilst the RANSAC fiber detection stage appears to follow a polynomial time curve.

B. User Study of Steel Fiber Reinforced Concrete

Steel Fiber Reinforced Concrete (SFRC) slabs were prepared using $50\mathrm{kg/m}^3$ of steel fibers with varying lengths and diameters and 10mm or 20mm aggregate (stones). Specific details can be seen in Table II. The SFRC was made into reinforced cuboid slabs, each of size $600^2 \times 100\mathrm{mm}^3$ with $\times 2$ cylindrical samples for each slab. The sample core volumes were $\pi 50^2 \times 100\mathrm{mm}^3$ (from volume of a cylinder). Each of these core samples were then imaged with XCT acquired using the aforementioned XCT parameters. Core sample voxel sizes were $0.01652^2 \times 1.0\mathrm{mm}^3$ with slice thickness $0.4\mathrm{mm}$. The conditional Entropy that combines elements of the distance and orientation of fibers, on a per voxel level were calculated for these XCT imaging data acquired from the steel fiber concrete cores with various properties.

The results can be seen in Figs. 12 and 13(a). The results in Fig. 12 appear to show greater fiber diameter results in increased conditional Entropy. The results in Fig. 13 included some comparisons with the multiscale Entropy that includes just orientation

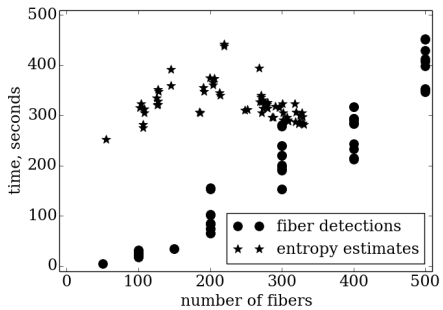


Fig. 11. Computation times for the RANSAC fiber detection stage and multiscale Entropy estimation stage for volumes consisting of a range of different numbers of fibers.

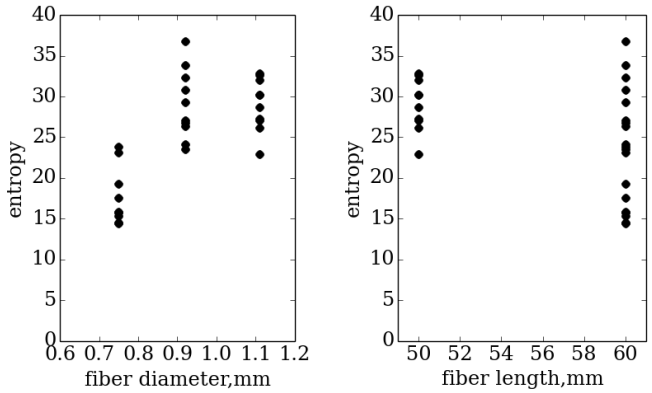


Fig. 12. Results of applying Entropy calculations to real steel fiber reinforced concrete where the diameter and length of the fibers for the different cores are included. These results appear to show that greater Entropy is seen for cores with thicker fibers (greater diameters).

TABLE II
PROPERTIES OF THE REAL STEEL FIBER CONCRETE CORES. RESULTS OF APPLYING THE ENTROPY CALCULATIONS TO XCT SCANS OF THESE CORES CAN BE SEEN IN FIGS. 12 AND 13.

length mm	**diameter** mm	**fibers** /m^3	**aggregate size** mm^3	**slab count**	**Peak** kN **Load**
50	1.11	140100	10	3	93
60	0.92	159150	10	3	106
60	0.75	229200	10	3	94
50	1.11	140100	20	3	68
60	0.92	159150	20	3	70
60	0.75	229200	20	3	82

information $H_{\mathrm{ms}}(q,r)$ in (2); Scalar Order Parameter (SOP) as used by Hermann et al. in [7], [8] $S \in \left[-\frac{1}{2}, 1\right]$; and efficiency factors e_{x}, e_{y} and e_{z} as used by Vicente et al. [2] and Abrishambaf et al. [3].

The comparative results in Fig. 13 appear to show a greater, consistent distinction between the 10mm aggregate core versus the 20mm aggregate core, than for the SOP or the efficiency factors. This can be quantified with an overlap calculation

$$\mathcal{O} = \left| \sum_{\forall \mathrm{cores}} \frac{(X - Y)}{\max(X, Y)} \right| \tag{12}$$

with corresponding values of (12) being shown in Table III. The calculated overlaps also show that the entropy based measures demonstrate greater difference between the 10mm and 20mm based aggregate cores.

User study perspective: From the user study point of view, these results are potentially very useful. Quantitative comparisons in the form of scatter plots in Fig. 13 show that the H_{vol} values possess a marked distinction between the different aggregate types and what appears to be a higher correlation with the peak load performance of the material in the different cores. Peak load values in kN can be seen in Table II, details of which were described by Ige et al. [48]. The SOP and efficiency factors results did not differentiate between aggregate types. The SOP values appear to have some correlation with the number of fibers, but not the underlying matrix that may otherwise affect the general strength properties of the material.

The 10mm aggregate cores all have fibers with lengths that are over three times the size of the aggregate size. As noted by Vandewalle [50], this is important because it enables the fibers to bridge any potential gaps and to provide sufficient bonding in the concrete matrix. However this is not the case for a few of the 20mm aggregate cores, where the 50mm length fibers (1.11mm diameter) fall below this length to aggregate size threshold. Another consideration is regarding the diameter of the steel fibers. A smaller diameter potentially increases the number of fibers for a given weight and therefore the potential for fibers to be distributed more densely across the matrix thereby improving the mechanical properties of the materials. The influence of these factors can be seen in the load bearing capacity of concrete panels as described by Ige et al. [48] where it can be seen that, for the same cores imaged here, the 10mm aggregate SFRC panels consistently possessed greater maximum

TABLE III
OVERLAP CALCULATIONS FROM (12) FOR THE DIFFERENT ENTROPY, SOP AND EFFICIENCY FACTORS RESULTS FOR THE REAL CORES; SHOWN IN FIG. 13.

	H_{vol}	S	H_{ms}	e_{x}	e_{y}	e_{z}
$\mathcal{O}$	2.43	0.78	1.21	0.50	0.42	0.57

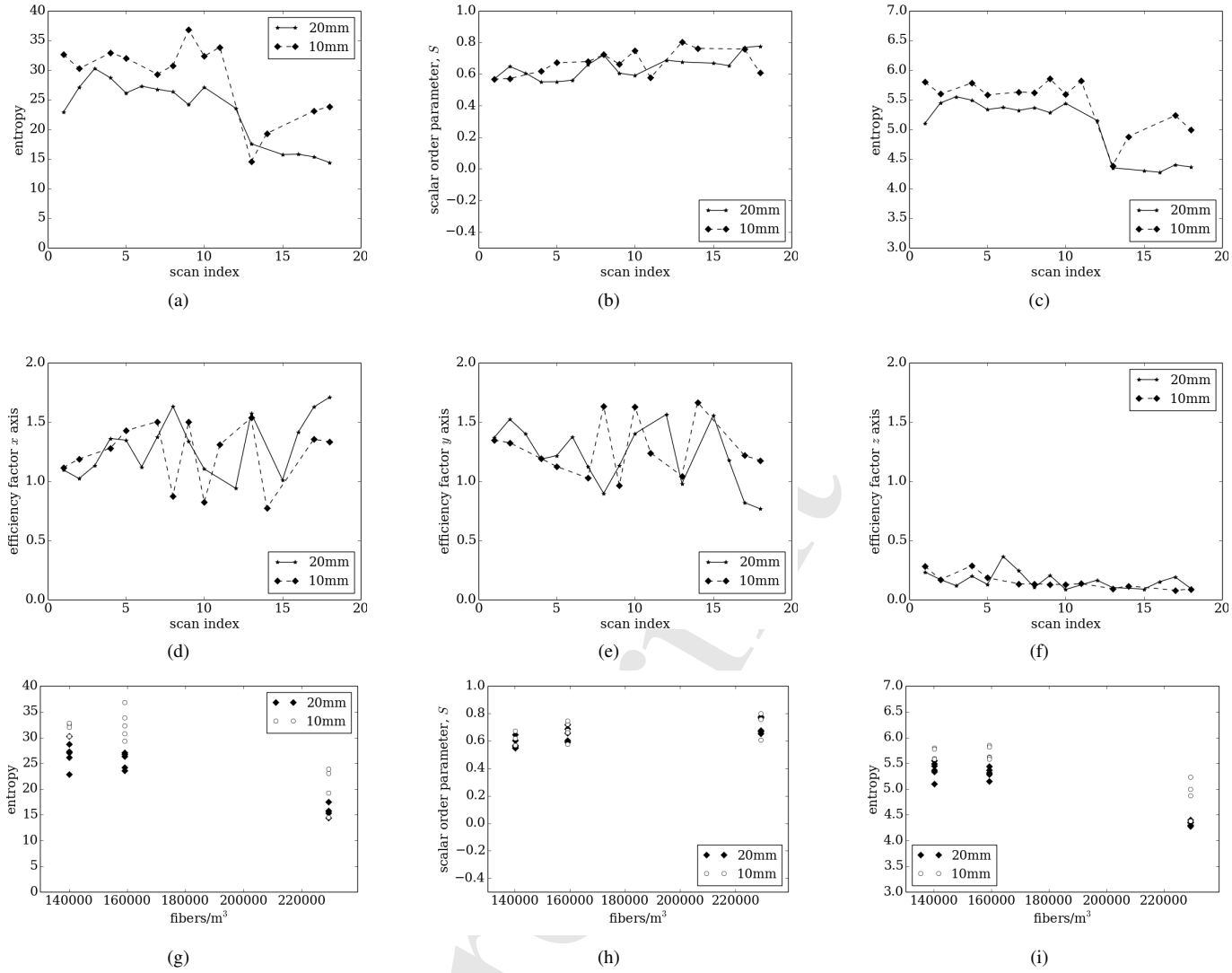


Fig. 13. Results of calculations applied to real steel fiber reinforced concrete. (a) shows conditional Entropy results using probabilities which include the distance between fibers; (b) are the SOP results; and (c) shows Entropy results using just fiber orientations but no distance information; (d-f) show the efficiency factors along each dimension of the data; (g-i) show the results of (a-c) but plotted as a function of fibers/m^3; The entropy based results (a,c) appear to consistently show that the 10mm based aggregate fiber concrete cores are more randomly distributed, however the Entropy calculations also appear to identify greater relative differences between the 20mm and 10mm aggregate cores in comparison to SOP values and the efficiency factors, also quantified in Table III.

load bearing capacity in comparison to the 20mm aggregate SFRC panels. This was also reflected well with the H_{vol} results as shown in Fig. 13.

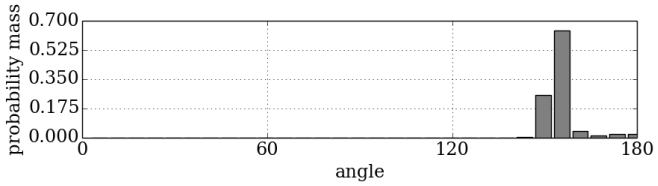
C. Entropy Comparison

The overlap results shown in Table III and in Fig. 13 indicate differences between Entropy based descriptions of randomness and variance based descriptions of fiber distributions (e.g. SOP). So a question we now ask is, *what properties of a distribution are each of the randomness descriptors capturing?*

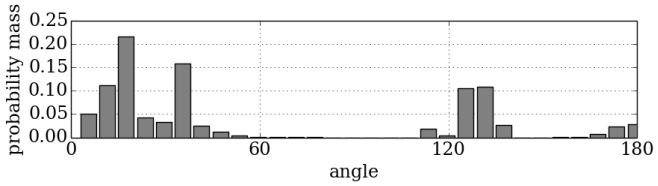
Some further simulations were undertaken in the form of multiple randomly sampled multi-modal mixture models, to help simulate the kind of scenarios that might be encountered in real world data. Some example randomly generated mixture models can be seen in Fig. 14. The mixture model takes the form

$$P(\theta) = \sum_{i=1}^{Q} \frac{w_i}{2\pi\sigma_i^2} \exp\left(-\frac{(\theta-\mu_i)^2}{2\sigma_i^2}\right) \tag{13}$$

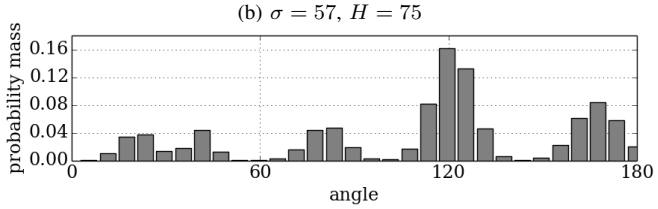
where the number of components Q, the means and standard deviations of each component μ_i and σ_i respectively and the priors w_i, were randomly sampled with the constraint of $\sum_i^N w_i = 1$. This enabled a large range of distributions (10000)



(a) $\sigma = 6$, $H = 36$



(b) $\sigma = 57$, $H = 75$



(c) $\sigma = 45$, $H = 88$

Fig. 14. Comparison of three randomly generated Probability Mass Functions (PMF). For (c) the PMF has the highest Entropy, but a lower standard deviation. It has a wide range of values and fewer gaps, indicating that a fiber could occur in any direction, rather than a select few, as would be the case for (b).

TABLE IV

CROSS-CORRELATION COEFFICIENT CALCULATIONS FOR ENTROPY CALCULATED WITH SINGLE SCALES H_{XX}; MULTIPLE SCALES (MS), H_{MS}; AND WITH THE STANDARD DEVIATION (σ_{var}) . ALSO INDICATED ARE, ACROSS ALL SCALES, THE MAXIMUM FOR EACH OF THE COEFFICIENTS IN BOLD AND THE MINIMUMS IN ITALICS. THESE RESULTS ARE ALSO SUMMARIZED IN FIG. 15. THESE RESULTS INDICATE THAT H_{MS} IS HIGHLY CORRELATED WITH A NUMBER OF THE TERMS IN THE MIXTURE BASED SAMPLING DISTRIBUTION GIVEN BY (13).

scale	Q	$\max_i(w_i)$	$\max_i(\mu_i) - \min_i(\mu_i)$	$\mathbf{E}[\sigma_i]$	$\max_i(\sigma_i)$
2	*0.43*	*0.49*	0.61	*0.04*	*0.24*
7	0.73	0.77	**0.67**	0.16	0.46
12	**0.76**	**0.79**	0.64	0.25	0.53
17	0.75	0.79	0.61	0.33	0.57
22	0.74	0.77	0.59	0.37	0.59
27	0.73	0.76	0.57	0.41	**0.60**
32	0.71	0.75	0.56	0.43	0.60
37	0.70	0.73	0.54	0.45	0.60
42	0.69	0.72	0.53	0.46	0.60
47	0.68	0.71	0.52	0.47	0.60
52	0.67	0.71	0.52	0.48	0.60
57	0.66	0.70	0.51	0.49	0.60
62	0.65	0.69	0.50	0.50	0.59
67	0.65	0.69	0.50	0.50	0.59
72	0.64	0.68	0.49	0.50	0.59
77	0.64	0.67	0.48	0.50	0.59
82	0.63	0.67	0.48	**0.51**	0.59
87	0.62	0.66	*0.47*	0.51	0.58
92	0.62	0.66	0.47	0.51	0.58
97	0.62	0.65	0.47	0.51	0.58
MS	0.69	0.73	0.55	0.45	0.59
σ_{var}	0.40	0.45	0.85	0.03	0.21

to be tested in terms of three randomness descriptors: multiple scale Entropy H_{ms}, standard deviation σ_{var} and single scale Entropy H_{XX} calculated at a number of scales. For each randomly generated distribution a number of measurements were made. The correlation coefficient was then calculated between the measurements and the randomness descriptors to determine what aspects of the generated distributions that the randomness descriptors capture. The measurements were: the number of mixture components Q; the maximum of the priors $\max_i(w_i)$; the difference between the maximum mean and the minimum mean $\max_i(\mu_i) - \min_i(\mu_i)$; the mean of the standard deviations $\mathbf{E}[\sigma_i]$; and the maximum of the standard deviations $\max_i(\sigma_i)$.

Table IV has the cross-correlation coefficients for the single scale entropies H_{XX} and also for the multiple scale Entropy H_{ms} and standard deviation σ_{var}. The results shown in Table IV can also be seen, summarized, in Fig. 15.

These results show that the standard deviation of the angles has the highest cross correlation 0.85 with the difference between

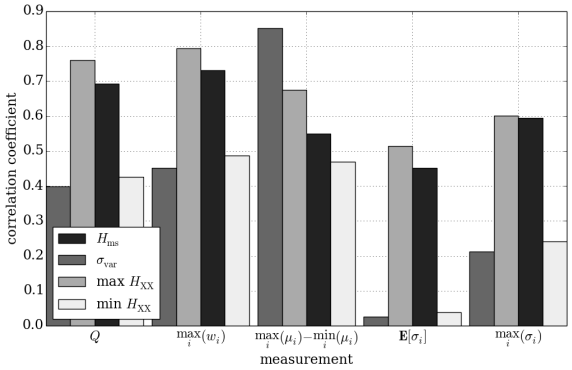


Fig. 15. Visual comparison of summaries of the cross-correlation coefficients shown in Table IV for parameters of the mixture based sampling distribution in (13). The highest cross-correlation is achieved with the standard deviation in relation to the maximum difference between means (0.85). In other cases the maximum of the single scales Entropy H_{XX} achieves the highest correlation with, e.g. the number of mixtures Q; the mixture with the greatest weight w_i; and even the greatest standard deviation $\max_i(\sigma_i)$. The multiscale Entropy $H_{\rm ms}$ is also highly correlated with these variables and is not dependent on a particular scale which makes it the preferred choice for the majority of cases.

the maximum mean and the minimum mean for a set of randomly generated components, i.e. $\max_i(\mu_i) - \min(\mu_i)$. This is reflected in the examples shown in Fig. 14. The single scale Entropy cross correlation values can take on a range of different values for each measurement, demonstrating (undesirable) scale dependency and the need for a multi-scale formulation. The multiple scale Entropy, however demonstrates higher cross correlation coefficient values for all other measurements (other than the difference between the means) completely independent of scale. This means that it can be computed without consideration of a particular scale. In particular, it is interesting to see the high correlation with the number of components present and the maximum prior, $\max_i(w_i)$. This latter measurement is a strong indicator of whether a distribution is uniform and hence the amount of randomness in a sample. This is also reflected in the examples in Fig. 14.

IV. Discussion and Conclusions

This work has included an extensive range of simulations of fiber distributions, a theoretical model and an investigation into the unique benefits of using multiscale Entropy to summarize the orientation distribution and the spatial distribution of fibers combined into a single model. The work is applied here to steel fibers for the reinforcement of concrete. The summarisation of the orientation distribution in this way appears to show important applicability in this field. This is because it helps to more accurately identify differences in the underlying properties of the material which is useful for mechanical applications and others.

It is also anticipated that a model of this nature could potentially be applicable to a wide range of other areas, principally where summarisation of the orientation and spatial distribution of fiber like materials is needed. This could include electrospun nano-fibers that have been used in tissue scaffolds for bone, see Stachewicz et al. [51], also for ligaments as indicated by Pauly et al. [52]. There are other fiber based materials including polymer based fibers as looked at by Salaberger et al. [41] and paper as shown by Axelsson and Svensson [6]. The work presented here may have some application to these other areas however some consideration will be needed with regards to applications at a smaller scale. For example, XCT of collagen fibers can present some difficulties, partly due to the smaller scale which can require additional modeling of partial volume effects or similar. Another consideration could potentially include the difficulty of accurately segmenting fibers that possess similar appearance to the surrounding materials in the imaging data which is not the case for steel in XCT. Also, the assumption of straight fibers may also need to be reconsidered for other applications.

Another aspect not considered here is the mechanical properties of the materials such as elasticity, which is essential when developing a model that can provide insight into how a particular material forms. An example of this is the formation of bundles of fibers which has been looked at by e.g. Hall et al. [53]. Hall et al. combined a variety of different mechanical properties such as cohesion and elasticity of the materials. On the other hand, an information theoretic model of a material could potentially include elements of this type of theory which is the topic of on-going research.

Acknowledgments

The authors would like to thank Chris Fox, Faculty of Engineering, University of Nottingham for acquiring the real microtomography data. We would also like to thank Dr David Gonzalez and Professor Asa Barber, School of Engineering, University of Portsmouth for helpful discussions and the anonymous reviewers for their helpful comments.

APPENDIX A
PROOF OF APPROXIMATE LOGARITHMIC MULTISCALE ENTROPY

Theorem. *For a volume consisting of fibres isotropically distributed in the range $\theta \in [0, q] \times 180$ and $\phi \in [0, r] \times 180$ where $q \in [0, 1]$ and $r \in [0, 1]$, the Multiscale Entropy $H_{\mathrm{ms}}(q, r)$ across a set of scales $M \in [1, \mathfrak{N}]$ and $N \in [1, \mathfrak{M}]$, each with $M \times N$ bins, is approximately $\log_2(qr\mathfrak{M}\mathfrak{N}) + \frac{2}{\ln(2)}$.*

Proof. At scale (M, N), which has $M \times N$ bins, the probabilities $P_{MN}(m, n, q, r)$ obtained from a volume consisting of fibres distributed in the range $[0, q] \times [0, r]$ for bin (m, n) where $0 \leq m < M$ and $0 \leq n < N$, are given by four possibilities. In two of the four possibilities, the range of the fibre distribution $[0, q]$ and $[0, r]$ are smaller than the bin size at a particular scale, i.e. $q \leq \frac{1}{M}$ and $r \leq \frac{1}{N}$ so that

$$\text{if } q \leq \frac{1}{M} \text{ and } r \leq \frac{1}{N} \text{ then}$$
$$P_{MN}(m, n, q, r) = \begin{cases} 1 & \text{for} \quad m, n = 0; \\ 0 & \text{elsewhere.} \end{cases} \tag{14}$$

The other two conditions occur when the fibre distribution is spread across multiple bins, i.e. $0 \leq m \leq \lceil qM \rceil$ and $0 \leq n \leq \lceil rN \rceil$, assuming (an approximation) the spread completely fills each of the bins with $\frac{1}{qM} \times \frac{1}{rN}$ as each bin is of size $\frac{1}{M} \times \frac{1}{N}$. Therefore

$$\text{if } q > \frac{1}{M} \text{ and } r > \frac{1}{N} \text{ then} \tag{15}$$
$$P_{MN}(m, n, q, r) \approx \begin{cases} \frac{1}{rN} \times \frac{1}{qM} & \text{for} \quad 0 \leq m \leq \lceil qM \rceil \\ & \text{and} \quad 0 \leq n \leq \lceil rN \rceil; \\ 0 & \text{elsewhere.} \end{cases}$$

For these four cases, abbreviating the notation, the product of the probability P_{MN} with the self information h_{MN} can be determined in the limit. For the first two cases ($q \leq 1/M$ & $r \leq 1/N$), we have $P_{MN}h_{MN} = 0$ as the log function grows more slowly in comparison to the probability, closer to zero. Similarly, for the other two cases we have ($q > 1/M$ & $r > 1/N$):

$$P_{MN}h_{MN} \approx \begin{cases} \frac{1}{qMrN} \log_2(qMrN) & \text{for} \quad 0 \leq m \leq \lceil qM \rceil \\ & \text{and} \quad 0 \leq n \leq \lceil rN \rceil; \\ 0 & \text{elsewhere.} \end{cases} \tag{16}$$

The entropy for a single scale (M, N) can then be determined (where $q > 1/M$ & $r > 1/N$, as otherwise it is equal to zero)

$$H_{MN}(q, r) = \sum_{m=1}^{M} \sum_{n=1}^{N} P_{MN}h_{MN} \tag{17}$$

Substituting in the values for the product term from (16):

$$\begin{aligned} H_{MN}(q, r) &\approx \sum_{m=1}^{\lceil qM \rceil} \sum_{n=1}^{\lceil rN \rceil} \frac{1}{qMrN} \log_2(qMrN) + 0 \\ &\approx \frac{\lceil qMrN \rceil}{qMrN} \log_2(qMrN) \approx \log_2(qMrN). \end{aligned} \tag{18}$$

This is the approximate entropy at a particular scale (M, N). The multiple scale entropy is then calculated with

$$\begin{aligned} H_{\mathrm{ms}}(q, r) &= \frac{1}{|\forall M|} \sum_{\forall M} \frac{1}{|\forall N|} \sum_{\forall N} H_{MN}(q, r) \\ &\approx \frac{1}{|\forall M||\forall N|} \sum_{\forall M} \sum_{\forall N} \log_2(qMrN). \end{aligned} \tag{19}$$

Assuming the scales are linearly progressed through the natural numbers to maximums $\mathfrak{M}$ and $\mathfrak{N}$ for scales M and N respectively, we have

$$\begin{aligned} H_{\mathrm{ms}}(q, r) &\approx \frac{1}{\mathfrak{M}\mathfrak{N}} \sum_{M=1}^{\mathfrak{M}} \sum_{N=1}^{\mathfrak{N}} \log_2(qMrN) \\ &\approx \log_2(qr) + \frac{1}{\mathfrak{M}\mathfrak{N}} \sum_{M=1}^{\mathfrak{M}} \sum_{N=1}^{\mathfrak{N}} \log_2(MN). \end{aligned} \tag{20}$$

The right most term can be expanded into factorials

$$H_{\mathrm{ms}}(q,r) \approx \log_2(qr) + \frac{1}{\mathfrak{M}\mathfrak{N}}\left\{\mathfrak{N}\log_2(M!) + \mathfrak{M}\log_2(N!)\right\}. \tag{21}$$

Stirling's approximation, i.e. $\ln(x!) \approx x\ln(x) - x$ can be used

$$H_{\mathrm{ms}}(q,r) \approx \log_2(qr) + \log_2(\mathfrak{M}) + \log_2(\mathfrak{N}) - \frac{2}{\ln(2)} \tag{22}$$

resulting in

$$H_{\mathrm{ms}}(q,r) \approx \log_2(rq\mathfrak{M}\mathfrak{N}) - \frac{2}{\ln(2)}. \tag{23}$$

□

Corollary. *The variance of the mean multiscale entropy estimator can be approximated by*

$$\sigma^2_{\mathrm{ms}} \approx 4\log_2^2(\mathfrak{N}\mathfrak{M}) - \log_2(\mathfrak{N}\mathfrak{M})(4 + 2/\ln(2)) + 4/\ln^2(2). \tag{24}$$

Proof. The variance of the estimator at a particular scale (M, N) can be defined as

$$\sigma^2_{MN} = \sum_{m=1}^{M}\sum_{n=1}^{N} P_{MN} \times \left(h_{MN} - H_{\mathrm{ms}}\right)^2. \tag{25}$$

Using a similar argument as was used for (16) and assuming $P_{MN} = \frac{1}{qMrN}$, results in

$$\begin{aligned}
\sigma^2_{MN} \approx & \sum_{m=1}^{\lceil qM\rceil}\sum_{n=1}^{\lceil rN\rceil} \frac{1}{qMrN} \times \\
& \left\{\log_2(qMrN) - \log_2(q\mathcal{M}r\mathcal{N}) + c\right\}^2 \\
= & \sum_{m=1}^{\lceil qM\rceil}\sum_{n=1}^{\lceil rN\rceil} \frac{1}{qMrN}\left\{\log_2\left(\frac{MN}{\mathcal{M}\mathcal{N}}\right) + c\right\}^2 \\
= & \frac{\lceil qMrN\rceil}{qMrN}\left\{\log_2\left(\frac{MN}{\mathcal{M}\mathcal{N}}\right) + c\right\}^2 \\
\approx & \left\{\log_2\left(\frac{MN}{\mathcal{M}\mathcal{N}}\right) + c\right\}^2.
\end{aligned} \tag{26}$$

where $c = 2/\ln(2)$. Assuming, as before, that the scales are linearly progressed to maximums $\mathcal{M}$ and $\mathcal{N}$ then the mean variance across all these scales can be described with

$$\begin{aligned}
\sigma^2_{\mathrm{ms}} \approx & \frac{1}{\mathcal{M}\mathcal{N}}\sum_{M=1}^{\mathcal{M}}\sum_{N=1}^{\mathcal{N}} \sigma^2_{MN} \\
\approx & \frac{1}{\mathcal{M}\mathcal{N}}\sum_{M=1}^{\mathcal{M}}\sum_{N=1}^{\mathcal{N}}\left\{\log_2\left(\frac{MN}{\mathcal{M}\mathcal{N}}\right) + c\right\}^2 \\
\approx & 3\log_2^2(\mathcal{M}\mathcal{N}) - 2c\log_2(\mathcal{M}\mathcal{N}) - c^2 \\
& + \frac{1}{\mathcal{M}\mathcal{N}}\sum_{M=1}^{\mathcal{M}}\sum_{N=1}^{\mathcal{N}}\log_2^2(MN)
\end{aligned} \tag{27}$$

The remaining summations are then approximated using the Euler-Maclaurin formula, ignoring the error terms results in

$$\sigma^2_{\mathrm{ms}} = 4\log_2^2(\mathfrak{N}\mathfrak{M}) - \log_2(\mathfrak{N}\mathfrak{M})(4 + c) + c^2. \tag{28}$$

□

REFERENCES

[1] A. Abrishambaf, V. Cunha, and J. Barros, "The influence of fibre orientation on the post-cracking tensile behaviour of steel fibre reinforced self-compacting concrete," *Frattura ed Integrita Strutturale, Fracture and Structural Integrity*, vol. 31, pp. 38–53, January 2015.

[2] M. Vicente, D. González, and J. Mínguez, "Determination of dominant fibre orientations in fibre-reinforced high-strength concrete elements based on compouted tomography scans," *Nondestructive Testing and Evaluation*, vol. 29, no. 2, pp. 164–182, 2014.

[3] A. Abrishambaf, J. Barros, and V. Cunha, "Relation between fibre distribution and post-cracking behaviour in steel fibre reinforced self-compacting concrete panels," *Cement and Concrete Research*, vol. 51, pp. 57–66, 2013.

[4] J. Weissenöck, A. Amirkhanov, W. Li, A. Reh, A. Amirkhanov, E. Gröller, J. Kastner, and C. Heinzl, "FiberScout: An interactive tool for exploring and analyzing fiber reinforced polymers," in *Pacific Visualization Symposium*. Yokohama: IEEE, 2014, pp. 153–160.

[5] M. Axelsson, "Estimating 3D fibre orientation in volume images," in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, Dec 2008, pp. 1–4.

[6] M. Axelsson and S. Svensson, "3D pore structure characterisation of paper," *Pattern Anal. Applic.*, vol. 13, pp. 159–172, 2010.

[7] H. Hermann and M. Eik, "Some comments on the theory of short fibre reinforced materials," *Proceedings of the Estonian Academy of Sciences*, vol. 60, no. 3, pp. 179–183, 2011.

[8] H. Hermann, E. Pastorelli, A. Kallonen, and J.-P. Suuronen, "Methods for fibre orientation analysis of X-ray tomography images of steel fibre reinforced concrete (sfrc)," *J. Mater. Sci.*, vol. 51, pp. 3772–3783, 2016.

[9] T. Jankun-Kelly and K. Mehta, "Superellipsoid-based, real symmetric traceless tensor glyphs motivated by nematic liquid crystal alignment visualization," *IEEE Trans. Vis. Comp. Graph.*, vol. 12, no. 5, pp. 1197–1204, Sept./Oct. 2006.

[10] N. Mottram and C. Newton, "Introduction to Q-tensor theory," *ArXiv e-prints*, Sep. 2014.

[11] P. Juan-García, J. Torrents, R. Lòpez-Carreño, and S. Pialarissi-Cavalaro, "Influence of fiber properties on the inductive method for the steel-fiber-reinforced concrete characterization," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 8, pp. 1937–1944, Aug. 2016.

[12] C.-F. Westin, S. Maier, H. Mamata, A. Nabavi, F. Jolesz, and R. Kikinis, "Processing and visualization for diffusion tensor MRI," *Medical Image Analysis*, vol. 6, pp. 93–108, 2002.

[13] C.-F. Westin and H. Knutsson, "Extraction of local symmetries using tensor field filtering," in *2nd Singapore Int. Conf. Image Processing*. IEEE, 1992.

[14] ——, "Estimation of motion vector fields using tensor field filtering," in *Int. Conf. Image Processing*. IEEE, 1994, pp. 237–242.

[15] X. Jiang, "On orientation and anisotropy estimation for online fingerprint authentication," *IEEE Trans. Signal Proc.*, vol. 53, no. 10, pp. 4038–4049, Oct. 2005.

[16] J. Bigun and G. Granlund, "Optimal orientation detection of linear symmetry," in *1st Int. Conf. Comp. Vision (ICCV)*. London: IEEE, 1987, pp. 433–438.

[17] H. Knutsson, "Representing local structure using tensors," in *Proc. Scandinavian Conf. Image Analysis*, Oulu, Oulu University, 1989, pp. 244–251.

[18] M. Liu and P.-T. Yap, "Invariant representation of orientation fields for fingerprint indexing," *Pattern Recognition*, vol. 45, pp. 2532–2542, 2012.

[19] C. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[20] R. Sharma, R. Pachori, and U. Acharya, "Application of entropy measures on intrinsic mode functions for the automated identification of focal electroencephalogram signals," *Entropy*, vol. 17, pp. 669–691, 2015.

[21] A. Renyi, "On measures of information and entropy," in *Proc. 4th Berkeley Symp. Math. Stat. Prob.*, 1960, pp. 547–561.

[22] S. Pincus, "Approximate entropy as a measure of system complexity," *Proc. Natl. Acad. Sci. USA*, vol. 88, pp. 2297–2301, Mar. 1991.

[23] J. Richman and J. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *Am. J. Physiol. Heart Circ. Physiol.*, vol. 278, pp. H2039–H2049, 2000.

[24] K. Ho, G. Moody, C. Peng, J. Mietus, M. Larson, D. Levy, and A. Goldberger, "Predicting survival in heart failure case and control subjects by use of fully automated methods for deriving nonlinear and conventional indices of heart rate dynamics," *Circulation*, vol. 96, no. 3, pp. 842–848, Aug. 1997.

[25] M. Costa, A. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of complex physiologic time series," *Phys. Rev. Lett.*, vol. 89, no. 6, pp. 68 102–68 102, 2002.

[26] S. Ramdani, V. Bonnet, G. Tallon, J. Lagarde, P. Bernard, and H. Blain, "Parameters selection for bivariate multiscale entropy analysis of postural fluctuations in fallers and non-fallers old adults," *IEEE Trans. Neural Systems Rehab. Eng.*, vol. 24, no. 8, pp. 859–871, Aug. 2016.

[27] P. Fino, A. Mojdehi, K. Adjerid, M. Habibi, T. Lockhart, and S. Ross, "Comparing postural stability entropy analyses to differentiate fallers and non-fallers," *Annals Biomed. Eng.*, vol. 44, no. 5, pp. 1636–1645, Oct. 2016.

[28] S. Mariani, A. Borges, T. Henriques, A. Goldberger, and M. Costa, "Use of multiscale entropy to facilitate artifact detection in electroencephalographic signals," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 2015, pp. 7869–7872.

[29] Z.-K. Gao, M.-S. Ding, H. Geng, and N.-D. Jin, "Multivariate multiscale entropy analysis of horizontal oil-water two-phase flow," *Physica A*, vol. 417, pp. 7–17, 2015.

[30] J. Valencia, A. Porta, M. Vallverdu, F. Claria, R. Baranowski, W. Orlowska-Baranowska, and P. Caminal, "Refined multiscale entropy: Application to 24-h holter recordings of heart period variability in healthy and aortic stenosis subjects," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 9, pp. 2202–2213, 2009.

[31] M. Costa and A. Goldberger, "Generalized multiscale entropy analysis: Application to quantifying the complex volatility of human heartbeat time series," *Entropy*, vol. 17, pp. 1197–1203, 2015.

[32] Y.-C. Zhang, "Complexity and 1/f noise. a phase space approach," *J. Phys. I France*, vol. 1, no. 7, pp. 971–977, July 1991, http://dx.doi.org/10.1051/jp1:1991180.

[33] H. Fogedby, "On the phase space approach to complexity," *J. Stat. Physics*, vol. 69, no. 1, pp. 411–425, 1992, doi:10.1007/BF01053799.

[34] A. Abbas, S. Mohsin, and D. Cotsovos, "A simplified finite element model for assessing steel fibre reinforced concrete structural performance," *Computers and Structures*, vol. 173, pp. 31–49, Sept. 2016.

[35] Y. Zhang, X. Wang, N. Pan, and R. Postle, "Weibull analysis of the tensile behavior of fibres with geometrical irregularities," *J. Mat. Sci.*, vol. 37, pp. 1401–1406, 2001.

[36] ACI-Committee-544, "State-of-the-art report on fiber reinforced concrete," *Journal Proc. American Concrete Institute*, vol. 70, no. 11, pp. 729–744, 2002.

[37] S. Abdallah, M. Fan, X. Zhou, and S. Le Geyt, "Achorage effects of various steel fibre architectures for concrete reinforcement," *Int. J. Concrete Structures and Materials*, vol. 10, no. 3, pp. 325–335, Sept. 2016.

[38] E. Weisstein, "Point-line distance–3-dimensional." From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/Point-LineDistance3-Dimensional.html, retrieved May 2017.

[39] L. Fritz, M. Hadwiger, G. Geier, G. Pittino, and M. Gröller, "A visual approach to efficient analysis and quantification of ductile iron and reinforced sprayed concrete," *IEEE Trans. Vis. Computer Graphics*, vol. 15, no. 6, pp. 1343–1350, Nov./Dec. 2009.

[40] C. Eberhardt and A. Clarke, "Automated reconstruction of curvilinear fibres from 3D datasets acquired by X-ray microtomography," *Journal of Microscopy*, vol. 206, no. 1, pp. 41–53, 2002.

[41] D. Salaberger, K. Kannappan, J. Kastner, J. Reussner, and T. Auinger, "Evaluation of computed tomography data from fibre reinforced polymers," *Int. Polymer Processing*, vol. 26, no. 3, pp. 283–291, 2011.

[42] M. Tausif, B. Duffy, H. Carr, S. Grishanov, and S. Russell, "Three-dimensional fibre segment orientation distribution using X-ray microtomography," *Microscopy and Microanalysis*, vol. 20, no. 4, pp. 1294–1303, 2014.

[43] M. Fischler and R. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, June 1981.
[44] R. Gonzalez and R. Woods, *Digital Image Processing*. Addison Wesley, 1993.
[45] E. Davies, *Computer and machine vision*, 4th ed. Academic Press, 2012.
[46] A. Antos and I. Kontoyiannis, "Convergence properties of functional estimates for discrete distributions," *Random structures and algorithms*, vol. 19, pp. 163–193, 2001.
[47] L. Paninski, "Estimation of entropy and mutual information," *Neural Computation*, vol. 15, pp. 1191–1253, 2003.
[48] O. Ige, S. Barnett, J. Chiverton, A. Nassif, and J. Williams, "Effects of steel fibre-aggregate interaction on mechanical behaviour of steel fibre reinforced concrete," *Advances in Applied Ceramics: Structural, Functional and Bioceramics*, vol. 116, no. 4, pp. 193–198, 2017.
[49] J. Chiverton, "fibsim: fibre simulation and characterization," https://github.com/chivertj, 2016.
[50] M. Vandewalle, "The use of steel fibre reinforced shotcrete for the support of mine openings," *The Journal of the South African Institute of Mining and Metallurgy*, pp. 113–120, May/June 1998.
[51] U. Stachewicz, T. Qiao, S. Rawlinson, F. Almeida, W. Li, M. Cattell, and A. Barber, "3D imaging of cell interactions with electrospun PLGA nanofiber membranes for bone regeneration," *Acta Biomater.*, vol. 27, pp. 88–100, Nov. 2015.
[52] H. Pauly, D. Kelly, K. Popat, N. Trujillo, N. Dunne, H. McCarthy, and T. Haut-Donahue, "Mechanical properties and cellular response of novel electrospun nanofibers for ligament tissue engineering: Effects of orientation and geometry," *J. Mech. Behavior Biomed. Materials*, vol. 61, pp. 258–270, Aug. 2016.
[53] D. Hall, I. Bruss, J. Barone, and G. Grason, "Morphology selection via geometric frustration in chiral filament bundles," *Nature Materials*, vol. 15, pp. 727–732, March 2016.