

## Estimating the 3D Pore Size Distribution of Biopolymer Networks from Directionally Biased Data

Nadine R. Lang,<sup>†</sup> Stefan Münster,<sup>†</sup> Claus Metzner,<sup>†</sup> Patrick Krauss,<sup>†</sup> Sebastian Schürmann,<sup>‡</sup> Janina Lange,<sup>†</sup> Katerina E. Aifantis,<sup>§</sup> Oliver Friedrich,<sup>‡</sup> and Ben Fabry<sup>†\*</sup>

<sup>†</sup>Department of Physics and <sup>‡</sup>Institute of Medical Biotechnology, University of Erlangen-Nuremberg, Erlangen, Germany; and <sup>§</sup>Laboratory of Mechanics and Materials, Aristotle University of Thessaloniki, Thessaloniki, Greece

**ABSTRACT** The pore size of biopolymer networks governs their mechanical properties and strongly impacts the behavior of embedded cells. Confocal reflection microscopy and second harmonic generation microscopy are widely used to image biopolymer networks; however, both techniques fail to resolve vertically oriented fibers. Here, we describe how such directionally biased data can be used to estimate the network pore size. We first determine the distribution of distances from random points in the fluid phase to the nearest fiber. This distribution follows a Rayleigh distribution, regardless of isotropy and data bias, and is fully described by a single parameter—the characteristic pore size of the network. The bias of the pore size estimate due to the missing fibers can be corrected by multiplication with the square root of the visible network fraction. We experimentally verify the validity of this approach by comparing our estimates with data obtained using confocal fluorescence microscopy, which represents the full structure of the network. As an important application, we investigate the pore size dependence of collagen and fibrin networks on protein concentration. We find that the pore size decreases with the square root of the concentration, consistent with a total fiber length that scales linearly with concentration.

### INTRODUCTION

The mesh size of the extracellular matrix (ECM) is an important parameter that governs its mechanical properties and influences the ability of cells to colonize and migrate through the ECM (1–3). Artificial three-dimensional extracellular matrices from self-assembled protein networks are widely used for tissue-engineering applications and for studying cell behavior in an environment that more closely resembles the *in vivo* physiological situation of mammalian cells (4,5). Knowing the exact pore size is crucial, because the ability of cells to migrate through steric constrictions drops sharply when the pore size falls below a critical value (6,7). Moreover, the pore size of the network matrix strongly influences cell behavior such as adhesion and polarization, and therefore needs to be accurately measured (1,8,9).

Typical examples of self-assembled biopolymer networks, ubiquitously used for 3D cell culture, are three-dimensional collagen matrices. They are composed of randomly oriented fibers that form when monomeric collagen polymerizes into a hierarchical structure (10–14).

Another important biopolymer network is fibrin, which provides the structural scaffold for blood clots but is also frequently used in tissue engineering applications and cell culture (15,16). Fibrin networks form during coagulation, when monomeric fibrin assembles into protofibrils that laterally aggregate into thicker fibers and occasionally branch to form a percolated, three-dimensional structure (17).

Changes in fiber diameter and density strongly affect the mechanics of both collagen and fibrin networks (18–21),

as well as the adhesion, spreading, polarization, and migration of embedded cells (1,3–5,22). These biological effects are attributable not only to the mechanical network properties or adhesive ligand density but also to the morphological structure of the network, most notably the pore size. Because of the low solid (protein) fraction in these networks, typically 0.05% to 0.5% (w/v), traditional methods of measuring porosity are not sensitive enough to be useful. In a similar way, hydrodynamic permeability can only serve as an indirect measure of pore size and critically depends on the validity of hydrodynamic models. Rather, network morphology is best characterized by a mesh size, or pore size, given by the 3D spacing of the fibers within the interstitial fluid, which can be directly obtained from microscopic images. Moreover, it is the pore size and interfiber cross-link distance that most critically sets the steric hindrance for the migrating cells and also the network mechanical properties (18,23–27).

There are several approaches for quantifying the network pore size from images of the network structure. Scanning electron microscopy (SEM) has excellent resolution (3,27–29) but requires the samples to be dehydrated and thus can only image a potentially collapsed network structure. By contrast, light microscopy methods can be applied to a fully hydrated sample even when it contains living cells. A widely used imaging modality is confocal reflectance microscopy (CRM) (5,6,10,11,30). This methodology offers a fundamental advantage over confocal fluorescence microscopy (CFM) in that the network need not be labeled with fluorophores, which is both time consuming and expensive. Moreover, less laser power is required to obtain the image stack with CRM, making it possible to avoid cell damage

Submitted May 23, 2013, and accepted for publication September 17, 2013.

\*Correspondence: [bfabry@biomed.uni-erlangen.de](mailto:bfabry@biomed.uni-erlangen.de)

Editor: Gijsje Koenderink.

© 2013 by the Biophysical Society  
0006-3495/13/11/1967/9 \$2.00



during live-cell imaging (4,22,30). These advantages make CRM a preferred method for simultaneous observation of cell migration and network structure.

CRM has a major disadvantage, however. Because CRM only detects light that is reflected back into the microscope lens, it preferentially visualizes horizontal fibers. Thus, CRM suffers from a blind spot in that it misses fibers with an angle steeper than a certain cutoff angle (31). Therefore, networks imaged with CRM appear anisotropic, and fewer fibers are visible, resulting in a substantial overestimation of the pore size of the network. In a similar way, second harmonic generation microscopy (SHG), another popular mode for imaging collagen, also suffers from an anisotropic transfer function (13). If, however, the cutoff angle for the missing fibers is known, it seems possible to correct for the blind spot effect. Such an approach would allow users to employ the convenient methods of CRM and SHG and still evaluate the pore size distribution of the network without bias.

Here, we introduce a method for determining the unbiased pore size of a biopolymer network when it is imaged with CRM and SHG. As a mathematically well-defined and robust measure for the network pore size, we introduce the nearest-obstacle distance (NOD). In the case of random networks, regardless of isotropy or anisotropy, it can be shown that the distribution of NODs follows a Rayleigh distribution. Furthermore, if fibers oriented at an angle above a cutoff angle are systematically removed from the network, the NOD still follows a Rayleigh distribution with a scaling parameter that is a monotonic function of the cutoff angle (see the [Supporting Material](#)).

Therefore, we can fit a Rayleigh distribution to the distribution of the NODs obtained from CRM and SHG images and then simply rescale the distribution function by a correction factor to predict the unbiased pore size distribution of the full network. We verify the validity of our approach by comparing the pore size distribution predicted from CRM data with the pore sizes directly measured by CFM on the same collagen samples. Furthermore, we demonstrate how the NOD can be converted to a previously established pore size measure—the covering radius transform. Finally, we show how the pore sizes of collagen and fibrin gels depend on protein concentration, and we compare our data to theoretical predictions.

## MATERIALS AND METHODS

### Preparation of collagen gels

Collagen matrices were produced under sterile conditions and with all ingredients held on ice to avoid premature polymerization during the mixing process. To avoid bubble formation, extra care was taken while mixing the ingredients. For every experiment, a stock solution was prepared and diluted with buffer solution consisting of 8 mL H<sub>2</sub>O, 1 mL 10× DMEM, and 1 mL NaHCO<sub>3</sub>, adjusted to pH 10 with 1 M NaOH, until the final concentration was reached. A 2.4 mg/mL collagen stock solution

was mixed out of 1.2 mL Collagen R (2 mg/mL rat collagen type I; Serva, Heidelberg, Germany) and 1.2 mL Collagen G (4 mg/mL bovine collagen type I; Biochrom, Cambridge, United Kingdom). Moreover, 270 μL of a 0.25 M NaHCO<sub>3</sub> buffer solution and 270 μL 10× DMEM (Biochrom) were added. To adjust the pH to 10, 43 μL of a 1 M NaOH solution was added. 1.2 mL of the mixture was pipetted in 35 mm cell culture dishes and polymerized in a cell culture incubator at 95% humidity, 5% CO<sub>2</sub>, at 37°C. After 2 h, 2 mL of 1× DMEM complete medium was added.

### Preparation of fluorescently labeled collagen gels

To obtain fluorescent images of collagen gels, a fraction of the Collagen G stock solution was labeled with 5-(and 6) carboxytetramethylrhodamine succinimidyl ester (TAMRA-SE; Invitrogen, Carlsbad, CA) at 4°C according to the manufacturer's protocol. To minimize possible alterations of the polymerized network due to the labeling process, the labeled Collagen G solution was mixed with unlabeled stock solution at a volume ratio of 1:6. The mixture of collagen solutions was used to prepare the gels as described above.

### Preparation of unlabeled and labeled fibrin gels

Lyophilized, plasminogen-free human fibrinogen and a lyophilized human α-thrombin solution (both from Enzyme Research Laboratories, South Bend, IN) were rehydrated according to manufacturer instructions and immediately frozen in aliquots at −80°C. Before the experiments, aliquots were thawed and thrombin was kept on ice, whereas fibrinogen was kept at room temperature. Both fibrinogen and thrombin were diluted with a buffer solution of 0.15 M NaCl and 20 mM HEPES, pH 7.4, to twice the final concentration for each, and fibrin gels were polymerized by mixing these solutions 1:1 v/v. Polymerization was allowed for at least 30 min at room temperature, after which 1 mL of buffer was added to prevent evaporation. For fluorescently labeled fibrin gels, the fibrinogen stock solution was labeled with TAMRA-SE at room temperature following the same protocol used for collagen. A 1:6 v/v mixture of labeled and unlabeled fibrinogen monomer was used to synthesize the gels.

### Confocal microscopy imaging

Stacks of optical sections were acquired with an upright SP5X confocal microscope (Leica, Wetzlar, Germany) equipped with a Leica 20× dip-in water-immersion objective (NA 1.0). To measure the effect of the NA on the cutoff angle, some of the images were also acquired with a Leica 20× water-immersion objective (NA 0.7) or a Leica 63× water-immersion objective (NA 1.2), both corrected for imaging through a 170 μm glass coverslip. CRM stacks were recorded by collecting the reflected light of the 488 nm argon laser line in one channel. CFM stacks were simultaneously acquired by collecting light between 571 nm and 772 nm in a second channel while the sample was also illuminated with a 543 nm HeNe laser. All images were recorded at 512 × 512 pixels with a digital magnification of 4.55, resulting in a pixel size of 317 × 317 nm; a total of 597 z-slices with a spacing of 335 nm were collected for each stack. Every line was averaged three times during scanning.

### SHG imaging

Collagen hydrogels were excited with a femtosecond-pulsed Ti:Sa-Laser (Vision II, Coherent, Santa Clara, CA) set to 810 nm. SHG signals were collected at 405 nm using a multiphoton microscope (TriM-Scope II, LaVision Biotech, Bielefeld, Germany) with a Zeiss 40×, NA 1.1 water immersion objective. Voxel size was 310 × 310 × 310 nm; scanning speed was set to 1200 Hz with a 2× line average.

## Binarization of the image stacks

A binarized data set with pixel values of 1 (representing the solid phase (fiber)) and 0 (representing the fluid phase) was obtained using a template-matching algorithm that simultaneously performs a binarization and skeletonization of the image stack (33). In the algorithm, small subsets of the image stack are compared to a template representing a diffraction-limited fiber cross section. Matching voxels are classified as fibers based on a mismatch threshold that is iteratively optimized for each image stack so that the final skeletonized network obeys a universal property of voxelized random line networks, that is, solid-phase voxels most likely have three solid-phase neighbors in a  $3 \times 3 \times 3$  pixel neighborhood. All isolated solid-phase voxels are regarded as image noise and are removed. The method is self-adapting, largely insensitive to the signal and noise in the image, and free of user-selected parameters (33). However, the pore size characterization and conversion presented in this article do not rely on a specific skeletonization and binarization method.

## Pore size evaluation from the distribution of NODs

In a three-dimensional binarized data set, the NOD,  $r_{nod}$ , describes the Euclidean distance from a 0-phase fluid point to the closest 1-phase solid point (see Fig. 3 A, inset). To obtain the characteristic distribution,  $p(r_{nod})$ , of each data set, we sampled  $r_{nod}$  at 10,000 randomly selected points.

Collagen networks closely resemble random networks, consisting of straight line segments with a spatially homogeneous density of the line centers and isotropic directions. For such a Mikado-like network, the distribution,  $p(r_{nod})$ , of NODs can be calculated analytically ((34) and see the Supporting Material). In the limiting case where the line segments are much longer than the average pore diameter,  $p(r_{nod})$  is given by a Rayleigh distribution with a single parameter,  $r_{mean}$ , which represents the mean value of  $r_{nod}$ .

$$p(r_{nod}) = \frac{r_{nod}}{\left(\frac{\sqrt{2}}{\pi}r_{mean}\right)^2} \exp\left(-\frac{r_{nod}^2}{2\left(\frac{\sqrt{2}}{\pi}r_{mean}\right)^2}\right). \quad (1)$$

Because  $r_{mean}$  is the only parameter needed to fit the Rayleigh distribution to the measured  $p(r_{nod})$ ,  $r_{mean}$  is a robust measure for the pore size.

## Pore size evaluation using the covering radius transform

A second method to determine the pore size distribution of a binarized fiber network is the covering radius transform (CRT) (35). For every voxel of the liquid phase, the Euclidean distance to the center of the next solid-phase voxel is determined. This distance can be interpreted as the radius of a pure liquid-phase sphere around that voxel. The CRT assigns every voxel of the fluid phase the value of the radius of the largest possible liquid-phase sphere, placed anywhere, that covers that pixel. This results in a complete coverage of the fluid space by overlapping spheres with a distribution of radii. For random biopolymer networks, the probability distribution of the radii more closely resembles a normal distribution, where the maximum of the distribution is approximately the average pore size of the gel.

## Evaluation of fiber orientation

We determined the direction vector of short fiber segments by treating their brightness distribution as a mass distribution, computing and diagonalizing the moment-of-inertia tensor, and finding the easy axis of minimal inertia. This easy axis points in the direction of the locally straight line segment.

From the binarized stack,  $10^5$  spherical volume segments with a radius of 3 pixels, with each sphere containing at least five solid-phase voxels, were chosen randomly. The voxels were treated as mass points, located at the voxel centers, with constant mass  $m = 1$  for all solid-phase voxels and  $m = 0$  for all liquid-phase voxels. After determining the easy axis of the inertia tensor, the corresponding unit direction vector in spherical coordinates was computed. This vector does not depend on the exact position of the sphere's center as long as the same solid voxels are enclosed. Finally, a histogram was generated for the polar angle,  $\theta$ , of the direction vectors from all fiber segments (see Fig. 3 B).

All analysis steps after acquiring the image data stack are summarized in a flow chart (Fig. S3 in the Supporting Material).

## RESULTS AND DISCUSSION

### The blind spot in CRM and SHG

CRM is widely used to visualize biopolymer networks but can only detect fibers that are oriented roughly horizontal to the imaging plane, up to a certain cutoff angle, thus missing the vertical fibers (the so-called blind-spot effect) (31). To illustrate this effect, we imaged collagen gels simultaneously in CRM and CFM. The data obtained with CFM show an isotropic distribution of fibers, whereas the data obtained with CRM preferentially show fibers aligned horizontally with the imaging plane (Fig. 1).

To test whether the blind-spot effect using CRM also occurs in other biopolymer systems, we repeated the same measurements on fibrin gels.  $xz$  slices of the fibrin gels obtained with CRM clearly demonstrate that in this case also, vertical fibers are missing (Fig. 2, A and B, and Movie S1).

Similar to CRM, SHG also exhibits an emission probability that depends on the orientation of the excited fibers. To test whether this leads to a blind spot, we acquired 3D data stacks of collagen gels with SHG in backscatter mode. Again, vertical fibers cannot be detected by this imaging modality (Fig. 2, C and D, and Fig. S2 B).

To evaluate the pore sizes of biopolymer networks from such directionally biased data, the missing vertical fibers have to be taken into account. The brightness of fiber segments imaged with CRM depends on their polar angle,  $\theta$  (31). Only fibers with angles larger than  $\theta_{cut}$  relative to the optical axis can be observed (Fig. 3 A). This cutoff angle depends predominantly on the NA of the imaging system. To determine the value of  $\theta_{cut}$  in our microscopy system, with a 1.0 NA objective, we quantified the probability density distribution of collagen and fibrin fiber angles in the CRM data (Fig. 3 B inset). The density distribution of polar angles is identical for both biopolymers (Fig. 3 B inset), indicating that the angular characteristics of light reflection of collagen and fibrin fibers are similar. For comparison, we also measured the probability density distribution of collagen and fibrin fiber polar angles in the CFM data. Both probability density distributions closely follow the  $\sin\theta$  dependency expected for an isotropic random line network (Fig. 3 B and inset), which demonstrates that our collagen and fibrin networks are indeed isotropic.

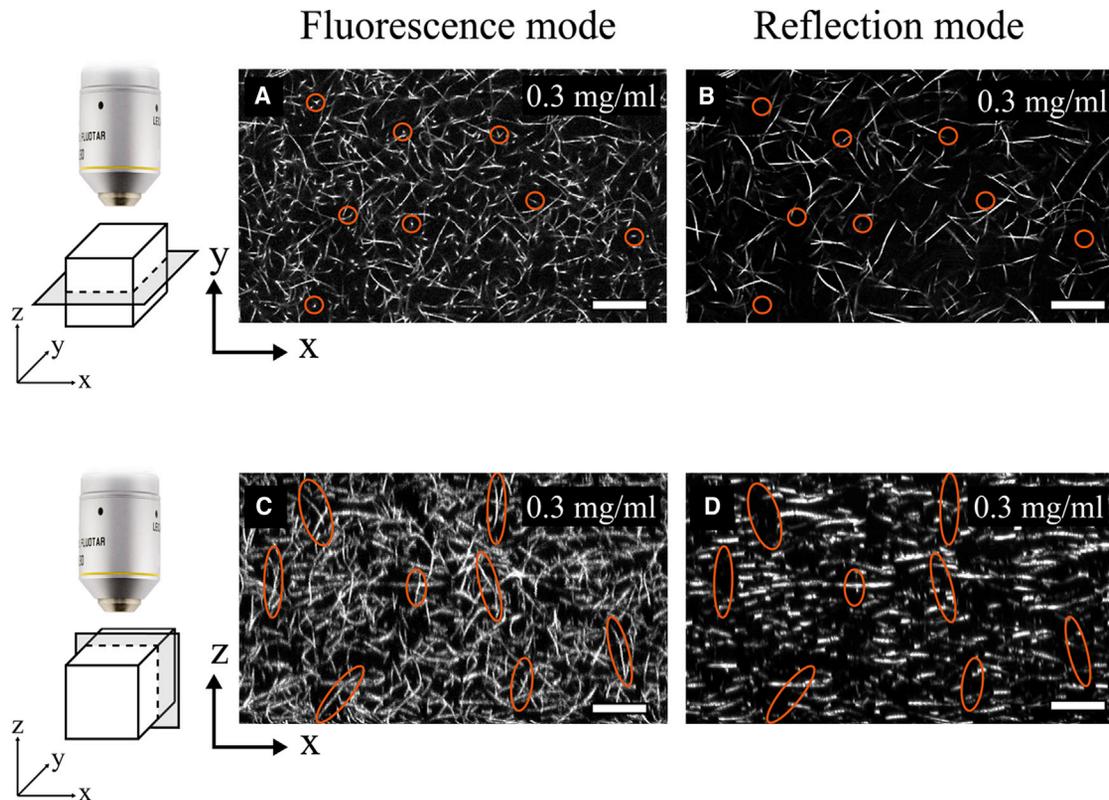


FIGURE 1 The blind spot in CRM. (A and B) Maximum-intensity projection of 3  $xy$  slices (total thickness  $1.0\ \mu\text{m}$ ) of a  $0.3\ \text{mg/mL}$  collagen gel imaged simultaneously with CFM (A) and CRM (B). (C and D) Projected view of 15  $xz$  slices (total thickness  $4.75\ \mu\text{m}$ ) of the same sample imaged with CFM (C) and CRM (D). Compared to CFM, CRM does not detect vertical fibers. Scale bars,  $20\ \mu\text{m}$ . Stacks were imaged with  $512 \times 512$  pixels with a size of  $317 \times 317 \times 335\ \text{nm}$ . To see this figure in color, go online.

Next, we define the cutoff angle,  $\theta_{cut}$ , such that the integral over the isotropic angle distribution in the interval  $(\theta_{cut}, 90^\circ)$  (Fig. 3 B, inset gray area) equals the total integral over the fiber angle distribution from the CRM data. The cutoff angle,  $\theta_{cut}$ , is  $51^\circ$  in our standard system with an NA 1.0 objective for both fibrin and collagen data. To measure how the cutoff angle depends on the NA, we measured the polar angle distribution also for fibrin gels imaged with CRM using either a water immersion  $63\times$  objective with NA 1.2 or a water immersion  $20\times$  objective with NA 0.7 (Fig. 3). As previously suggested (31), the cutoff angle increases with lower NA (Fig. 3 B); for a  $63\times$ , NA 1.2 objective, we find that  $\theta_{cut} = 46^\circ$ , and for a  $20\times$ , NA 0.7 objective,  $\theta_{cut} = 62^\circ$ . Furthermore, similar cutoff angles are also observed in SHG images of collagen networks in backscatter mode (Fig. S2), where we find a cutoff angle of  $\theta_{cut\ SHG} = 48^\circ$  for a  $40\times$ , NA 1.1 water immersion objective.

### Rayleigh distribution of NOD as a measure of pore size distribution

For a simulated random line network with a line thickness of one voxel, the distribution of NODs calculated for

randomly chosen points of the fluid phase follows a Rayleigh distribution (see the Supporting Material). Networks with a high density have narrow distributions with a prominent peak. For networks with lower densities, the peak shifts to the right and the distribution broadens (Fig. 4 A). Conveniently, the Rayleigh distribution relates the peak and the width of the distribution curve using only one free parameter—the average distance,  $r_{mean}$ . Moreover, it can be shown analytically that the distribution of NODs in a random line network follows a Rayleigh distribution even when the network is anisotropic (see the Supporting Material). To demonstrate this also numerically, we simulated an isotropic random line network and then removed fibers at cutoff angles of  $40^\circ$  and  $55^\circ$ , respectively. For both anisotropic networks, we found that the Rayleigh distribution fits the data perfectly, albeit with a right-shifted peak and a broader distribution compared to the isotropic network (Fig. 4 B), as expected. Therefore, the characteristic mean value,  $r_{mean}$ , of the Rayleigh distribution fitted to the  $p(r_{nod})$  distribution of NODs can be taken as a measure of the average pore size of the network imaged with either CRM or CFM.

This independence of the Rayleigh distribution from fiber anisotropy offers the opportunity to reconstruct the

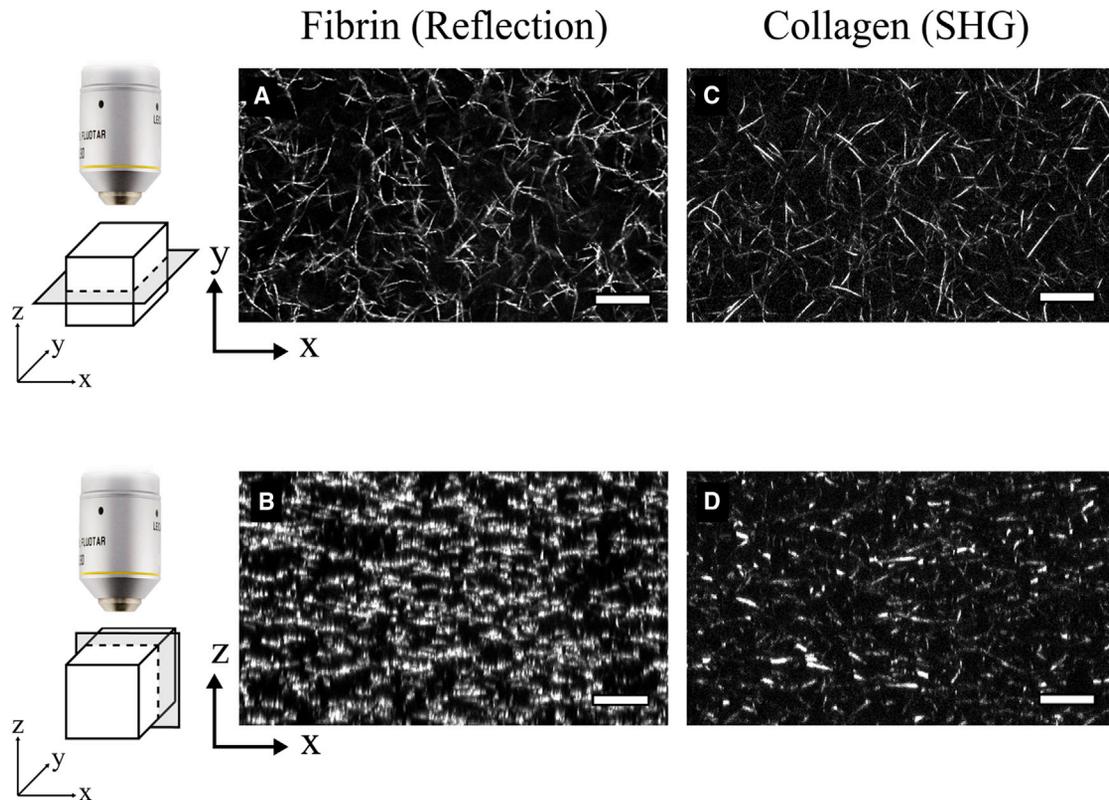


FIGURE 2 The blind spot in fibrin gels imaged with CRM and in collagen gels imaged with SHG. (A) Maximum-intensity projection of 3  $xy$  slices (total thickness  $1.0\ \mu\text{m}$ ) of a  $1.0\ \text{mg/mL}$  fibrin gel imaged with CRM. (B) Maximum-intensity projection of 15  $xz$  slices (total thickness  $4.75\ \mu\text{m}$ ) of the sample image in A. Similar to the collagen gels shown in Fig. 1, vertical fibers in fibrin gels are missing. (C and D) SHG images of a  $0.3\ \text{mg/mL}$  collagen gel. (C) Maximum-intensity projection of 3  $xy$  slices (total thickness  $0.93\ \mu\text{m}$ ). (D) Maximum-intensity projection of 15  $xz$  slices (total thickness  $4.65\ \mu\text{m}$ ) of the sample image in C. Similar to CRM (Figs. 1 and 2 B), SHG images preferentially show horizontal fibers. To see this figure in color, go online.

unbiased network pore size from CRM and SHG images and to correct for the blind-spot effect. It can be analytically shown that the  $r_{\text{mean,unbiased}}$  of an isotropic network and the  $r_{\text{mean,based}}$  of the same network but with invisible fibers above a cutoff angle are related by a factor that is simply the square root of the fraction of the visible fibers (see the Supporting Material). For a cutoff angle of  $51^\circ$ , this correction factor is  $\sqrt{\cos(51^\circ)} = 0.793$ . The correction factor is constant for different network densities and only depends on the optical cutoff angle,  $\theta_{\text{cut}}$ , of the imaging system:

$$r_{\text{mean,unbiased}} = r_{\text{mean,based}} \times \sqrt{\cos \theta_{\text{cut}}}. \quad (2)$$

At first glance, an overestimation of the pore size by 25% from uncorrected CRM images may not seem dramatic. However, several recent reports have demonstrated that cell migration in a porous environment critically depends on pore size, with a surprisingly sharp cutoff below which cells cannot migrate (6,7). A 25% change in pore size around a mean diameter of  $2\ \mu\text{m}$ , for instance, was associated with a  $>10$ -fold change in transmigration efficiency (6).

### Comparison with the CRT

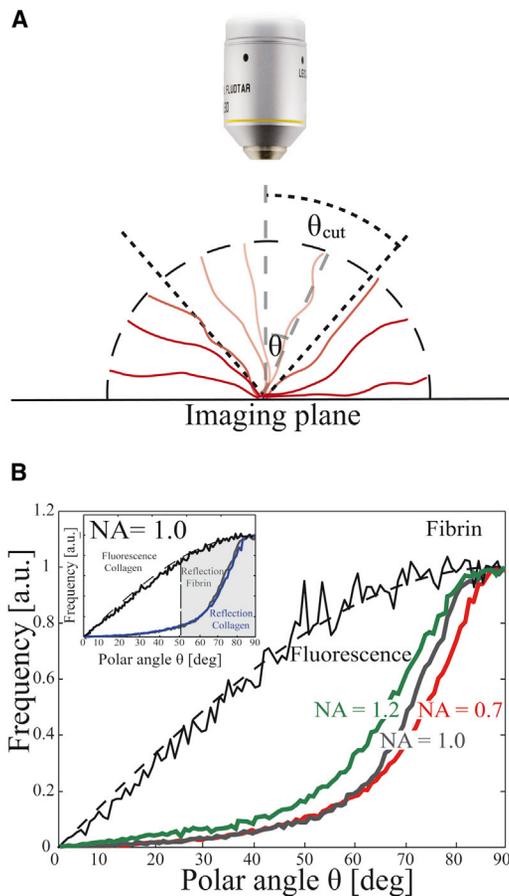
To compare the pore sizes obtained using the NOD method with the pore sizes obtained using the established CRT method (35), the scaling factor between the two measures was numerically determined for simulated isotropic networks (see the Supporting Material). The average pore radius given by the CRT,  $r_{\text{CRT}}$ , is larger by a factor of  $f = 1.82$  than the  $r_{\text{mean,unbiased}}$  obtained with the NOD method.

$$\begin{aligned} r_{\text{CRT}} &= 1.82 r_{\text{mean,unbiased}} \\ &= 1.82 \times r_{\text{mean,based}} \times \sqrt{\cos \theta_{\text{cut}}}. \end{aligned} \quad (3)$$

For random isotropic networks, the factor,  $f$ , relating the pore sizes obtained with the two methods is independent of the imaging setup and the cutoff angle.

### Application to collagen data

To test whether these analytical and numerical results also hold for real biopolymer networks, we applied our method to a set of collagen gels polymerized at different monomer concentrations imaged with the  $20\times$  dip in objective



**FIGURE 3** Distribution of fiber orientation. (A) In CRM, the brightness of fiber segments depends on their polar angle  $\theta$ , as illustrated by different shades of red (more saturated color indicates brighter CRM signal). Only fibers with angles larger than  $\theta_{cut}$  relative to the optical axis, can be observed. This cutoff angle depends on the NA of the objective lens. (B) Distribution of polar angles of a fibrin network measured with CRM and three different water immersion objectives ( $20\times$ , 0.7 NA (red line),  $20\times$ , 1.0 NA (gray line), and  $63\times$ , 1.2 NA (green line)) compared to the distribution of polar angles for a fibrin network measured with CFM (black solid line). The polar angle distribution of the fluorescent data set follows the expected  $\sin(\theta)$  distribution of an ideal, isotropic network (black dotted line). The fraction of fiber segments that are visible in CRM is given by the ratio of the integrals of both frequency distributions. As expected, a higher NA increases the fraction of visible fibers. (Inset) Distribution of polar angles of a fibrin (gray line) and a collagen (blue line) network measured with a  $20\times$ , NA 1.0 water immersion objective with CRM, and in a collagen network measured with CFM (black solid line) that follows the expected  $\sin(\theta)$  distribution of an isotropic network (black dotted line). The cutoff angle,  $\theta_{cut}$ , is chosen so that the gray area equals the integral of the polar angle distribution for the CRM data. To see this figure in color, go online.

(NA 1.0). For each gel, we obtained a CRM and CFM data set and measured the distributions of NODs. For all concentrations imaged, the  $r_{nod}$  distributions followed Rayleigh distributions for both CRM and CFM. As expected, the CRM data consistently showed larger pores than the CFM data (Fig. 5, A–D). With decreasing collagen concentrations, these distributions broadened and the maxima shifted to

larger pore sizes. The quality of the Rayleigh fit was very high, with correlation coefficients between  $r^2 = 0.987$  for 2.4 mg/mL collagen gels and  $r^2 = 0.997$  for 0.3 mg/mL collagen gels. The CRM data of fibrin gels and the SHG data of collagen gels also followed a Rayleigh distribution (Fig. S2 A), thus further confirming that the NOD of random biopolymer networks is Rayleigh distributed, irrespective of imaging mode or protein composition.

We then computed the  $r_{mean,bias}$  by fitting a Rayleigh distribution (Eq. 1) to the anisotropic CRM data, and predicted the unbiased distribution of NODs for an isotropic network according to Eq. 2. The Rayleigh distribution predicted from the CRM data and the measured distribution from CFM data show excellent agreement for all collagen concentrations (Fig. 5, A–D), confirming the validity of our method.

We next predicted the CRT pore size,  $r_{CRT}$ , for different collagen gels from the fit of the Rayleigh distribution to the CRM data according to Eq. 3. The predicted  $r_{CRT}$  was then compared to the directly measured  $r_{CRT}$  from the CFM images. We found a close agreement between predicted and directly measured pore sizes for the different collagen concentrations (Fig. 5 E), confirming the validity of Eq. 3 for converting the different pore size measures.

### Relationship between pore size and protein concentration

Our data show that collagen gels polymerized at higher monomer concentrations have smaller pore sizes, consistent with previous findings (6,19). Moreover, by analyzing the relationship between protein concentration and pore size for both collagen and fibrin, it is possible to address the question of whether higher monomer concentrations are stoichiometrically incorporated into the fibers, and whether the incorporated monomers contribute to increased fiber length or instead to increased fiber thickness.

We can take advantage of the fact that the scaling factor,  $r_{mean}$ , of the Rayleigh function only depends on the total fiber length/unit volume,  $\lambda = Nl/V$ , and not on fiber thickness, according to  $r_{mean} = 1/2\sqrt{\lambda}$  (see the Supporting Material).  $N$  is the number of fibers, and  $l$  is the average length of the fibers.

As  $Nl \sim c$ , with polymer concentration  $c$ , it follows that  $\lambda \sim c/V$  and  $r_{mean} \sim 1/2\sqrt{c}$ .

If protein monomers are stoichiometrically incorporated into fibers, and if they contribute only to fiber lengthening and not to fiber thickening, we expect that the pore size increases with concentration  $c_m$  according to  $r_{mean} \sim c_m^{-1/2}$ . The pore sizes determined for collagen polymerized at concentrations between 0.6 and 2.4 mg/mL and for fibrin polymerized at concentrations between 0.125 and 8 mg/mL closely follow this prediction (Fig. 6). However, for very high fibrin concentrations and accordingly small

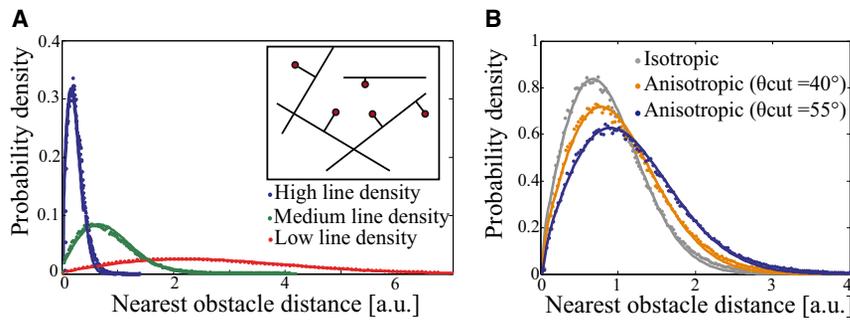


FIGURE 4 Rayleigh distribution of NODs. (A) The distribution of NODs in simulated line networks is determined for three different line densities by computing the Euclidean distance to the nearest solid-phase pixel (*inset*). Regardless of line density, the resulting probability density calculated for  $10^5$  test points randomly placed in the liquid phase (*dots*) follows a Rayleigh distribution (*solid lines*). (B) Probability densities of NODs calculated for an isotropic line network (*gray*) and two anisotropic networks with cut-off angles of  $40^\circ$  and  $55^\circ$ . In all cases, the probability distribution of the NOD follows a Rayleigh distribution (*solid lines*). To see this figure in color, go online.

pore sizes, the relationship deviates from the expected square-root dependency (Fig. 6), suggesting that pore sizes  $< 1 \mu\text{m}$  cannot be reliably resolved.

Collagen polymerization is known to be strongly influenced by the source of the collagen (36) but also by pH and temperature (18,19,27,38), which differ between protocols from different laboratories. Previous studies (39) have reported a stronger relationship between collagen monomer concentration and pore size, according to  $r_{mean} \sim c_m^{-1}$ . Our gels were polymerized at  $37^\circ\text{C}$  and with pH 10, however, which leads to a faster polymerization (38) compared to the more widely used pH 7 or lower (19,20,38), and therefore may have contributed to these differences. Furthermore, it has been suggested that fluorescent labeling of collagen might affect its polymerization behavior (40). Even though we do not see systematic pore size differences between fluorescently labeled and unlabeled collagen, our CRM-based method avoids this issue altogether.

### CONCLUSION

We have shown here analytically, numerically, and experimentally that the NOD distribution of isotropic and anisotropic random line networks, such as reconstituted biopolymer networks, follows a Rayleigh distribution. Moreover, if a fraction of the network fibers cannot be visualized for technical reasons such as the blind-spot effect in CRM or SHG imaging, the unbiased pore size distribution of the network can still be recovered by a simple rescaling of the Rayleigh distribution, if the fraction of invisible fibers is known. For an isotropic network, the fraction of invisible fibers can be estimated from the polar angle distribution of fiber orientations. As an important biophysical application of this approach, we investigated the dependence of the pore sizes of self-assembled collagen and fibrin networks on protein concentration. We find that the pore size decreases with the square root of the protein concentration, consistent with a total fiber length that scales linearly with concentration.

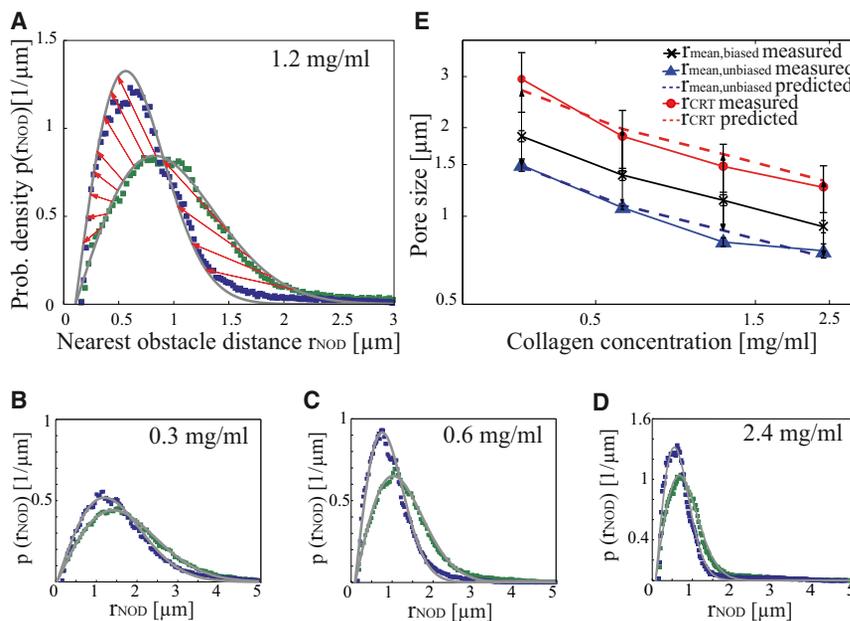


FIGURE 5 Blind spot correction. (A) From the binarized reflection data, the NOD is determined (*green squares*) and fitted with a Rayleigh distribution (*gray line*). Rescaling this distribution with a correction factor according to Eq. 2 yields a prediction for the distribution of NOD for CFM data, as indicated by the red arrows. The distribution of the NODs calculated from a binarized fluorescent data set obtained for the same collagen sample (*blue squares*) is correctly predicted by the rescaled Rayleigh curve (this curve is not a fit to the measurements). (B–D) Repeating this procedure for different collagen concentrations from 0.3 mg/mL to 2.4 mg/mL gave excellent agreement between predicted and measured distributions. (E) Average pore size,  $r_{mean, biased}$ , for different collagen concentrations calculated from reflection data (*black line*), blind-spot corrected ( $r_{mean, unbiased}$ ; *blue dashed line*) with Eq. 2, and converted to the corresponding mean CRT values ( $r_{CRT}$ ; *red dashed line*) with Eq. 3. All predicted values are in good agreement with the measured data from fluorescent images (*solid lines*). The error bars indicate the SD between different fields of view ( $n = 8$ ) of the same samples. To see this figure in color, go online.

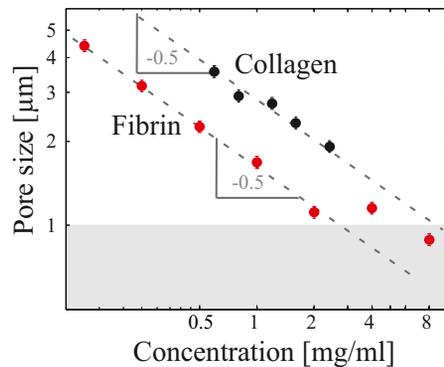


FIGURE 6 Concentration dependence of the average pore size of fibrin (red) and collagen (black) gels. Average pore size,  $r_{mean,unbiased}$ , reconstructed from 3D CRM data versus protein concentration. The error bars indicate the mean  $\pm$  SD from different collagen gels ( $n \geq 3$ ) and different fields of view ( $n \geq 5$  for each gel). Both data sets are in good agreement, with a power law with an exponent of  $-0.5$  (gray dashed line), as expected for stoichiometrically polymerizing networks where monomer addition contributes to fiber lengthening but not to fiber thickening. The gray area illustrates the range ( $r < 1 \mu\text{m}$ ) where the pore sizes cannot be reliably measured. To see this figure in color, go online.

## SUPPORTING MATERIAL

Five figures, one movie, reference (41) and supplemental information are available at [http://www.biophysj.org/biophysj/supplemental/S0006-3495\(13\)01086-2](http://www.biophysj.org/biophysj/supplemental/S0006-3495(13)01086-2).

We thank Dr. Martin Vielreicher, Institute of Medical Biotechnology, for help with SHG imaging.

This work was supported by grants from the German Science Foundation (DFG), the Emerging Fields Initiative of the University of Erlangen-Nuremberg, and the European Research Council (Starting Grant 211166 MINATRAN).

## REFERENCES

- Cukierman, E., R. Pankov, ..., K. M. Yamada. 2001. Taking cell-matrix adhesions to the third dimension. *Science*. 294:1708–1712.
- Pedersen, J. A., and M. A. Swartz. 2005. Mechanobiology in the third dimension. *Ann. Biomed. Eng.* 33:1469–1490.
- Zaman, M. H., L. M. Trapani, ..., P. Matsudaira. 2006. Migration of tumor cells in 3D matrices is governed by matrix stiffness along with cell-matrix adhesion and proteolysis. *Proc. Natl. Acad. Sci. USA*. 103:10889–10894.
- Friedl, P., and E.-B. Bröcker. 2000. The biology of cell locomotion within three-dimensional extracellular matrix. *Cell. Mol. Life Sci.* 57:41–64.
- Friedl, P., and K. Wolf. 2010. Plasticity of cell migration: a multiscale tuning model. *J. Cell Biol.* 188:11–19.
- Wolf, K., M. Te Lindert, ..., P. Friedl. 2013. Physical limits of cell migration: control by ECM space and nuclear deformation and tuning by proteolysis and traction force. *J. Cell Biol.* 201:1069–1084.
- Tozluoğlu, M., A. L. Tournier, ..., E. Sahai. 2013. Matrix geometry determines optimal cancer cell migration strategy and modulates response to interventions. *Nat. Cell Biol.* 15:751–762.
- Petroll, W. M., L. Ma, and J. V. Jester. 2003. Direct correlation of collagen matrix deformation with focal adhesion dynamics in living corneal fibroblasts. *J. Cell Sci.* 116:1481–1491.
- Brábek, J., C. T. Mierke, ..., B. Fabry. 2010. The role of the tissue microenvironment in the regulation of cancer cell motility and invasion. *Cell Commun. Signal.* 8:22.
- Helseth, Jr., D. L., and A. Veis. 1981. Collagen self-assembly in vitro. Differentiating specific telopeptide-dependent interactions using selective enzyme modification and the addition of free amino telopeptide. *J. Biol. Chem.* 256:7118–7128.
- Gross, J., J. H. Highberger, and F. O. Schmitt. 1954. Collagen structures considered as states of aggregation of a kinetic unit. The tropocollagen particle. *Proc. Natl. Acad. Sci. USA*. 40:679–688.
- Christiansen, D. L., E. K. Huang, and F. H. Silver. 2000. Assembly of type I collagen: fusion of fibril subunits and the influence of fibril diameter on mechanical properties. *Matrix Biol.* 19:409–420.
- Orgel, J. P., T. C. Irving, ..., T. J. Wess. 2006. Microfibrillar structure of type I collagen in situ. *Proc. Natl. Acad. Sci. USA*. 103:9001–9005.
- Kadler, K. E., D. F. Holmes, ..., J. A. Chapman. 1996. Collagen fibril formation. *Biochem. J.* 316:1–11.
- Janmey, P. A., J. P. Winer, and J. W. Weisel. 2009. Fibrin gels and their clinical and bioengineering applications. *J. R. Soc. Interface.* 6:1–10.
- Arkudas, A., A. Balzer, ..., U. Kneser. 2013. Evaluation of angiogenesis of bioactive glass in the arteriovenous loop model. *Tissue Eng. Part C Methods*. 19:479–486.
- Chernysh, I. N., C. Nagaswami, and J. W. Weisel. 2011. Visualization and identification of the structures formed during early stages of fibrin polymerization. *Blood*. 117:4609–4614.
- Yang, Y. L., and L. J. Kaufman. 2009. Rheology and confocal reflectance microscopy as probes of mechanical properties and structure during collagen and collagen/hyaluronan self-assembly. *Biophys. J.* 96:1566–1585.
- Yang, Y. L., S. Motte, and L. J. Kaufman. 2010. Pore size variable type I collagen gels and their interaction with glioma cells. *Biomaterials*. 31:5678–5688.
- Raub, C. B., A. J. Putnam, ..., S. C. George. 2010. Predicting bulk mechanical properties of cellularized collagen gels using multiphoton microscopy. *Acta Biomater.* 6:4657–4665.
- Ryan, E. A., L. F. Mockros, ..., L. Lorand. 1999. Structural origins of fibrin clot rheology. *Biophys. J.* 77:2813–2826.
- Friedl, P. 2004. Dynamic imaging of cellular interactions with extracellular matrix. *Histochem. Cell Biol.* 122:183–190.
- Piechocka, I. K., R. G. Bacabac, ..., G. H. Koenderink. 2010. Structural hierarchy governs fibrin gel mechanics. *Biophys. J.* 98:2281–2289.
- MacKintosh, F. C., J. Käs, and P. A. Janmey. 1995. Elasticity of semiflexible biopolymer networks. *Phys. Rev. Lett.* 75:4425–4428.
- Storm, C., J. J. Pastore, ..., P. A. Janmey. 2005. Nonlinear elasticity in biological gels. *Nature*. 435:191–194.
- Gardel, M. L., J. H. Shin, ..., D. A. Weitz. 2004. Elastic behavior of cross-linked and bundled actin networks. *Science*. 304:1301–1305.
- Raub, C. B., J. Unruh, ..., S. C. George. 2008. Image correlation spectroscopy of multiphoton images correlates with collagen mechanical properties. *Biophys. J.* 94:2361–2373.
- Brown, A. E. X., R. I. Litvinov, ..., J. W. Weisel. 2009. Multiscale mechanics of fibrin polymer: gel stretching with protein unfolding and loss of water. *Science*. 325:741–744.
- Weisel, J. W. 2004. The mechanical properties of fibrin for basic scientists and clinicians. *Biophys. Chem.* 112:267–276.
- Brightman, A. O., B. P. Rajwa, ..., S. L. Voytik-Harbin. 2000. Time-lapse confocal reflection microscopy of collagen fibrillogenesis and extracellular matrix assembly in vitro. *Biopolymers*. 54:222–234.
- Jawerth, L. M., S. Münster, ..., D. A. Weitz. 2010. A blind spot in confocal reflection microscopy: the dependence of fiber brightness on fiber orientation in imaging biopolymer networks. *Biophys. J.* 98:L1–L3.
- Reference deleted in proof.

33. Krauss, P., C. Metzner, ..., B. Fabry. 2012. Parameter-free binarization and skeletonization of fiber networks from confocal image stacks. *PLoS ONE*. 7:e36575.
34. Metzner, C., P. Krauss, and B. Fabry. 2011. Pore sizes in random line networks. ArXiv1110.1803v1.
35. Mickel, W., S. Münster, ..., G. E. Schröder-Turk. 2008. Robust pore size analysis of filamentous networks from three-dimensional confocal microscopy. *Biophys. J.* 95:6072–6080.
36. Wolf, K., S. Alexander, ..., P. Friedl. 2009. Collagen-based cell migration models in vitro and in vivo. *Semin. Cell Dev. Biol.* 20:931–941.
37. Reference deleted in proof.
38. Li, Y., A. Asadi, ..., E. P. Douglas. 2009. pH effects on collagen fibrillogenesis in vitro: electrostatic interactions and phosphate binding. *Mater. Sci. Eng. C*. 29:1643–1649.
39. Kaufman, L. J., C. P. Brangwynne, ..., D. A. Weitz. 2005. Glioma expansion in collagen I matrices: analyzing collagen concentration-dependent growth and motility patterns. *Biophys. J.* 89:635–650.
40. Yang, Y. L., L. M. Leone, and L. J. Kaufman. 2009. Elastic moduli of collagen gels can be predicted from two-dimensional confocal microscopy. *Biophys. J.* 97:2051–2060.
41. Münster, S., and B. Fabry. 2013. A simplified implementation of the bubble analysis of biopolymer network pores. *Biophys. J.* 104:2774–2775.

## **Supporting Material**

### **Estimating the 3D pore size distribution of biopolymer networks from directionally biased data**

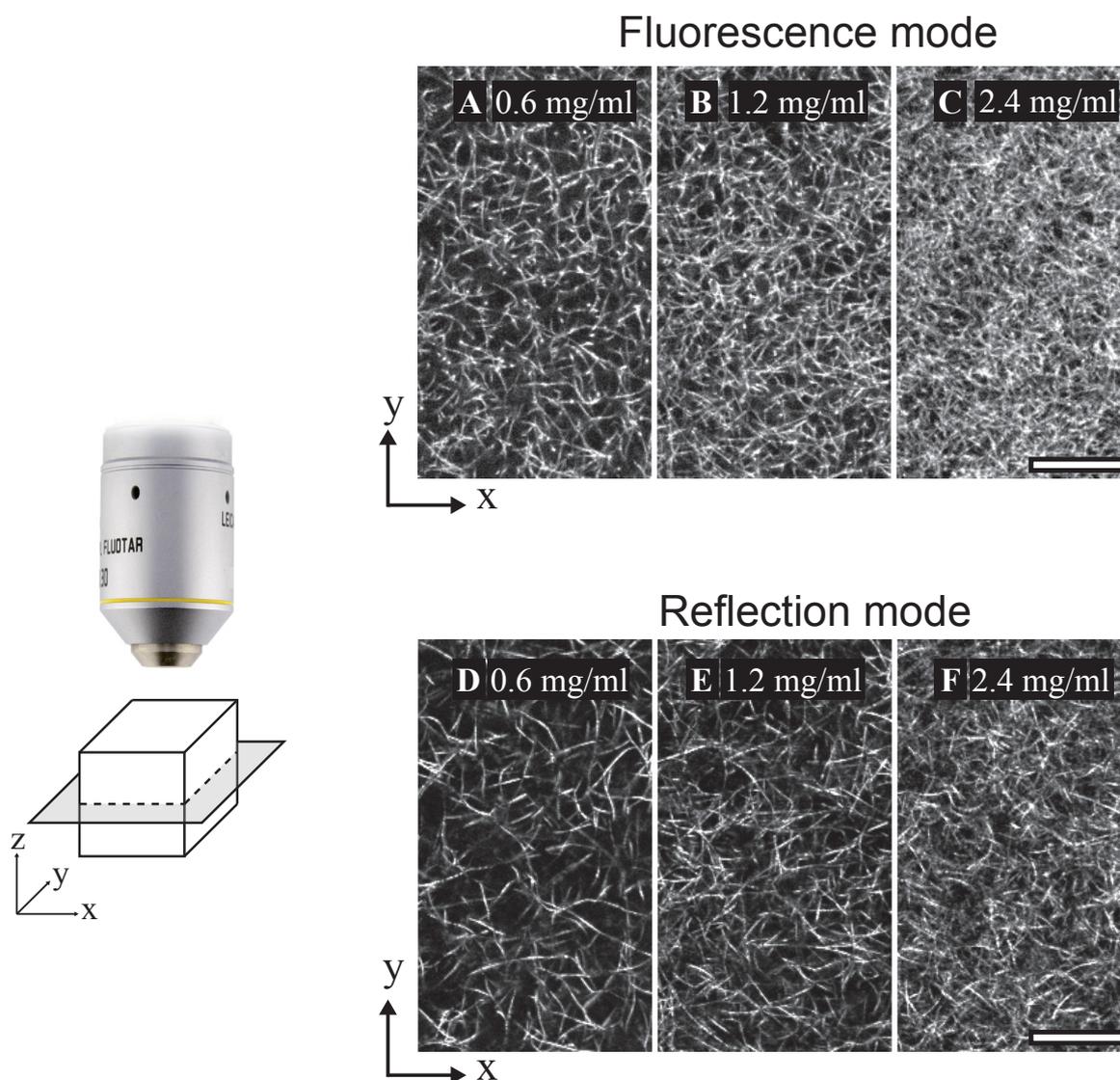
Nadine R. Lang, Stefan Münster, Claus Metzner, Patrick Krauss, Sebastian Schürmann, Janina Lange, Katerina E. Aifantis, Oliver Friedrich, Ben Fabry

Corresponding author:

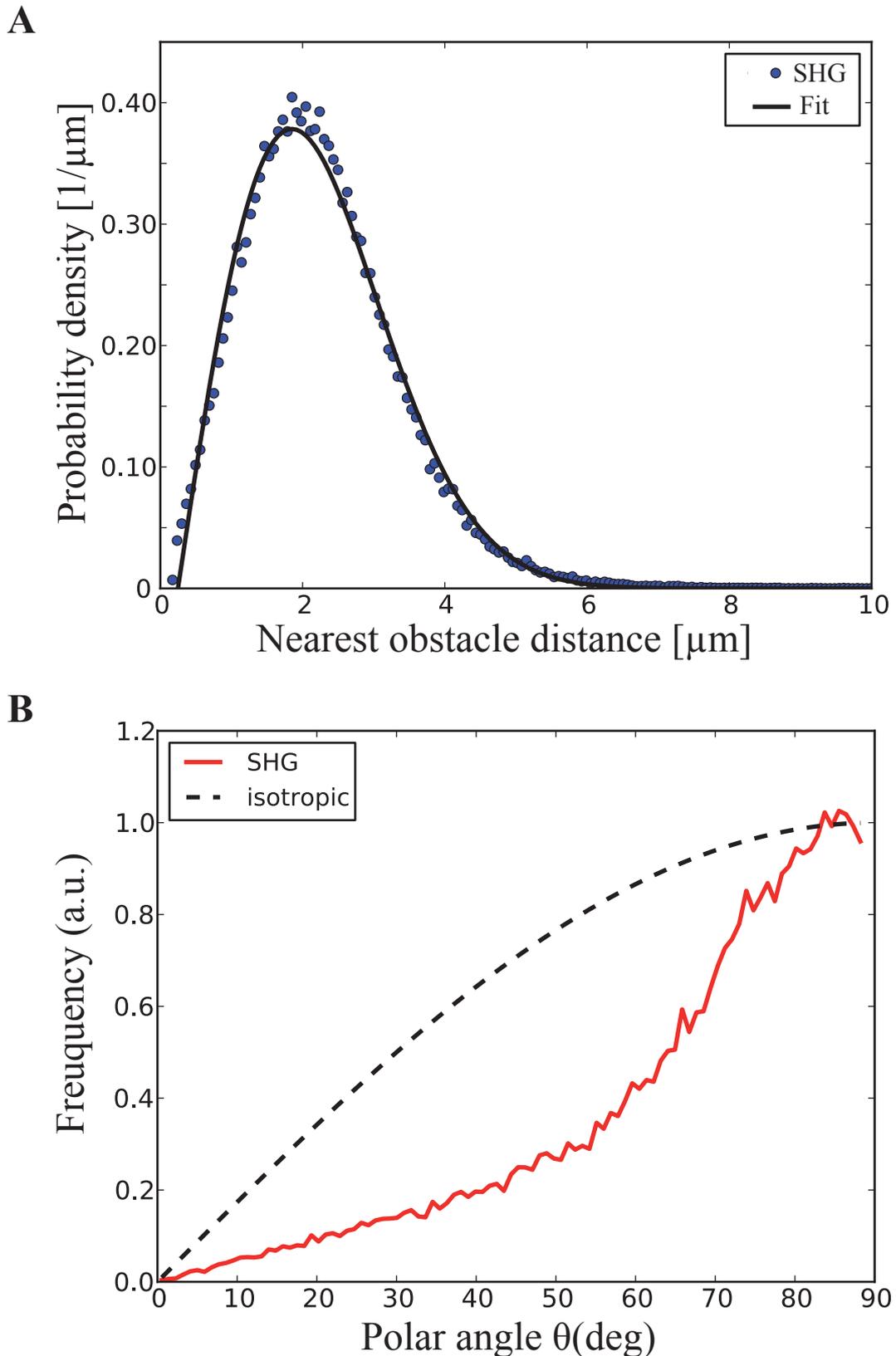
Ben Fabry  
Department of Physics  
University of Erlangen-Nuremberg  
Henkestr. 91  
91052 Erlangen  
phone: +49 9131 85 25610  
fax: +49 9131 85 25601  
[bfabry@biomed.uni-erlangen.de](mailto:bfabry@biomed.uni-erlangen.de)

## **SUPPLEMENTARY INFORMATION**

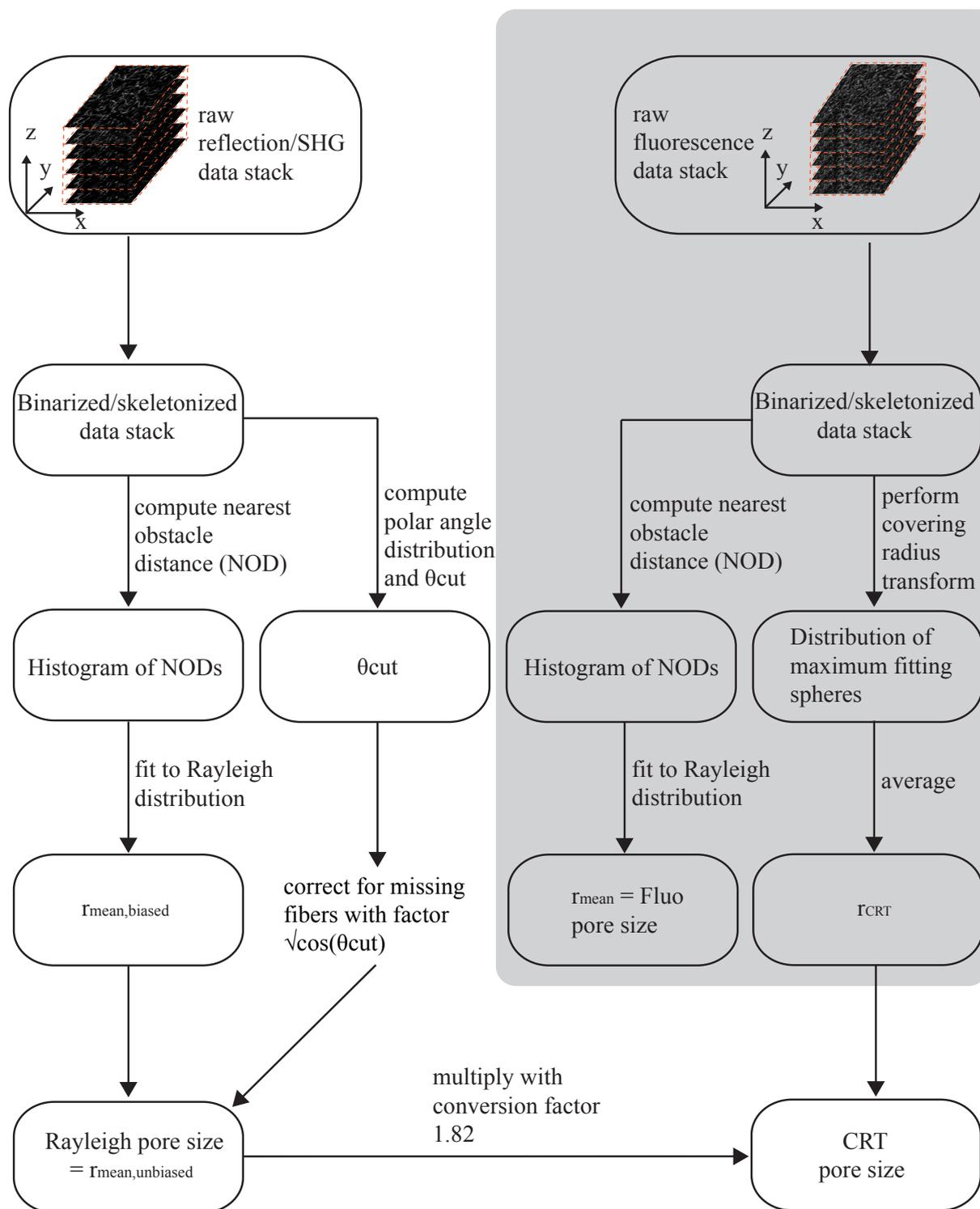
**Movie S1 Comparison of CRM and CFM.** Projected views of 3 x-y slices (total thickness 1  $\mu\text{m}$ ) (top) and 15 x-z slices (total thickness 4.75  $\mu\text{m}$ ) (bottom) of the same field-of-view of a 1 mg/ml fibrin network are imaged simultaneously in fluorescence mode and reflectance mode, and are switched back and forth in rapid succession to visualize missing fibers, similar to a blink comparator for astronomical images. Vertical fibers that can be seen in fluorescence mode are not detected in reflectance mode due to the blind spot effect.



**Fig. S1 Different concentrations of collagen monomers influences pore size.** X-y projections of CFM (A-C) and CRM (D-F) images (3 slices, total thickness 1  $\mu\text{m}$ ) from collagen gels polymerized at different monomer concentrations of 0.6 mg/ml (A,D), 1.2 mg/ml (B,E) and 2.4 mg/ml (C,F). With increasing concentrations, the pore sizes become smaller and the network denser. The scale bar of 20  $\mu\text{m}$  applies to all panels. Stacks were imaged with 512 x 512 pixels with a size of 317 nm x 317 nm x 335 nm. Fluorescence mode images appear denser compared to reflection mode images because of the blind spot effect.



**Fig. S2: NOD distribution and blind spot effect from SHG images of collagen.** (A) Distribution of the nearest obstacle distance of a 0.3 mg/ml collagen network from an image stack obtained with SHG follows a Rayleigh distribution with a very high fit quality of  $r^2=0.994$ ; same raw data as shown in Fig. 2C,D. (B) Distribution of the polar angles of the SHG signal from a 0.3 mg/ml collagen network shows systematic deviations from the  $\sin\theta$  dependency expected for an isotropic network. The polar angle distribution is similar to data obtained with CRM. The cut-off angle under SHG for a 40x NA=1.1 objective is  $48^\circ$ .



**Fig. S3: Data analysis flow chart.** Analysis steps in the shaded region from confocal fluorescence data are not required for obtaining the pore size distribution of networks and were only performed for validation of our method, which is based on confocal reflectance data.

## Theory of pore sizes in random line networks

In order to characterize the stochastic geometry of random line networks, a frequently used parameter is the average pore size  $r_{pore}$ . It is determined by finding, for a representative fraction of network pores, the largest spheres that can be fit into that pores and then computing the average of the radii of these maximum spheres. While this can be done numerically in a straight forward way, this definition of pore size is not suited very well for exact analytical calculations. Therefore, we suggest as an alternative measure the most probable nearest obstacle distance  $\sigma$  for randomly chosen test points and show that it is directly related to the pore size.

We start our investigations with a ‘Mikado’-like network model that has two parameters, the length  $l$  of the line segments and the volume density  $\rho$  of their centers. It is possible to compute the distribution  $p(r_{nod})$  of nearest obstacle distances analytically in this model. In the limit of zero line length, the Mikado model contains the case of point networks. More interesting is the opposite limit, where  $l$  is much larger than the average pore size. In this case, the Mikado model converges towards a more general model that represents any random line network with a large persistence length. Indeed, the single system parameter in this limiting case is the overall line density  $\lambda$ , i.e. the total line length per unit volume. This parameter only sets the spatial scale of the network, and no other details matter for the distributions  $p(r_{nod})$  or  $W(r_{pore})$ . For example, a network composed of random circles with identical overall line density would yield the same universal Rayleigh distribution  $p(r_{nod})$  as the Mikado model, provided the radius of the circles is much larger than the average pore size. We compare these analytical results to a numerical simulation that is directly based on the exact analytic geometry of points and lines, thus avoiding any possible artifacts arising from voxelation.

After demonstrating perfect agreement of the simulations with the analytic results, we use the simulations to determine the pore size distribution for line networks of various density parameters  $\lambda$ . As expected from scaling arguments, the average pore size  $r_{pore} = c\sigma$  is simply proportional to  $\sigma$ , allowing us to determine the conversion factor as  $c \approx 1.82$ .

### Distribution of nearest obstacles distances $p(r_{nod})$ , accessible volume fraction $Q(r)$ and pore sizes

We consider random biphasic networks, in which every point of 3-dimensional space either belongs to phase 0 (pore, liquid) or phase 1 (material, solid). In order to map out the stochastic geometry of the network, one can repeatedly choose a random point  $R_0=(x,y,z)$  within the 0-phase of the network and the find its ‘nearest obstacle distance’  $r_0(R_0)$ , defined as the Euclidean distance from that point  $R_0$  to the closest point of the 1-phase (Fig. S4(a)). The network is then characterized by the distribution  $p(r_{nod})$  of the nearest obstacle distances.

Closely related to  $p(r_{nod})$  is the ‘accessible volume fraction’  $Q(r)$ , defined as the fraction of the 0-phase in which a sphere of radius  $r$  (from now on called a  $r$ -sphere) could be centered without overlapping the 1-phase (Fig. S4(b)). In general, the dimensionless quantity  $Q(r)$  has the value  $Q(r=0)=1$  and decreases monotonically for all radii  $r>0$ .

The complementary quantity  $1 - Q(r)$  is the fraction of 0-phase for which an  $r$ -sphere overlaps the 1-phase. It corresponds to the probability that a random 0-phase point  $R_0$  has a nearest obstacle distance  $r_{nod}$  smaller than  $r$ , or

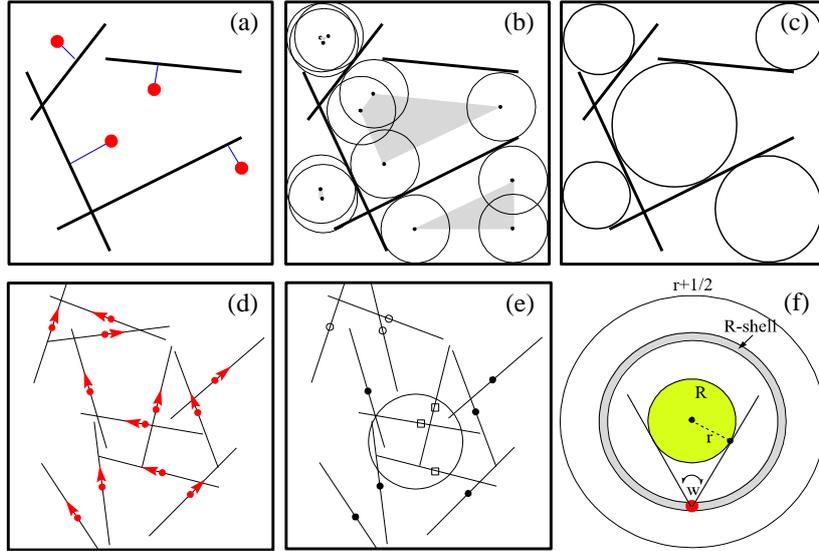
$$\begin{aligned} 1 - Q(r) &= Prob(r_{nod} < r) \\ &= \int_0^r p(r_{nod}) dr_{nod}. \end{aligned} \quad (1)$$

The derivative of this equation with respect to  $r$  shows that  $Q(r)$  is just the negative cumulative probability of  $p(r_{nod})$ :

$$p(r = r_{nod}) = -\frac{d}{dr}Q(r). \quad (2)$$

While both quantities carry the same information about the network, the cumulative  $Q(r)$  is more convenient for analytical considerations, as will be demonstrated below.

Another way to characterize pores of a network is to find the maximum sphere that fits to each pore and to define the ‘pore size’  $r_{pore}$  as the radius of this maximum sphere. The concept is also illustrated in Fig. S4(c). We denote the distribution of pore sizes by  $W(r_{pore})$ .



**Figure S4:** 2D illustration of various statistical measures used for networks of line segments. (a) Nearest obstacle distances  $r_{nod}$  (*thin lines*) for a few selected points (*circles*). (b) Accessible volume (shaded areas) for spheres of a given radius. (c) Maximum spheres fitting into network pores; thereby defining the pore sizes  $r_{pore}$ . (d) A homogeneous, isotropic random distribution of straight line segments. The segments have a prescribed length  $l$  and their center points a spatial density of  $\rho$ . (e) Classification of line segments in the 2D Mikado model. 1-group (*squares*): Centers within  $r$ -sphere. 2-group (*full circles*): Centers outside  $r$ -sphere but with partial overlap. 3-group (*empty circles*): Remote segments without overlap. (f) 2D illustration of an  $r$ -sphere (*green*), a concentric spherical shell of radius  $R$  (*gray*) and a specific point (*red*) within this shell. From all line segments centered at the red point, only those can intersect the  $r$ -sphere with orientations falling into a cone of apex angle  $\omega$ .

## Random Line Networks: The Mikado model

In the following, we consider random networks in which the 1-phase consists of straight line segments of fixed length, with isotropic orientations and a homogeneous distribution throughout the 3D volume. We refer to this model as the Mikado model.

Each individual line segment (LS) can be described by its center point and a unit direction vector. In order to avoid ambiguities, we require that all unit vectors have a positive z-component and thus ‘point upwards’ (Fig. S4(d)). The two parameters of the Mikado model are the length  $l$  of the LSs and the volume density  $\rho = \frac{N}{V}$  of their center points, where  $N$  is the number of line segments within a volume  $V$ .

Consider first the extreme case  $l \rightarrow 0$ , where all LSs degenerate into their center points. The configuration of LS-centers throughout the volume is a spatial Poisson process with ‘event rate’ that is identical to the volume density  $\rho$ . On average, a randomly placed r-sphere will contain a number of

$$n_{av,l \rightarrow 0}(r) = \rho \frac{4}{3} \pi r^3 \quad (3)$$

LS-centers. The probability  $Q(r)$  that not a single LS-center lies within the r-sphere is given by the Poisson probability for  $k=0$  events, which is

$$Q(r) = \text{Poisson} \{k=0, n_{av} = n_{av,l \rightarrow 0}(r)\} = e^{-n_{av,l \rightarrow 0}(r)} \quad (4)$$

Therefore, in the case of the random point network, the accessible volume fraction is given by

$$Q(r)_{l \rightarrow 0} = e^{-\frac{4\pi}{3} \rho r^3} \quad (5)$$

We now turn back to the general case  $l > 0$ . As before, we can write

$$Q(r) = e^{-n_{av}(r)} \quad (6)$$

In order to compute  $n_{av}(r)$ , we note that with respect to a given r-sphere, the LSs can be classified into 3 groups (Fig. S4 (e)):

- 1-group with LS-centers inside the r-sphere.
- 2-group with LS-centers outside the r-sphere, but with partial overlap of the fibers.
- 3-group with LS-centers outside the r-sphere and with no fiber overlap.

Only the groups 1 and 2 contribute to  $n_{av}(r)$ . The contribution of the 1-group is identical to the case of point networks above:

$$n_{av,1}(r) = \rho \frac{4}{3} \pi r^3. \quad (7)$$

Fibers belonging to group 2 consists of LSs with centers in a sphere of radius  $r + (l/2)$  around the center of the r-sphere. We now consider in more detail the fibers in an infinitesimal

spherical shell of radius  $R$  around the center of the  $r$ -sphere, with  $0 < R < r + (l/2)$ . This  $R$ -shell contains a number of

$$dN' = \rho dV = \rho 4\pi R^2 dR \quad (8)$$

candidates for intersection. Among them, only those LSs will actually overlap the  $r$ -sphere that have orientations within a certain cone (Fig. S4(f)). This cone has an apex angle of  $\omega = 2 \arcsin(r/R)$  and the corresponding solid angle is

$$\begin{aligned} \Omega(R) &= 4\pi \sin^2\left(\frac{\omega}{4}\right) \\ &= 4\pi \left[ \sin\left(\frac{1}{2} \arcsin\left(\frac{r}{R}\right)\right) \right]^2 \\ &= 2\pi \left(1 - \sqrt{1 - \left(\frac{r}{R}\right)^2}\right). \end{aligned} \quad (9)$$

Since the total solid angle available for LS orientations is  $\Omega_{tot} = 2\pi$  (according to our convention that all unit direction vectors are pointing upward), the intersecting LSs amount to a fraction of  $\Omega(R)/\Omega_{tot} = \left(1 - \sqrt{1 - \left(\frac{r}{R}\right)^2}\right)$ . We conclude that the average number of actual intersections from LSs within the  $R$ -shell is

$$\begin{aligned} dN(R) &= dN' \left(1 - \sqrt{1 - \left(\frac{r}{R}\right)^2}\right) \\ &= 4\pi\rho \left(1 - \sqrt{1 - \left(\frac{r}{R}\right)^2}\right) R^2 dR. \end{aligned} \quad (10)$$

The total contribution from all LSs of group 2 fibers is obtained by integration over the relevant  $R$ -shells:

$$n_{av,2}(r) = \int_{R=r}^{R=r+\frac{l}{2}} dN(R). \quad (11)$$

This integral can be performed analytically. Using the abbreviation

$$f(s) := \frac{1}{3} \left[ s^3 - (s^2 - 1)^{3/2} \right], \quad (12)$$

one obtains

$$n_{av,2}(r) = 4\pi\rho r^3 \left[ f\left(1 + \frac{l}{2r}\right) - f(1) \right] \quad (13)$$

By adding the contributions of both relevant groups,  $n_{av}(r)=n_{av,1}(r) + n_{av,2}(r)$ , and using  $Q(r) = \exp(-n_{av}(r))$ , we arrive at an analytic expression for the accessible volume fraction in the Mikado model. Defining another useful abbreviation

$$g(x) := 3 \left[ f \left( 1 + \frac{x}{2} \right) - \frac{1}{3} \right], \quad (14)$$

the result can be cast into the form

$$Q(r) = e^{-\frac{4\pi}{3}\rho r^3 [1+g(\frac{l}{r})]}. \quad (15)$$

It correctly contains the limit of point networks, since  $g(l/r) \rightarrow 0$  for  $l \rightarrow 0$ . All the differences between point and LS networks are included in the ‘perturbation function’  $g(l/r)$ .

From the accessible volume fraction  $Q(r)$ , we immediately obtain the distribution of nearest obstacle distances  $p(r_{nod} = r) = -d/dr Q(r)$  in the Mikado model. With increasing  $r_{nod}$ , this distribution starts with  $p(r_{nod} = 0) = 0$ , develops a single peak and then decays exponentially for distances much larger than the average pore size  $R_{pore}$  of the network.

#### Mikado model in the long fiber limit

We next consider the case  $l \gg r$ , where the LSs are much longer than the typical distances of interest. Since  $p(r_{nod})$  is exponentially small for distances beyond the average pore size, this limit can also be interpreted as  $l \gg R_{pore}$ . Note that this is a typical situation for networks of semi-flexible fibers, such as collagen.

It is straight-forward to show that in this limit the perturbation function diverges

as  $g\left(\frac{l}{r}\right) \rightarrow \frac{3l}{4r}$ . One therefore obtains

$$Q_{l \gg r}(r) = e^{-(\pi\rho l)r^2} = e^{-\frac{1}{2}\left(\frac{r}{\sigma}\right)^2} \quad (16)$$

which is the ‘right half’ of a Gaussian bell curve with standard deviation

$$\sigma = 1/\sqrt{2\pi\rho l}. \quad (17)$$

The corresponding distribution of the nearest obstacle distances is a Rayleigh distribution

$$p_{l \gg r}(r) = \frac{r}{\sigma^2} e^{-\frac{1}{2}\left(\frac{r}{\sigma}\right)^2}. \quad (18)$$

The most probable nearest obstacle distance, i.e. the value of  $r$  at which  $p(r)$  is maximum, is given by  $\sigma$ . We note that the accessible volume fraction in Eq.(16) depends only on the ratio  $r/\sigma$ . Therefore, all nearest obstacle distance distributions  $p_{l \gg r}(r)$  should collapse onto a universal distribution when the distance  $r$  is measured in units of  $\sigma$ . In the long fiber limit, a dense and a dilute Mikado network cannot be distinguished from each other, if the spatial scale is unknown.

### Relating $\sigma$ to line density

It is remarkable that in the long fiber limit of the Mikado model, the properties of the network are completely determined by the parameter combination  $\rho l$ , which appears in the quantity  $\sigma = 1/\sqrt{2\pi\rho l}$ .

Recalling the definition of  $\rho$  as the volume density of LS centers, we can write

$$\rho l = \frac{N}{V} l = \frac{L_{tot}}{V} =: \lambda, \quad (19)$$

where  $L_{tot}$  is the total length of all LSs. The new density parameter  $\lambda$  corresponds to the total ‘fiber’ length per unit volume. It follows that

$$\sigma = 1/\sqrt{2\pi\lambda}. \quad (20)$$

### Numerical test of the Mikado model

In order to test the predictions of the Mikado model, we have simulated random line networks and compared the resulting numerical  $p(r_{nod})$  with the analytical results above.

In the simulation, each line segment (of constant length  $l$ ) was numerically represented by its center coordinates and a unit direction vector, as depicted schematically in Fig. S4(d). Initially, a list of  $N$  such line objects was generated, with the center points distributed randomly throughout a cubic simulation box of linear dimension  $L$  (with homogenous density  $\rho = \frac{N}{V} = \frac{N}{L^3}$ ) and with random, isotropic direction vectors (more precisely, in order to avoid boundary effects, we extended the simulation box on each side by  $l/2$  and distributed a correspondingly larger number of  $N^* = N(L + l)^3 / L^3$  line centers within this extended box.).

The distribution  $p(r_{nod})$  was determined by randomly choosing  $K = 10^5$  test points  $R_{k=1\dots K}$  within the simulation box, finding the nearest obstacle distance  $r_{nod}(R_k)$  for each test point and then computing a histogram of these distances. The distance  $r_{nod}(R_k)$  is found by first computing the distances  $r_{kn}$  between test point  $R_k$  and all the lines  $n$  of the network and then finding the smallest of those values. Note that the distance  $r_{kn}$  between a point and a line segment can be obtained exactly (without any ‘voxelization’ required).

For the numerical test of the Mikado model in the long fiber limit, we prescribed the density parameter  $\lambda$ , set  $L = l = 1$  and computed the required number of fibers as  $N = \frac{\lambda L^3}{l}$ . We found an excellent agreement between the analytical prediction and the simulation (Fig. S5(a)).

## Relationship between the most probable nearest obstacle distance $\sigma$ and the average pore size

$r_{pore}$

For any real network, it is possible to compute the nearest obstacle distance  $r_{nod}(R_0)$  for each spatial point  $R_0$ , resulting in a so-called ‘Euclidean distance map’ (EDM). The pore centers of the network can then be defined as the positions  $R_0=R_{max}^{(i)}$  of the local maxima of the EDM and the pore size distribution  $W(r_{pore})$  is the distribution of the distance values  $r_{pore}^{(i)}=r_{nod}(R_{max}^{(i)})$  taken at these local maxima (1).

Based on our numerically exact simulation of random line networks we have computed the pore size statistics  $W(r_{pore})$  and compared it to the corresponding distribution  $p(r_{nod})$  of nearest obstacle distances. As expected,  $W(r_{pore})$  is peaked at a larger value than  $p(r_{nod})$  (Fig. S5(c)). In the long fiber limit, the ratio  $r_{pore}/\sigma$  between the average pore size and the most probable obstacle distance is a constant, i.e. independent from the density parameter  $\lambda$  of the network. This follows from the fact that the distribution  $p(r_{nod})$  is universal in length units of  $\sigma$ . To demonstrate the constant ratio, we have plotted  $r_{pore}(\lambda)$  and  $\sigma(\lambda)$  double-logarithmically (Fig. S5(d)).

### Summary

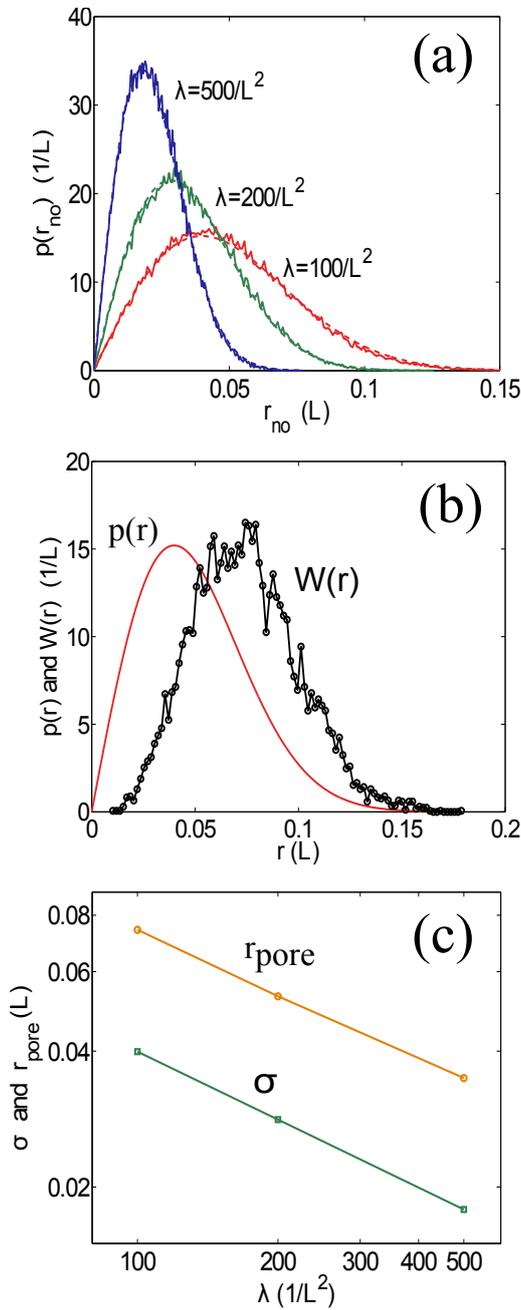
Here, we have theoretically explored random line networks, modelled as isotropic and macroscopically homogeneous distributions of straight line segments in 3D space. In the limiting case when the line segments are much longer than the average pore size  $r_{pore}$ , the distances  $r_{nod}$  of random test points to the nearest line segment are distributed according to a Rayleigh distribution

$$p(r_{nod}) = \frac{r_{nod}}{\sigma^2} e^{-\frac{1}{2}\left(\frac{r_{nod}}{\sigma}\right)^2}. \quad (21)$$

The most probable distance  $\sigma$  (peak position of the distribution) is determined by the overall line density  $\lambda$ , i.e. the total line length per unit volume, by

$$\sigma = 1/\sqrt{2\pi\lambda}. \quad (22)$$

The average pore size  $r_{pore}$ , defined as the radii of maximum spheres fitting into the pores, is proportional to  $\sigma$ , with  $r_{pore} \approx 1.82\sigma$ .



**Figure S5:** (a) Distribution of nearest obstacle distances in 3D networks of straight line segments, for three different density parameters  $\lambda$ . Analytical predictions of the Mikado model in the long fiber limit (dashed lines) are compared to numerical simulations (solid lines). The unit of length was set equal to the size of  $L$  of the simulation box, which in turn was equal to the length  $l$  of the line segments. (b) Distribution of the nearest obstacle distances  $p(r)$  (red line) and pore size distribution  $W(r)$  (black line with symbols) in a 3D network of straight line segments, for a density parameter of  $\lambda = 100/L^2$ . (c) Most probable obstacle distance  $\sigma$  (squares) and average pore size  $r_{pore}$  (circles) as a function of the density parameter  $\lambda$ . In the long fiber limit, the ratio is constant with  $r_{pore}/\sigma \approx 1.82$

1. Münster, S., and B. Fabry. 2013. A Simplified Implementation of the Bubble Analysis of Biopolymer Network Pores. *Biophys. J.* 104: 2774–2775.