Extracting poresize statistics of collagen networks from confocal images

Claus Metzner, Patrick Krauss, Stefan Münster, ..., Ben Fabry*
Biophysics Group, University of Erlangen, Henkestr.91, D-91052 Erlangen, Germany
(Dated: May 4, 2011)

We present a method to extract and characterize the statistical geometric properties of semiflexible random biopolymer networks, such as collagen gels, from confocal 3D image stacks. Our method is particularly useful for microscope images obtained in the reflection mode, where a part of fibers with steep angles is invisible. From a properly binarized image stack (defining the solid and liquid phase), we first determine the distribution of nearest obstacle distances $P_{nod}(r)$ and fit it to an analytical model of random line networks. Based on the model, we then correct the distribution for the fraction of insible fibers. The full 3D distribution can also be reconstructed from a 2D slice of the 3D image stack. We demonstrate the feasability of the invisible fiber correction by comparing the statistics of the same gel measured in the relection and fluorescence mode.

Keywords: collagen, poresizes, confocal microscopy, fiber networks

I. APPENDIX

A. Distribution of nearest obstancles distances $p(r_{no})$, accessible volume fraction $Q(r)$ and poresizes

We consider in the following 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, we repeatedly choose a random point $\vec{R}_0 = (x, y, z)$ within the 0-phase of the network and then find its 'nearest obstacle distance' $r_{no}(\vec{R}_0)$, defined as the Euklidean distance from that point $\vec{R}_0$ to the closest point of the 1-phase (compare Fig.1)(a). The network is then characterized by the distribution $p(r_{no})$ of the nearest obstancle distances.

Closely related to $p(r_{no})$ 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 (compare Fig.1(b)). In general, the dimensionless quantity $Q(r)$ has the value $Q(r=0)=1$ and decreases monotonically for all radii $r>0$.

The complemental 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 $\vec{R}_0$ has a nearest obstacle distance r_{no} smaller than r, or

$$\begin{aligned} 1-Q(r) &= \mathrm{Prob}(r_{no} < r) \\ &= \int_0^r p(r_{no}) dr_{no}. \end{aligned} \tag{1}$$

The derrivative of this equation with respect to r shows that $Q(r)$ is just the negative cumulative probability of $p(r_{no})$:

$$p(r = r_{no}) = -\frac{d}{dr} Q(r). \tag{2}$$

While both quantities carry the same information about the network, the cumulative $Q(r)$ is more convenient for analytical considerations, as will be demonstated 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 maximium sphere. The concept is also illustrated in Fig.1(c). We denote the distribution of poresizes by $W(r_{pore})$.

B. 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' (compare Fig.2). 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.

Consider first the extreme case $l \to 0$, where alle LSs degenerate into their center points, and place a r-sphere randomly into the system. Note that the configuration of LS-centers throughout the volume is a spatial Poisson process with 'event rate' ρ. On average, the r-sphere will contain a number of

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

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

$$Q(r) = \mathrm{Poisson}\{k=0, n_{av} = n_{av,l\to 0}(r)\} = e^{-n_{av,l\to 0}(r)}. \tag{4}$$

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

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

* bfabry@biomed.uni-erlangen.de

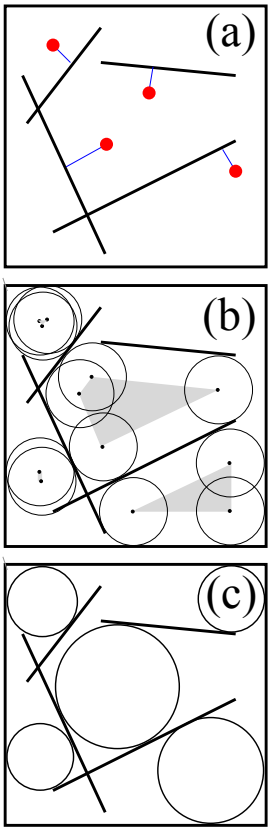

FIG. 1. Illustration of various statistical measures for 2D network of line segments (thick lines, also including system borders). (a) Nearest obstacle distances r_{no} (thin lines) for a few selected points (circles). (b) Accessible volume (shaded areas) for spheres of a given radius. (c) Maximum spheres fitting into network pores, thereby defining the pore sizes r_{pore}.

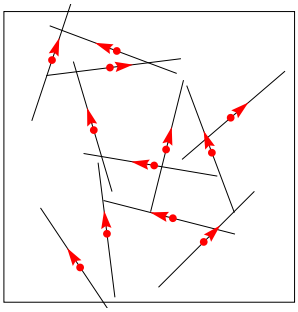

FIG. 2. Sketch of the Mikado model in 2D: A homogeneous, isotropic random distribution of straight line segments. The segments have a prescribed length l and their center points a spatial density of ρ.

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

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

In order to compute $n_{av}(r)$, we note that with respect

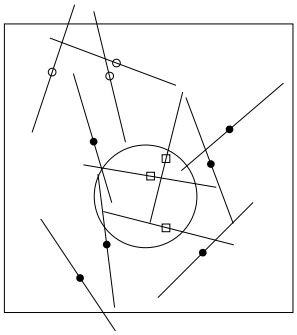

FIG. 3. Classification of line segments in the 2D Mikado model. 1-group (squares): Centers within r-sphere. 2-group (full circles): Centers outside r-sphere, yet with chance of overlap. 3-group (empty circles): Remote segments without chance of overlap.

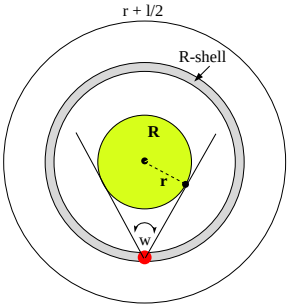

FIG. 4. 2D sketch of an r-sphere (green), a concentric spherical shell of radius R (gray) and a specific point (red) within this shell. From all line segments centered at the red point, only those can intersect the r-sphere with orientations falling into a cone of apex angle ω.

to a given r-sphere, the LSs can be classified into 3 groups (compare Fig.3):

- 1-group with LS-centers inside the r-sphere.
- 2-group with LS-centers outside the r-sphere, but yet with a possibility of intersecting the r-sphere.
- 3-group with LS-centers too far away to touch the r-sphere.

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

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

The 2-group 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 ones in an infinitesimal sperical shell of radius R around the center of the r-sphere, with $0<R<r+(l/2)$. This R-shell contains a number of

$$dN' = \rho dA = \rho 4\pi R^2 dR \tag{8}$$

candidates for intersection. Among them, only those LSs will actually overlap the r-sphere that have orientations within a certain cone (compare Fig.4). This cone has an apex angle of $\omega = 2\arcsin(r/R)$ and the corresponding solid angle is

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

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 $\Omega(R)/\Omega_{tot} = \left(1-\sqrt{1-(r/R)^2}\right)$. We conclude that the average number of actual interections from LSs within the R-shell is

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

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

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

This integral can be performed analytically. Using the abbreviation

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

one obtaines

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

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

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

the result can be cast into the form

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

It correctly contains the limit of point networks, since $g(l/r) \to 0$ for $l \to 0$. 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 $p(r_{no} = r) = -\frac{d}{dr}Q(r)$ in the Mikado model. With increasing r_{no}, this distribution starts with $p(r_{no}=0)=0$, developes a single peak and then decays exponentially for distances much larger than the average pore size $\overline{R}_{pore}$ of the network.

C. 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 $p(r_{no})$ is exponentially small for distances beyond the average pore size, this limit can also be interpreted as $l \gg \overline{R}_{pore}$. Note that this is a typical situation for networks of semiflexible fibers, such as collagen.

It is straight-forward to show that in this limit the perturbation function diverges as $g(l/r) \to \frac{3}{4}\frac{l}{r}$. One therefore obtains

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

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

$$\sigma = 1/\sqrt{2\pi\rho l}. \tag{17}$$

The corresponding nearest obstacle distribution is a Rayleigh distribution

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

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

D. Computing σ from experimental parameters

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 ρl, which appears in the quantity $\sigma = 1/\sqrt{2\pi\rho l}$.

Remembering the definition of ρ as the volume density of LS centers, we can write

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

where L_{tot} ist the total length of all LSs. The new density parameter λ corresponds to the *total 'fiber' length per unit volume*. It follows that

$$\sigma = 1/\sqrt{2\pi\lambda}. \tag{20}$$

Summing up, the Mikado model predicts that for sufficiently straight and long fibers, the distribution of nearest obstancle distances, $p(r_{no})$, is a Rayleigh distribution. The most probable distance, $r_{no}^{peak} = \sigma$, can be directly computed from the density parameter λ.

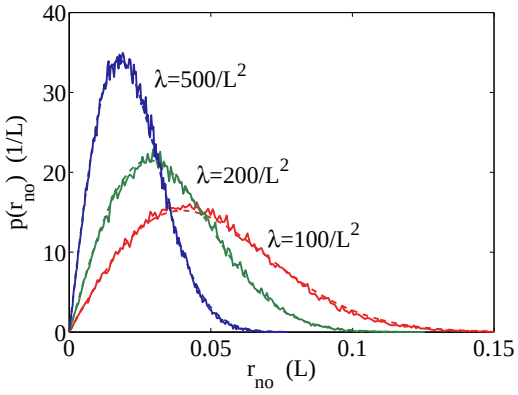

FIG. 5. Distribution of nearest obstacle distances in 3D networks of straight line segments, for three different density parameters λ. Analytical predictions of the Mikado model in the long fiber limit (dashed lines) are compared to numerical simulations (solid lines). The unit of length was set equal to the linear size L of the simulation box, which in turn was equal to the length l of the line segments.

E. 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 $p(r_{no})$ 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.2. Initially, a list of N such line objects was genenerated, with the center points distributed randomly throughout a cubic simulation box of linear dimension L (with homogeneous density $\rho = \frac{N}{V} = \frac{N}{L^3}$) and with random, isotropic direction vectors [1].

The distribution $p(r_{no})$ was determined by randomly choosing $K = 10^5$ test points $\vec{R}_{k=1...K}$ within the simulation box, finding the nearest obstacle distance $r_{no}(\vec{R}_k)$ for each test point and then computing a histogram of these distances. The distance $r_{no}(\vec{R}_k)$ is found by first computing the distances r_{kn} between test point $\vec{R}_k$ 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 λ, 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 (compare Fig.5).

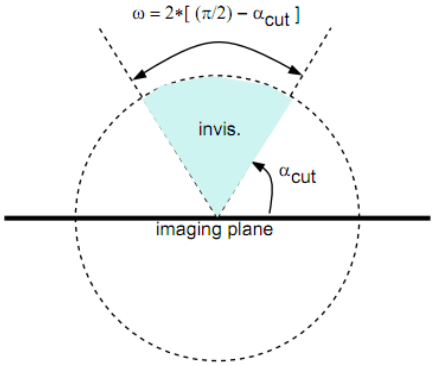

FIG. 6. In confocal reflection microscopy, only fiber segments with direction angles smaller than a cut-off angle α_{cut} are detected, leaving a cone of apex angle ω invisible.

F. Correction for invisible fibers

In confocal reflection microscopy, only fiber segments with directions up to a cut-off angle α_{cut} can be detected (Here, α_{cut} is measured from the image plane as illustrated in Fig.6). This leaves invisible a cone of apex angle ω, corresponding to a solid angle of $\Omega_{inv} = 2\pi\,(1 - cos(\omega/2))$. Compared to all possible fiber directions (using our convention that all unit direction vectors are pointing upward), this amounts to a fraction

$$\Omega_{inv}/\Omega_{tot} = (1 - cos(\omega/2))\,. \tag{21}$$

of invisible fiber segments.

Consequently, if a density parameter λ_{app} is estimated for a given network by fitting Eq.(18) to the measured $p(r_{no})$, this apparent density will be smaller than the actual density λ by a factor $1 - \frac{\Omega_{inv}}{\Omega_{tot}}$, or

$$\lambda = \frac{\lambda_{app}}{1 - (\Omega_{inv}/\Omega_{tot})} = \frac{\lambda_{app}}{\cos(\omega/2)}. \tag{22}$$

Using Eq.(20), a similar relation can be derived between the apparent and actual peak value σ of the distribution $p(r_{no})$:

$$\sigma = \sigma_{app}\sqrt{\cos(\omega/2)}. \tag{23}$$

G. Relation between the most probable obstacle distance σ and the average pore size $\overline{r}_{pore}$

For any concrete network, it is possible to compute the nearest obstacle distance $r_{no}(\vec{R}_0)$ for each spatial point $\vec{R}_0$, resulting in a so-called 'Euklidean distance map' (EDM). The pore centers of the network can then be defined as the positions $\vec{R}_0 = \vec{R}^{(i)}_{max}$ of the local maxima of the EDM and the pore size distribution $W(r_{pore})$ is the distribution of the distance values $r^{(i)}_{pore} = r_{no}(\vec{R}^{(i)}_{max})$ taken at these local maxima.

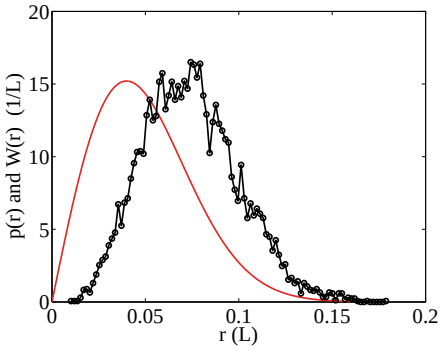

FIG. 7. Distribution of nearest obstacle distances $p(r)$ (line) and pore size distribution $W(r)$ (line with symbols) in a 3D network of straight line segments, for a density parameter of $\lambda = 100/L^2$.

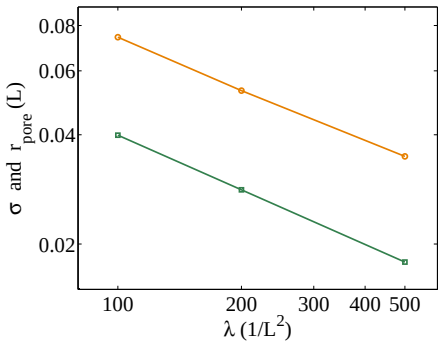

FIG. 8. Most probable obstacle distance σ (squares) and average pore size r_{pore}^{av} (circles) as a function of the density parameter λ. In the long fiber limit, the ratio is constant with $r_{pore}^{(av)}/\sigma \approx 1.86$.

Based on our numerically exact simulation of random line networks, as described in Sect.I E, we have computed the pore size statistics $W(r_{pore})$ and compared it to the correponding distribution $p(r_{no})$ of nearest obstacle distances. As expected, $W(r_{pore})$ is peaked at a larger value than $p(r_{no})$ (compare Fig.7). In the long fiber limit, the ratio $r_{pore}^{(av)}/\sigma$ between the average pore size and the most probable obstacle distance is a constant, i.e. independent from the density parameter λ of the network. This follows from the fact that the distribution $p(r_{no})$ is universal in length units of σ. To demonstrate the constant ratio, we have plotted $r_{pore}^{(av)}(\lambda)$ and $\sigma(\lambda)$ double-logarithmically (compare Fig.8).

H. Finite thickness of fibers

In the case of a real polymer network, the fibers are not lines but, in a more accurate approximation, cylinders of finite radius r_f. The total volume of the 1-phase is then given by

$$V_1 = \pi r_f^2 L_{tot} \tag{24}$$

and we find that

$$\lambda = \frac{1}{\pi r_f^2}\frac{V_1}{V} = \frac{\nu_1}{\pi r_f^2}, \tag{25}$$

where $\nu_1 = \frac{V_1}{V}$ is the volume fraction of the 1-phase.

ACKNOWLEDGMENTS

This work was supported by grants from Deutsche Forschungsgemeinschaft.

---

[1] 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.