

# An algorithm for extracting the network geometry of three-dimensional collagen gels

ANDREW M. STEIN\*, DAVID A. VADER†, LOUISE M. JAWERTH‡, DAVID A. WEITZ†, LEONARD M. SANDER‡

\*Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, MN 55403, USA, +astein@ima.umn.edu+

†School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA

‡Michigan Center for Theoretical Physics and Department of Physics, University of Michigan, Ann Arbor, Michigan 48109, USA

**Key words:** 3d reconstruction, biopolymer network, collagen gel, confocal microscopy, image processing.

## Summary

The geometric structure of a biopolymer network impacts its mechanical and biological properties. In this paper, we develop an algorithm for extracting the network architecture of three-dimensional (3d) fluorescently labeled collagen gels, building on the initial work of Wu *et al.*, (2003). Using artificially generated images, the network extraction algorithm is then validated for its ability to reconstruct the correct bulk properties of the network, including fiber length, persistence length, cross-link density, and shear modulus.

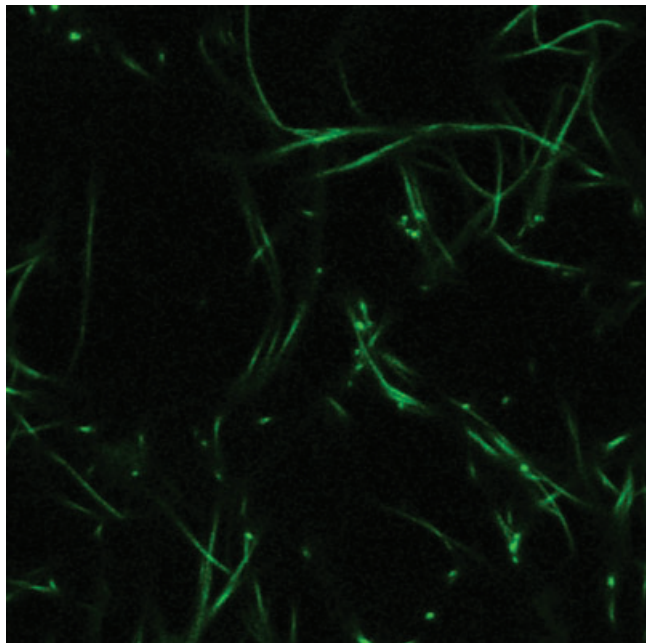
## Introduction

Biopolymer networks pervade living systems. Actin networks are critical to cell motility, microtubule networks provide structural support for the cell cytoskeleton, collagen networks are a primary component of the extracellular matrix, and fibrin is an important component of blood clot formation. Understanding the mechanical and structural properties of biopolymer networks is critical in understanding their biological role. Within the last 10 years, much work has been done on understanding how the mechanical properties of a network are influenced by its microstructure Chandran & Barocas (2006; Head *et al.*, 2003a, 2005; Heussinger *et al.*, 2007; Huisman *et al.*, 2007; Stylianopoulos & Barocas 2007; Van Dillen *et al.*, 2006; Wilhelm & Frey 2003). This work is mainly theoretical and is based on networks that are formed by randomly placing sticks in two or three dimensions and introducing a cross-link whenever two sticks overlap. With the exception of Chandran & Barocas (2006), which includes an analysis of a real collagen network that was characterized by hand, there has been no validation that these artificial networks accurately model the true network architecture. To

apply this theory to experimental data, it is critical that we measure the reticular properties of biopolymers such as the fiber lengths, angles and cross-link locations.

A 2d slice obtained by confocal microscopy of a fluorescently labeled collagen-I gel is shown in Figure 1. The complexity of fiber networks makes quantification of the geometry challenging. While software tools have been developed for helping humans to identify the network architecture (Baradet *et al.*, 1995), to the author's knowledge, there has been only one automated algorithm developed for this task (Wu *et al.*, 2003) and there has been very little validation that this algorithm accurately measures the network architecture. The algorithm by Wu *et al.*, as discussed in more detail below, skeletonizes the image by tracing along the medial axis of the fibers. Nissler *et al.*, (2007) have also used skeletonization for quantifying 2d projections of 3d networks, but they are unable to identify 3d network architecture directly. A variety of 3d skeletonization algorithms for vascular networks are reviewed by Boskamp *et al.*, (2005) and Suri (2002), but to our knowledge, these algorithms have not been validated for polymer networks. An important difference between vascular and polymer networks is that vascular networks are frequently acyclic whereas polymer networks are not. There also exist in the literature many 2d algorithms for extracting network architectures, (Can *et al.*, 1999; Tupin *et al.*, 1998; Stoica *et al.*, 2004). In these papers, the original image is initially correlated with a set of filters that are designed to detect line features. While such methods can be generalized to 3d using steerable filters (Aguet *et al.*, 2005), doing so is nontrivial and to our knowledge, the use of such code to extract 3d networks has only been briefly mentioned (Aguet *et al.*, 2005). Other related work includes the development of beamlets (Donoho & Huo 2002), where the focus is on identifying a single fiber in a noisy image rather than on the identification of an entire network.

Correspondence to: Andrew M. Stein +astein@ima.umn.edu+



**Fig. 1.** A fluorescently labeled  $25\ \mu\text{m} \times 25\ \mu\text{m}$  slice through a collagen gel, approximately  $1\ \mu\text{m}$  thick, obtained by confocal microscopy.

The one existing algorithm for identifying 3d collagen networks (Wu *et al.*, 2003) has been verified in only a limited way. They have shown that their algorithm faithfully matches the angle distribution of the fibers, but they have not tested their algorithm's ability to accurately identify fiber length, persistence, or cross-link location. Furthermore, this algorithm has only been used on images acquired by reflectance microscopy and it has recently been observed that this imaging modality misses many fibers that are oriented perpendicular to the focal plane (results in preparation). In this paper, a new algorithm for FIBer Extraction (FIRE) is presented. This algorithm is then tested on actual collagen networks and on artificially generated images of known structure and both the geometric and mechanical properties of these networks are compared to experimental measurement.

## Materials and Methods

### Collagen labeling

We labeled bovine collagen type I (Inamed Biomaterials, Fremont, CA) with the fluorescent molecule TAMRA (Invitrogen, Carlsbad, CA), using a protocol closely derived from Baici *et al.*, (1980). Briefly, we injected 0.75 mL of 5.9 mg/mL unlabeled collagen (Inamed Biomaterials) into a 3 mL dialysis cassette (Pierce, Rockford, IL) and dialyzed it against a 0.25 M sodium bicarbonate (Sigma, St. Louis, MO) buffer containing 0.4 M sodium chloride (Sigma), and a pH adjusted to 9.5 through the addition of 10 M sodium hydroxide (Sigma). An equal volume of TAMRA solution was prepared

from the same buffer, with a final molar concentration of dye 20 times that of collagen. The dye solution was then injected into the cassette, which was removed from the dialysis buffer; we left the dye to diffuse for 12–24 hrs. The collagen+TAMRA solution was then dialyzed once more against a solution of 0.2% (v/v) acetic acid (Sigma) in deionized water, with pH adjusted to 4.0. The concentration of dyed collagen stock is about 3.0 mg/mL.

### Collagen gel preparation

Collagen solution at final concentrations of 0.5, 1.0, and 1.5 mg/mL were prepared in batches of 1 mL, which yields 5 samples for confocal imaging. The solution was composed of the following: i) 5% 10X MEM (Invitrogen); ii) 1:4 ratio of labeled to unlabeled collagen; iii) a few  $\mu\text{L}$  of 1 M NaOH to neutralize the pH; iv) an appropriate amount of DMEM (Invitrogen) to bring the total volume to 1 mL. Because collagen monomers interact quickly once the pH is increased, collagen was always added last to the solution. To ensure reproducible results, we always added and mixed the components in the following order: 10X MEM, DMEM, NaOH and finally collagen.

### Collagen imaging

We used a Leica SP5 (Leica Microsystems Inc., Bannockburn, IL) resonant confocal microscope with a 63X 1.2-NA water immersion objective to image the collagen samples. Excitation was provided by a 543 nm HeNe laser, while emitted photons with wavelengths between 550 and 650 nm were collected via photomultiplier tubes. The laser intensity was set at 75% of its maximum value, at which minimal photobleaching was observed; line- and frame- averaging increased the signal-to-noise ratio (SNR) significantly. Each image was  $1024 \times 512$  pixels in-plane, with each pixel  $100\text{ nm} \times 100\text{ nm}$ ; in the axial direction, usually 256 slices were imaged, with a typical spacing of 100 nm, though 200 nm was sometimes used, especially for sparser (0.5 mg/mL) collagen networks. The pinhole aperture was set to 1 Airy unit, the default value of the imaging software; this corresponds to a depth-of-field of approximately half a micron. To convert images from the Leica .LIF format to Matlab-readable TIFF stacks, an extended version of ImageJ was used, which is provided by the McMaster Biophotonics Laboratory (<http://www.macbiophotonics.ca/imagej/>).

### Bulk rheology

An AR-G2 (TA Instruments, New Castle, DE) rheometer with a  $4^\circ$ , 40 mm cone-plate geometry with a  $109\ \mu\text{m}$  gap was used to measure the shear modulus of gels. Approximately 1.2 mL of collagen solution was pipetted onto the  $37^\circ\text{C}$  preheated bottom plate of the rheometer and the cone was

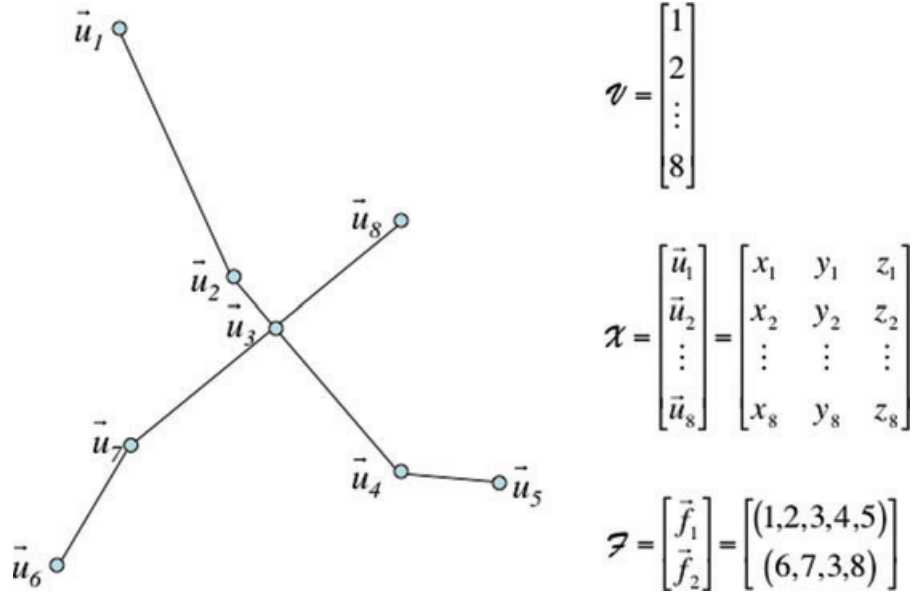


Fig. 2. A sample network illustrating the definitions in Section 3.

lowered onto the sample. A solvent trap was used to prevent the sample from drying during the measurement. During polymerization, the increase in the shear and viscous moduli was probed by continuously oscillating the sample at a fixed 0.5% strain amplitude and at a frequency of 0.2Hz. We verified that a 0.5% shear strain was within the linear regime by performing oscillatory strain sweeps (data not shown).

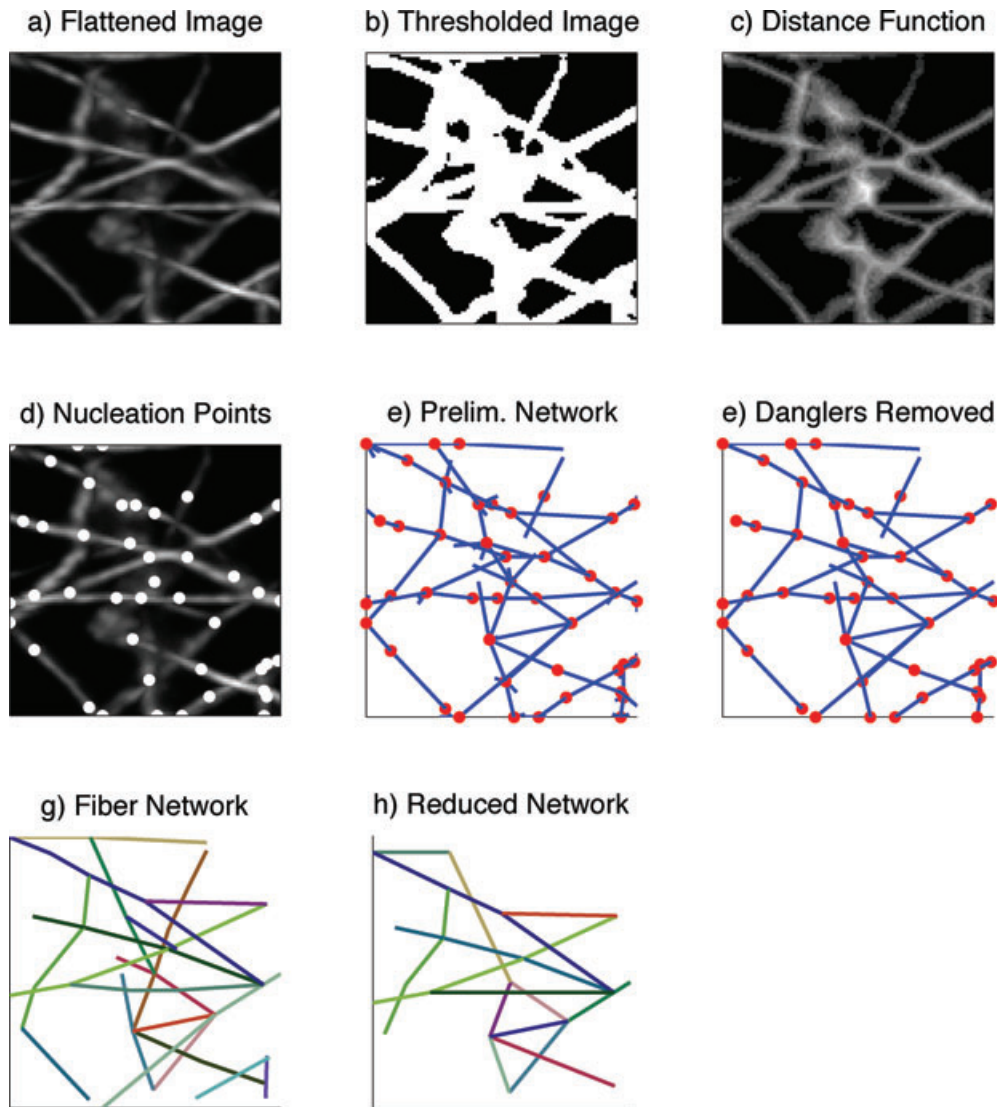
### FibeR Extraction (FIRE)

Before describing the network extraction algorithms, we first define a network. A network  $\mathcal{N} = \{\mathcal{V}, \mathcal{X}, \mathcal{F}\}$  consists of a list of  $n$  vertices  $\mathcal{V} = \{1, 2, \dots, n\}$ , with coordinates  $\mathcal{X} = \{\bar{u}_1, \bar{u}_2, \dots, \bar{u}_n\}$  and  $m$  fibers  $\mathcal{F} = \{\bar{f}^1, \bar{f}^2, \dots, \bar{f}^m\}$ . Each vertex  $v_i$  has an associated 3d vector  $\bar{u}_i = (x_i, y_i, z_i)$  indicating the vertex location in space, referenced to a global coordinate origin. Each fiber  $\bar{f}^k$  is a vector of  $p^k$  vertex identifiers  $\bar{f}^k = (v_1^k, v_2^k, \dots, v_{p^k}^k)$  which indicate the vertices that fiber  $k$  passes through. The superscript of  $v_i^k$  identifies the fiber and the subscript identifies the ordered position of the vertex in that fiber. The order is important, although a reversal of the order gives the identical fiber. A simple network is shown in Figure 2.

An 8 bit image  $I(\bar{u})$  for a volume of size  $L \times M \times N$  is defined to be a mapping  $I(\bar{u}) : U \rightarrow \mathbf{Z}_{[0,255]}$ , where  $U = \{\mathbf{Z}_{[0,L]} \times \mathbf{Z}_{[0,M]} \times \mathbf{Z}_{[0,N]}\}$ . Here  $\mathbf{Z}_{[0,L]}$  denote the set of integers  $\{0, 1, \dots, L\}$ . A flattened image is defined as  $J(x, y) = \max_z \{I(x, y, z)\}$ , where throughout this manuscript,  $\hat{z}$  is taken to be the direction perpendicular to the focal plane. A flattened image, shown in Figure 3(a) is used to illustrate the 3d images that are analyzed.

A new algorithm is now described that is based on the work of Wu *et al.*, (2003) but with significant modifications that considerably improve the results and reduce the computation time by over an order of magnitude. The algorithm requires a substantial number of parameters to be adjusted by the user, and these parameters are summarized in Appendix A, along with a guide for choosing these parameters.

The 3d image is first smoothed with a gaussian filter and then binarized, such that pixels of value one are likely to be fibers and pixels of value zero are likely to be the background. Two typical methods for binarizing an image are to keep a percentage of the brightest pixels or to keep all the pixels that are brighter than a particular value. If a percentage is used, then the number of “on” pixels is conserved, no matter how dense the network is. Thus the more fibers present, the thinner each fiber will appear so that total number of bright pixels in the binarized image is conserved. To reflect that a denser network should have more “on” pixels, the binarization method of selecting a pixel value threshold is used. It should be noted that for the algorithm to be able to obtain a network that qualitatively matches the image, the threshold must be chosen appropriately. If it is too high, one long fiber may be broken into many shorter, disjoint fibers. If the threshold is too low, then fibers that are clearly separated in the original image become blurred together. We have found that with confocal images of fluorescently labeled collagen, it is relatively straight forward to choose a threshold that qualitatively appears to maintain network topology, c. f. Figure 3(b)). Furthermore, we found that the mechanical properties of the extracted network were relatively insensitive to the choice of threshold, as discussed in Appendix A. However, for noisier imaging modalities such as confocal reflectance microscopy used by Wu *et al.*,



**Fig. 3.** Typical results after each step of FIRE. The images all come from a cube that is 5  $\mu\text{m}$  on a side. a) Flattened 3d image along  $\hat{z}$  b) Flattened thresholded image c) The flattened distance function ( $D = 0$  is black, and  $D = 5.1$  is white) d) Nucleation points, identified by dots, overlaying flattened image e) Preliminary network after each nucleation point is extended f) Preliminary network after artificial short fibers are removed. Note that many nucleation points near the center of the image have fewer short fibers when compared to (e). g) Fiber network after edges of like orientation are linked together. Different colors indicate different fibers. h) Reduced network for FEA modeling. Fibers are interpolated and portions of the network that don't contribute to network stiffness are removed.

(2003), it may be more difficult to choose an appropriate threshold.

Next, the minimal distance from a fiber pixel to a background pixel is calculated using Matlab's `bwdist.m` function in the image processing toolbox, see Figure 4. Then the distance function is smoothed with a Gaussian filter of standard deviation  $\sigma_d < 1$  pixel and is represented by  $D(\vec{u})$ , with  $\vec{u} \in U$ . The specific value for  $\sigma_d$  does not significantly impact the results. Smoothing the distance function improved the overall fiber tracking results, but was not done in Wu *et al.*, (2003). It has been found in systems of blood vessels that the network can also be extracted by tracing the ridge of the image intensity

itself (Lorigo *et al.*, 2001), but here, significantly better results are attained using the distance function.

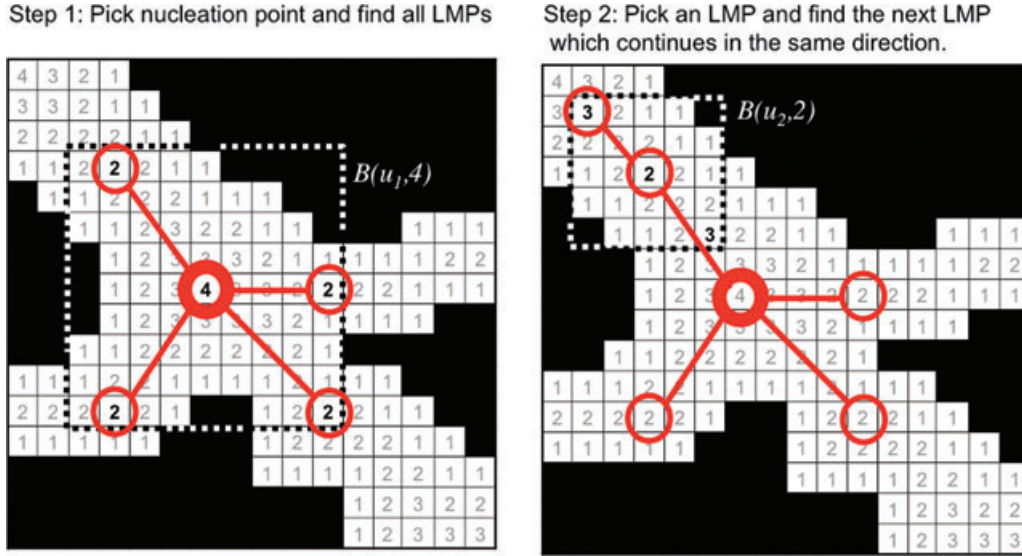
The goal now is to trace along the maximal ridges of the distance function. The process is illustrated in Figure 4. For clarity, some definitions are provided. A box  $B(\vec{u}, r)$  and its boundary  $\partial B(\vec{u}, r)$  with radius  $r$  around a point  $\vec{u} \in U$  are defined as follows.

$$B(\vec{u}, r) = \{\vec{v} \in U \mid \|\vec{u} - \vec{v}\|_\infty \leq \lceil r \rceil\}$$

$$\partial B(\vec{u}, r) = \{\vec{v} \in U \mid \|\vec{u} - \vec{v}\|_\infty = \lceil r \rceil\}.$$

Here,  $\lceil r \rceil$  denotes the smallest integer that is greater than  $r$ . The set of nucleation points,  $\mathcal{P}$  in image space  $U$  is defined to





**Fig. 4.** A visual outline for the tracing of the ridges of the distance function in 2d. Generalization to 3d is straight forward. Here, a binarized image is represented, where the black squares indicate background and the white squares indicate fibers. The number inside each white square indicates an approximate value that the smoothed distance function may take. In practice, it can take on floating point values, but integers are used for clarity. In step 1, a nucleation point ( $\tilde{u}_1$ , indicated by the thick circle) is selected. A box of radius  $\lceil D(\tilde{u}_1) \rceil$  is drawn (dashed line) and 4 LMPs are identified (thin circles). The vertex list  $\mathcal{X}$  and fiber list  $\mathcal{F}$  are updated accordingly, with one new fiber for each LMP. In step two, the fiber in the upper left corner is extended by one vertex to the next LMP. Note that the 3s in the upper left and lower right corners of  $B(\tilde{u}_2, 2)$  are both LMPs. The 3 in the upper left corner is selected as the next vertex in the fiber because it lies in the most similar direction to the fiber orientation. This new vertex is then added to  $\mathcal{V}$  and  $\mathcal{X}$ , and  $\mathcal{F}$  is again updated accordingly. This branch will be extended until either no more LMPs can be found (indicating the branch has reached the end of the fiber) or another nucleation point is contained within  $B(\tilde{u}_n, D(\tilde{u}_n))$ .

be:

$$\mathcal{P} = \{\tilde{u} \in U \mid D(\tilde{u}) \geq \max\{\theta_{nuc}, D(\tilde{v})\} \quad \forall \tilde{v} \in B(\tilde{u}, s_{xbox})\}$$

where  $\theta_{nuc} = 2$  and  $s_{xbox} = 5$ , are typical parameters that are set by the user. In words, nucleation points occur at local maxima of  $D(\tilde{u})$ , where the smoothed distance function also exceeds a threshold  $\theta_{nuc}$ . The nucleation points for a sample network are shown in Figure 3(d). A single nucleation point is denoted by the thick circle in Figure 4. After identifying the nucleation points, the next step is to trace the branches extending from the nucleation points. For this, we define the set of Local Maximum Points (LMPs) at nucleation point  $i$  as follows.

$$\begin{aligned} \{\text{LMP}\}_i &= \{\tilde{v} \in \partial B(\tilde{u}_i, r_i) \mid \\ &D(\tilde{v}) \geq \max\{\theta_{LMP}, D(\tilde{w})\} \quad \forall \tilde{w} \in \partial B(\tilde{u}_i, r_i) \\ &\text{s. t. } \|\tilde{v} - \tilde{w}\|_\infty = 1\}, \end{aligned}$$

where  $r_i = \lceil D(\tilde{u}_i) \rceil$ . In words LMPs occur where there is a local maximum in  $D(\tilde{u})$  on the surface of the box  $\partial B(\tilde{u}_i, r_i)$  that exceeds a critical threshold  $\theta_{LMP}$  (typically 0.2). In selecting  $\theta_{LMP}$ , if it is too small, many short, artificial side branches are created at each nucleation point. However, if it is too large, some fibers are missed. We found it was best to make  $\theta_{LMP}$  small and subsequently remove the artificial side branches in the next processing step. In the unlikely incident that two

LMPs occur less than  $\theta_{LMPdist}$  (typically equal to 2) apart, one of them is removed to save time. This process is illustrated in “Step 1” of Figure 4. Each new branch is then extended until it reaches an end and a new fiber is added to the fiber list containing the start and end point of the branch. If the end point is not a nucleation point, it is added to the vertex list. Once all branches have been extended from one nucleation point, the algorithm goes to the next one until all nucleation points have been explored.

A single step of the fiber extension process is shown in “Step 2” of Figure 4. The tip of the current fiber is denoted by  $\tilde{l}$ . During the first extension step, the fiber direction  $\hat{d}_1$  is defined to be  $\hat{d}_1 = (\tilde{l}_1 - \tilde{u}_i) / \|\tilde{l}_1 - \tilde{u}_i\|_2$ , where the subscript on  $\tilde{l}$  denotes the vertex index along the new fiber and the subscript on  $\tilde{u}$  denotes the nucleation point index. At  $\tilde{l}_n$ , a box of radius  $\lceil D(\tilde{l}_n) \rceil$  is drawn. If a nucleation point lies inside the box, that nucleation point is added to the end of the fiber and the fiber has reached its end. If no nucleation point lies in the box, then a new set of LMPs  $\{\tilde{l}_j\}$  are identified. If none are found, then the fiber has reached its end. If at least one LMP is present, then the direction of each LMP is calculated by  $\hat{\delta}_j = (\tilde{l}_n - \tilde{l}_j) / \|\tilde{l}_n - \tilde{l}_j\|_2$ . Only LMPs with the property that  $\hat{d}_n \cdot \hat{\delta}_j < \cos(\theta_{ext})$ , that could extend the fiber in a similar direction are considered and if none exist, the fiber has reached its end. A typical value for  $\theta_{ext} = 60^\circ$ . If exactly one LMP meets this criterion, then

it is added to  $\mathcal{V}$  and  $\mathcal{X}$  and the current fiber is extended. If more than one LMP meets this criterion, then the one with the largest  $D(\vec{l}_j)$  is chosen. Finally, the fiber tip direction is updated in the following way:

$$\hat{d}_{n+1} = \frac{\lambda_{dirdecay} \hat{d}_n + (1 - \lambda_{dirdecay}) \delta_j}{\|\lambda_{dirdecay} \hat{d}_n + (1 - \lambda_{dirdecay}) \delta_j\|_2}$$

Here,  $\lambda_{dirdecay}$  is a number between zero and one, with a typical value of 0.5. Its purpose is to give some memory to the fiber direction update.

As mentioned earlier, this algorithm is similar to the one developed by Wu *et al.*, (2003), but it has been found that the differences have improved the computation time by an order of magnitude while also greatly improving the accuracy of the network. The main difference is the use of multiple nucleation points and the restriction that each branch extend only in a single direction from a nucleation point. In Wu *et al.*, (2003), the network has a single nucleation point at the global maximum of  $D(\vec{u})$ . A similar process of tracing along LMP trajectories is done, but at each new LMP, multiple branches may form. This results in the creation of many artificial side branches. This approach also requires frequent checks to ensure that the algorithm does not get stuck in a loop that continually traces the same path. Loops in the network, which are quite common, caused serious issues in our implementation of the algorithm developed by Wu *et al.*, (2003). The introduction of nucleation points also greatly improves the algorithm's ability to identify crosslinks.

As mentioned above, a low LMP threshold means that there are often short fibers extending from each nucleation point, as shown in Figure 3(e). Dangling fibers are identified as those having length less than  $\theta_{dang-L}$  and having a tangent angle that differs from any other incoming fiber by at least  $\theta_{dang-extend}$ . The dangling fibers are removed in Figure 3(f). When two fibers touch at their end points, or have end points that are close together, and share similar orientation at the ends, then these two fibers are linked together to form a single, longer fiber, as seen when comparing Figure 3(f) to Figure 3(g). This step begins by first calculating the directions of the two fiber ends as follows:

$$\begin{aligned} v_1 &= \vec{f}_1^k, & v_2 &= \vec{f}_{s_{fiberdir}}^k, & v_3 &= \vec{f}_{p^k - s_{fiberdir}}^k, & v_4 &= \vec{f}_{p^k}^k \\ d_1^k &= (\mathcal{X}(v_2) - \mathcal{X}(v_1)) / \|\mathcal{X}(v_2) - \mathcal{X}(v_1)\|_2 \\ d_2^k &= (\mathcal{X}(v_3) - \mathcal{X}(v_4)) / \|\mathcal{X}(v_3) - \mathcal{X}(v_4)\|_2 \end{aligned}$$

where  $p^k$  is the number of vertices in the fiber and  $s_{fiberdir}$  (typically equal to 4) denotes the spacing along the fiber used to calculate the vertex direction. Next, the algorithm searches for fiber tips that are less than  $\theta_{linkd}$  (typically equal to 8°) apart. If the angle between the two tip directions is between  $\theta_{linka}$  (typically equal to 130°) and 180°, then those fibers are linked to form a single fiber. If there is more than one fiber tip that satisfies these properties, then the two with

an angle between them closest to 180° are linked. The end output of the algorithm is the network  $\tilde{\mathcal{N}} = \{\tilde{\mathcal{V}}, \tilde{\mathcal{X}}, \tilde{\mathcal{F}}\}$ , where the tilde indicates approximation to the true network. In the following two sections, we test this algorithm by comparing the estimated geometric and mechanical properties of the collagen gels extracted by FIRE to those of artificially generated networks.

## Artificial Network Generation

We test the ability of our algorithm to extract a network geometry using artificially generated networks with known structure. To get the parameters for generating such networks, we first extract the networks of gels at three different collagen densities (0.5, 1.0, and 1.5 mg/ml). Results for a single network at 0.5 mg/ml are shown in Figure 5 and the results for all three densities are summarized in Figure 7. The fiber length ( $L_f$ ) is exponentially distributed with  $\langle L_f \rangle \approx 6 \mu\text{m}$ . The spacing between *apparent* cross-links ( $L_c$ ) is Poisson distributed as expected (Head *et al.*, 2003b), with  $\langle L_c \rangle \approx 2.5 \mu\text{m}$ . Given a point spread function and its full width at half maximum (FWHM), an apparent cross-link is defined to exist wherever two fibers are less than one FWHM apart. This is because it is impossible to distinguish unconnected fibers that are close together from cross-linked fibers. In an attempt to address this issue, we deconvolved the image stacks using the Classic Maximum Likelihood Estimation restoration method provided with Huygens Deconvolution Software (Hilversum, The Netherlands). Deconvolution significantly reduced the noise and also partially reduced the apparent fiber width in the  $\hat{x}\hat{y}$  plane. However, the fibers still appeared considerably thicker in the  $\hat{z}$  direction compared to the imaging plane and distributions for  $L_f$ ,  $L_c$  and  $L_p$  were not significantly altered (data not shown). The fiber persistence length is log-normally distributed with a median length  $\langle L_p \rangle' \approx 20 \mu\text{m}$ , where median is denoted by a prime. The variance of  $\log(L_p)$  is  $\sigma_p = 0.85$ . The mean of the persistence length is not well defined because it is possible for a fiber to be perfectly straight and have infinite persistence length. The details for how the persistence length is calculated are given in Appendix B. It should be noted that the persistence length as defined here is not a thermal persistence length, which is on the order of centimeters, but rather a geometrical persistence length due to innate curvature of the fibers. To approximate the average radius  $\langle r \rangle$  of a fiber, we assume the fibers are curvilinear with circular cross-section and thus the total volume of the collagen fibers  $V_c$  is given by  $V_c = \pi \langle r \rangle^2 L_{tot}$ , where  $L_{tot}$  is the total fiber length in the network image. Assuming that no collagen is left in solution and that the volume of each fiber is entirely composed of collagen gives  $V_c = V_g \rho_c v_c$ , where  $V_g$  is the volume of the gel,  $\rho_c$  is the mass density of collagen in solution (0.5, 1.0, 1.5 mg/ml) and  $v_c = 0.73 \text{ ml/g}$  is the specific volume of collagen (Hulmes & Miller

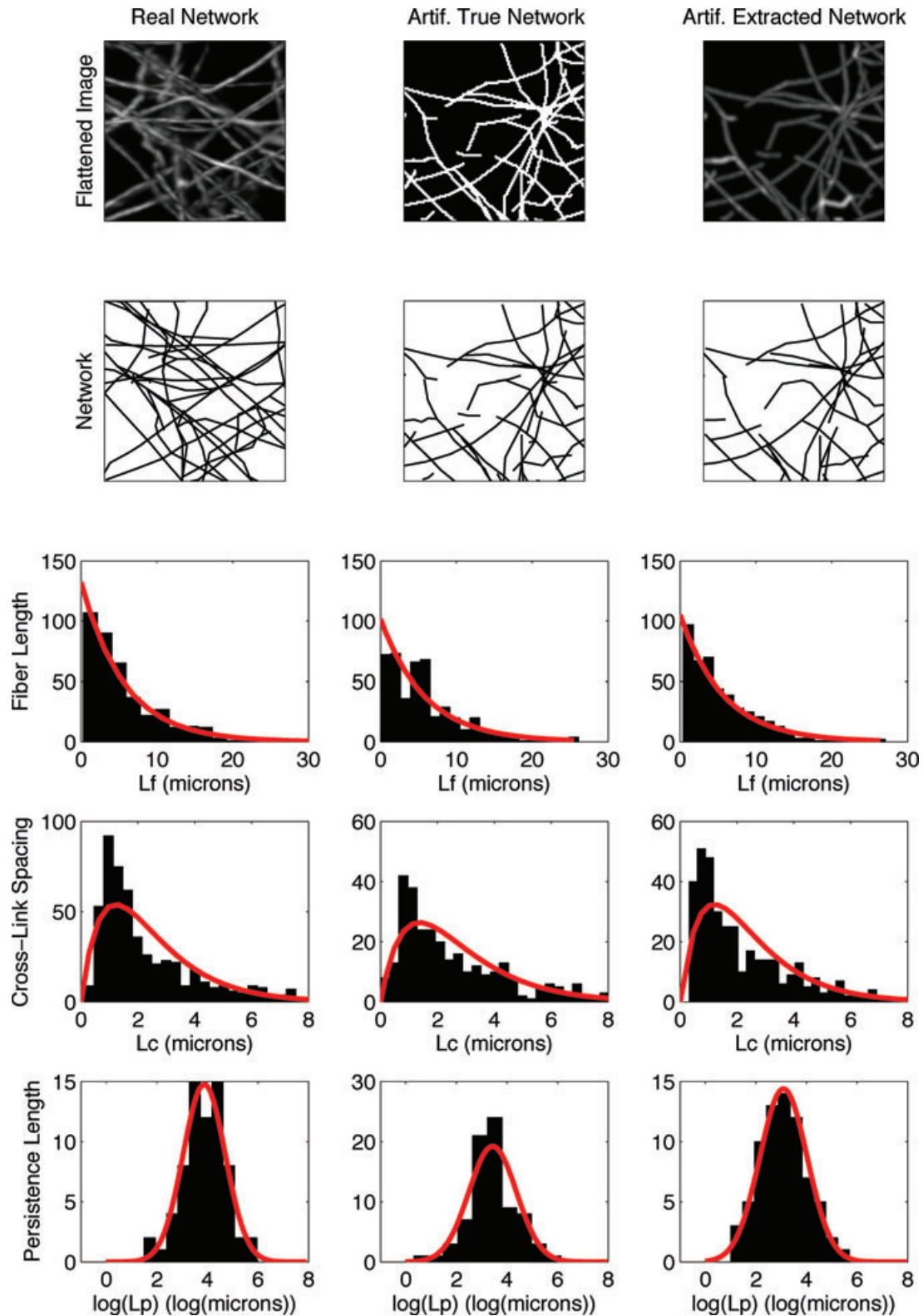


Fig. 5. Geometric properties of a single network with collagen density 0.5 mg/ml. We look at a real network extracted by FIRE, an artificial network of known geometry, and an artificial network that is extracted from the corresponding artificial image. The histograms show the length distributions for the actual networks and the solid lines show fits to the data, where fiber length is fit to an exponential distribution, cross-link spacing is fit to a Poisson distribution, and persistence length is fit to a log-normal distribution. The specific parameters for the distributions are given in the text and in Figure 7.

1979). These two equations are combined below.

$$\langle r \rangle = \sqrt{\frac{V_g \rho_c v_c}{\pi L_{tot}}}$$

From this equation, we find that the estimate of  $\langle r \rangle \approx 30$  nm. This matches what has been observed using scanning electron microscopy (Raub *et al.*, 2007).

The artificial networks are generated in a cube 25.5  $\mu\text{m}$  on each side to match experiment. The networks are grown one fiber at a time until the network achieves the target collagen density. When fiber  $k$  is added, it is assigned an initial point of nucleation  $\vec{u}_0^k$  inside the volume, and an initial tangent direction  $\vec{t}_0^k$ . Both are chosen from a uniform distribution. The fiber is also assigned a target fiber length  $L_f^k \sim \text{Exp}(1/\langle L_f \rangle)$  and persistence length  $\log L_p^k \sim N(\log \langle L_p \rangle', \sigma_p)$ . The fiber is assigned a plus and minus end and with equal probability, the initial nucleation point may be extended by one micron to  $\vec{u}_1^k = \vec{u}_0^k + \vec{t}_0^k$  or  $\vec{u}_{-1}^k = \vec{u}_0^k - \vec{t}_0^k$ . On each ensuing step, an end is chosen at random and the fiber extends again, following a persistent random walk. In such a scheme, the tangent angle changes as the fiber grows and the rule for updating the tangent direction comes from the definition of the persistence length  $L_p$  for a fiber with length parameterized by  $l$  (Howard 2001).

$$\langle \cos(\theta(l + \Delta l) - \theta(l)) \rangle_l = \exp[-\Delta l / L_p] \quad (1)$$

Here,  $\theta(l)$  is the angle between the tangent vector main fiber axis which links the two end points. A typical fiber is shown in Figure 6. Letting  $\Delta\theta = \theta(l + \Delta l) - \theta(l)$  and Taylor expanding about  $\Delta l$  gives

$$\left\langle 1 - \frac{1}{2} \Delta\theta^2 \right\rangle_l \approx 1 - \frac{1}{L_p} \Delta l$$

$$\langle \Delta\theta^2 \rangle_l \approx \frac{2}{L_p} \Delta l$$

This is a discretized diffusive process in  $\theta$ . Hence, the tangent angle of a particular fiber end after  $n$  extensions in a particular direction is given by  $\theta_{n+1}^k = \theta_n^k + \Delta\theta$ , where  $\Delta\theta$  is a random variable with mean 0 and variance  $2\Delta l / L_p$ . Any distribution satisfying these properties could be used and here,  $\Delta\theta$  is chosen from a normal distribution. Specifying  $\Delta\theta$  is not enough to

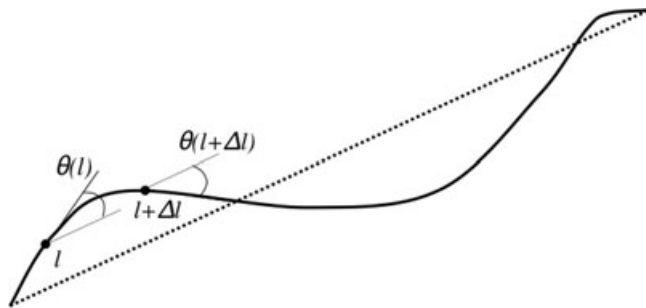


Fig. 6. A typical fiber with the arc length parameterized  $l$ .

give a unique direction because there is a cone of possible new direction vectors that differ in angle by  $\Delta\theta$  from the previous direction. Hence once  $\Delta\theta$  is chosen, a specific direction vector in the cone is chosen from a uniform distribution.

The fiber continues to extend in 1  $\mu\text{m}$  increments until its length exceeds  $L_f^k$  or it reaches the edge of the 25.5  $\mu\text{m}$  box. If the fiber exceeds its maximum length, the appropriate length is subtracted off the most recent extension and growth halts. If the fiber intersects one of the faces of the volume, the most recent addition is cut such that the fiber ends at that face and extension is no longer permitted in that direction. If both fiber ends intersect the box before the fiber reaches its target length, then growth for that fiber halts. Once a fiber is finished growing, the total density of the network is checked and if it is below the target density, then a new fiber is added.

Once the network has reached its target size, the final step is to cross-link fibers at appropriate locations. As noted above, the FIRE algorithm cannot distinguish cross-linked fibers from fibers that appear to be cross-linked because of blurring. Our goal here is to test whether FIRE is able to correctly identify apparent cross-links. Calculating the minimum distance between two segments using an algorithm described by Eberly (1999), we create cross-links between segments that are closer than  $d_{||} = 150$  nm in the focal plane or  $d_{\perp} = 600$  nm in the perpendicular direction, where  $d_{||}$  and  $d_{\perp}$  are the FWHM of the fiber intensity function in the focal plane and perpendicular to the focal plane, respectively. When a cross-link is created between two fibers  $\vec{f}^k$  and  $\vec{f}^l$ , a new vertex is added halfway between the two fibers and  $\vec{f}^k$  and  $\vec{f}^l$  are updated accordingly. An unfortunate effect is that this cross-linking algorithm is that it reduces the fiber persistence length *c. f.* Figure 7(b). This could be addressed with a more realistic algorithm for network polymerization where the fibers diffuse freely through space as they grow, but development of such an algorithm is beyond the scope of this manuscript.

From this artificial network, an artificial image is formed by first creating a binary image with 0.1  $\mu\text{m}$  resolution that is one at every pixel that is less than  $\langle r \rangle$  from a fiber and zero elsewhere. The image is then convolved with a 3d gaussian with standard deviation  $d_{||}$  in the  $\hat{x}$  and  $\hat{y}$  directions and  $d_{\perp}$  in the  $\hat{z}$  direction. Finally, uncorrelated gaussian white noise is added to the image with a standard deviation of  $\sigma_{noise} = 5$ . Because these are 8 bit images, the noise is not permitted to drive the signal below 0 or above 255. FIRE is then used to extract networks from these images. An example image is shown next to an actual network in the top row of Figure 5.

### Algorithm Performance

In Figures 5 and 7, the geometric properties of real, extracted (RE) collagen networks are compared to those of both the true, artificial networks (AT) and the artificial networks extracted by FIRE (AE). In Figure 5, we see that for an example network



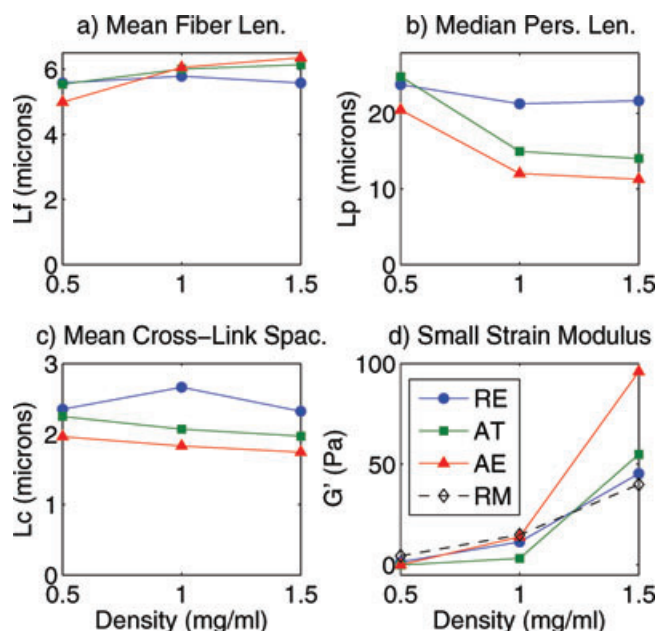


Fig. 7. The geometric (a-c) and mechanical (d) properties of the real extracted (RE) networks, the artificial true (AT) networks, the artificial extracted (AE) networks (extracted from the artificial images). In (d), we also have real measured (RM) mechanical stiffness of actual gels. Overall, the agreement between artificial and experimental networks is good, and so is the agreement between the stiffness of RM and the finite element model of RE.

at 0.5 mg/ml,  $L_f$ ,  $L_p$ , and  $L_c$  are similar for all three networks. Ideally, we would have also compared this algorithm to the one developed by Wu *et al.*, (2003), but the authors were unable to make their code available, and due to the issues of loops in the network, our attempt to implement the code was not successful. One significant discrepancy between the experimental and the artificial networks is that the artificial networks contain short fibers that are disconnected from the rest of the network. Fibers such as these are typically not seen in experimental networks and this discrepancy is likely due to the fact that in real networks, there is a large degree of branching and fewer fiber nucleation events. Branching was not included in the artificial network algorithm because at present, we do not distinguish between cross-overs and branches. Somewhat surprisingly, despite this discrepancy, RE, AT, and AE still had similar cross-link spacing and mechanical properties, as discussed in more detail below. The geometric properties  $L_f$ ,  $L_p$ , and  $L_c$  of AT and AE agree to within 10% in Figure 7(a-c).

In evaluating the mechanical properties of the network, we first perform two postprocessing steps: fiber interpolation and removal of portions of the network that do not contribute to the network stiffness. The postprocessing is described in Appendix C. Afterwards, we use a small strain finite element model, treating each fiber as an Euler-Bernoulli Beam (Bathe

1995), with a Young's Modulus of 500 MPa (van der Rijt *et al.*, 2006). The fibers are treated as cylindrical rods with  $r = 30$  nm. Such a model for a collagen fiber has been experimentally validated by Yang *et al.*, (2008). Modeling the cross-links between fibers is more challenging, since little is known about their mechanical properties. Thus we choose a simple model that assumes each apparent cross-link acts as a pin joint with a torsional spring attached. The torsional spring constant is chosen to be  $K_{xlink} = 1 \mu\text{N}\cdot\mu\text{m}$ , which gave a good match between the shear moduli of extracted networks and the shear moduli of actual collagen gels at three different densities, shown in Figure 7(d). While increasing  $K_{xlink}$  increases the network stiffness, the power-law observation that  $G \sim \rho^{2.2}$  does not change. Because FIRE is only able to detect apparent cross-links, one could choose to remove some fraction of the cross-links in the model and a larger  $K_{xlink}$  would be needed to match experiment. Since little is known about the dynamics of polymerization, it would be difficult to justify any particular fraction and so we did not explore this further.

The small strain shear modulus  $G_{bulk}$  is estimated by imposing a shear displacement  $u$  on one face of the network while keeping the opposite face fixed. The equilibrium configuration is solved for using the finite element method and the energy  $U$  of the deformed network is computed. The bulk modulus  $G_{bulk}$  of the network can then be solved for by noting that for small deformations  $U = G_{bulk} \gamma^2/2$ , where  $\gamma = u/L_{bulk}$  is the engineering strain. We see in Figure 7(d) that the stiffness of the extracted network is in relatively good agreement with what is measured experimentally by a cone-plate rheometer. The stiffness of AE is about 50% larger than for AT and this is likely due to the decrease in  $L_c$  and hence a larger number of cross-links present in the extracted network over the true network. We believe a 50% agreement to be relatively good since a particular protocol often yields variations of this magnitude across different rheology experiments.

## Conclusions

A new algorithm has been presented for extracting the biopolymer network geometry from a 3d image. This algorithm has been tested on both artificial and real networks. It has been shown that the algorithm is able to reliably reproduce the geometries of artificial networks of known architecture. By tuning a single unknown parameter  $K_{xlink} = 1 \mu\text{N}\cdot\mu\text{m}$ , we find that the small strain modulus of the network at three different densities, measured by cone-plate rheometry, is in agreement with both the extracted collagen networks and artificial networks with similar geometric properties.

A fundamental weakness of this algorithm is that the FWHM of a fiber is much larger than the fiber diameter and thus it is impossible to tell whether an apparent cross-link is truly a cross-link or simply two fibers in close proximity. Ultimately,

this issue could be addressed by observing the network response to localized deformations or observing the network behavior during polymerization. Either approach ultimately requires an algorithm for extracting network geometry and thus while FIRE does not solve this issue, it does provide a tool that will ultimately be needed for further investigation of the fibril cross-link properties.

## Acknowledgements

We would like to thank T. Jackson, T. Deisboeck, E. A. Chiocca, M. Berens, P. Smereka, P. Chandran, A. Gilbert, and S. Esedoglu for useful discussions, and also V. Barocas and T. Stylianopoulos for providing their artificial network generation code. This work is supported by NIH Bioengineering Research Partnership grant R01 CA085139-01A2, the Institute for Mathematics and its Applications, the Swiss National Science Foundation, and NSF-Integrative Graduate Education and Research Traineeship Program.

## References

- Aguet, F., Jacob, M. & Unser, M. (2005) Three-dimensional feature detection using optimal steerable filters. In *IEEE International Conference on Image Processing 2005* vol. 2, II-1158–61.
- Baici, A., Cohen, G., Fehr, K. & Boni, A. (1980) A handy assay for collagenase using reconstituted fluorescein-labeled collagen fibrils. *Anal. Biochem.* **108**, 230–2.
- Baradet, T.C., Haselgrove, J.C. & Weisel, J.W. (1995) Three-dimensional reconstruction of fibrin clot networks from stereoscopic intermediate voltage electron microscope images and analysis of branching. *Biophys. J.* **68**, 1551–1560.
- Bathe, K.-J. (1995) *The Finite Element Method* 399–420. Prentice-Hall, Upper Saddle River.
- Boskamp, T., Hahn, H., Hindennach, M., Zidowitz, S., Oeltze, S., Preim, B. & Peitgen, H.-O. (2005) *Geometrical and structural analysis of vessel systems in 3D medical image datasets* vol. 5, chap. 1, 9–18. World Scientific, Singapore.
- Can, A., Turner, J.N., Tanenbaum, H.L. & Roysam, B. (1999) Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms. *IEEE Trans. Info. Tech. Biomed.* **3**, 125–138.
- Chandran, P.L. & Barocas, V.H. (2006) Affine versus non-affine fibril kinematics in collagen networks: Theoretical studies of network behavior. *J. Biomech. Eng.* **128**, 259–269.
- Donoho, D. & Huo, X. (2002) Beamlets and multiscale image analysis. In *Multiscale and Multiresolution Methods* edited by T.J. Barth, T. Chan & H.R., Springer Lecture Notes in Computational Science and Engineering. Springer, New York.
- Eberly, D. (1999) Distance between two line segments in 3D. <http://www.geometrictools.com/Documentation/DistanceLine3Line3.pdf>.
- Head, D., Levine, A.J. & MacKintosh, F. (2003a) Distinct regimes of elastic response and deformation modes of cross-linked cytoskeletal and semiflexible polymer networks. *Phys. Rev. E* **68**, 061907.
- Head, D., Levine, A.J. & MacKintosh, F. (2005) Mechanical response of semiflexible networks to localized perturbations. *Phys. Rev. E* **72**, 061914.
- Head, D., MacKintosh, F. & Levine, A.J. (2003b) Nonuniversality of elastic exponents in random bond-bending networks. *Phys. Rev. E* **68**, 025101(R).
- Heussinger, C., Schaefer, B. & Frey, E. (2007) Nonaffine rubber elasticity for stiff polymer networks. *Phys. Rev. E* **76**, 031906.
- Howard, J. (2001) *Mechanics of motor proteins and the cytoskeleton* 110–111. Sinauer Associates, Inc, Sunderland.
- Huisman, E.M., van Dillen, T., Onck, P. & van der Giessen, E. (2007) Three-dimensional cross-linked f-actin networks: Relation between network architecture and mechanical behavior. *Phys. Rev. Lett.* **99**, 208103.
- Hulmes, D.J.S. & Miller, A. (1979) Quasi-hexagonal molecular packing in collagen fibrils. *Nature* **282**, 878–880.
- Lorigo, L.M., Faugeras, O.D., Grimson, W. E.L., Keriven, R., Kikinis, R., Nabavi, A. & Westin, C.F. (2001) Curves: Curve evolution for vessel segmentation. *Medical Image Analysis* **5**, 195–206.
- Nisslert, R., Kvarnstrom, M., Loren, N., Nyden, M. & Rudemo, M. (2007) Identification of the three-dimensional gel microstructure from transmission electron micrographs. *J. Microsc.* **225**, 10–21.
- Raub, C.B., Suresh, V., Krasieva, T., Lyubovitsky, J., Mih, J.D., Putnam, A.J., Tromberg, B.J. & George, S.C. (2007) Noninvasive assessment of collagen gel microstructure and mechanics using multiphoton microscopy. *Biophys. J.* **92**, 2212–2222.
- Stoica, R., Descombes, X. & Zerubia, J. (2004) A Gibbs point process for road extraction from remotely sensed images. *Int. J. Comp. Vision* **57**, 121–136.
- Stylianopoulos, T. & Barocas, V.H. (2007) Volume-averaging theory for the study of the mechanics of collagen networks. *Comput. Methods Appl. Mech. Eng.* **196**, 2981–2990.
- Suri, J.S. (2002) A review of MR vascular image processing: skeleton versus nonskeleton approaches: part II. *IEEE Trans. Info. Tech. Biomed.* **6**, 338–349.
- Tupin, F., Maitre, H., Mangin, J.-F., Nicolas, J.-M. & Pechersky, E. (1998) Detection of linear features in SAR images: application to road network extraction. *IEEE Trans. Geosci. Remote Sensing* **36**, 434–453.
- van der Rijt, J. A.J., van der Werf, K.O., Bennink, M.L., Dijkstra, P.J. & Feijen, J. (2006) Micromechanical testing of individual collagen fibrils. *Macromol. Biosci.* **6**, 697–702.
- Van Dillen, T., Onck, P. & van der Giessen, E. (2006) Stiffening of semiflexible biopolymers and cross-linked networks. arXiv 0611230.
- Wilhelm, J. & Frey, E. (2003) Elasticity of stiff polymer networks. *Phys. Rev. E* **91**, 108103.
- Wu, J., Rajwa, B., Filmer, D.L., Hoffmann, C.M., Yuan, B., Chiang, C.-S., Sturgis, J.E. & Robinson, J.P. (2003) Automated quantification and reconstruction of collagen matrix from 3D confocal datasets. *J. Microsc.* **210**, 158–165.
- Yang, L., van der Werf, K.O., Koopman, B.F., Subramaniam, V., Bennink, M.L., Dijkstra, P.J. & Feijen, J. (2008) Micromechanical bending of single collagen fibrils using atomic force microscopy. *J. Biomed. Mat. Res. A* **94**, 2204–2211.

## Appendix A: Image Processing Parameters

Image processing parameters at 0.1 micron resolution in the  $\hat{x}$  and  $\hat{y}$  directions are given below. For each processing

stage, the most sensitive parameters are described and the methodology used for choosing these parameters is given. While there are many parameters to choose, we find that the parameters in each section can be chosen independently of the rest.

#### Image Preprocessing

$\sigma_{im}$	1	standard deviation of Gaussian smoothing filter for the image function (in pixels)
$\theta_{im}$	0.08	binary threshold for image, given as a percentage of the maximum pixel intensity, below which pixels are identified as background
$\sigma_d$	0.3	standard deviation of Gaussian smoothing filter for the distance function (in pixels)

The threshold  $\theta_{im}$  is chosen such that Figure 3(b) preserves as much as the network topology from Figure 3(a) as possible. If  $\theta_{im}$  is too high, fibers tend to get broken up into pieces whereas if it is too low, parts of the image get blurred together. To assess the algorithm sensitivity to this choice of threshold, we varied  $\theta_{im}$  and measured the overall change in stiffness of the extracted network. We found that by setting  $\theta_{im} = 0.01$ , more pixels were included in the network and the network stiffness roughly doubled. Doubling  $\theta_{im}$  to 0.16, and reducing the number of pixels in the network halved the network stiffness. Given that slightly different experimental protocols may also generate networks that differ in modulus by a factor of two, we consider the use FIRE for modeling the mechanical properties of a network to be relatively insensitive to this parameter for fluorescently labeled collagen.

#### Nucleation Point Identification

$\theta_{nuc}$	2	minimal value of $D(\vec{u})$ for which a point can be considered a nucleation point (in pixels)
$s_{xbox}$	5	local neighborhood of $\vec{u}$ , used to determine whether or not the point $\vec{u}$ is a local maxima (in pixels).

The most robust results are found to be when  $\theta_{nuc}$  is relatively low and there are many nucleation points along each fiber, as shown in Figure 3(d).

#### Fiber Extension

$\theta_{LMP}$	0.2	minimal distance function value for which a point may be considered an LMP
$\theta_{LMPdist}$	2	minimal distance (in pixels) between two LMPs
$\theta_{ext}$	60°	maximal value of the angle between the direction of the fiber and the direction of the candidate extension direction.
$\lambda_{dirdecay}$	0.5	decay rate for fiber direction update

If  $\theta_{LMP}$  is too small, many short, artificial side branches are created at each nucleation point. However, if it is too large,

some fibers are missed. We found it was best to make  $\theta_{LMP}$  small and subsequently remove the artificial side branches in the next processing step.

#### Removal of Dangling Fibers

$\theta_{dang-L}$	6	a fiber is removed if it is shorter than his value (in pixels), contains only a single cross-link, and is unaligned (defined below) with all other fibers that share the same cross-link.
$\theta_{dang-extend}$	10°	maximum difference in angle between a short fiber (defined above) such that it is considered to be aligned with another fiber going through the same cross-link

The purpose of this step is to remove the short fibers near the nucleation points in Figure 3(e). These parameters should be adjusted if the user finds that there are still many too many artificial, dangling fibers.

#### Fiber Processing

$s_{fiberdir}$	4	number of vertices to use in calculating the orientation of a fiber end
$\theta_{linkd}$	8	maximal distance between two fibers ends for which the fibers may be linked
$\theta_{linka}$	130°	minimum angle between two fiber ends for linking of the two fibers
$\theta_{flen}$	10	minimal fiber length for a fiber with one or zero crosslinks. shorter fibers are removed

The purpose of this step is to link fibers edges of like orientation. If the user finds that what looks to be one long fiber is broken into many smaller fibers,  $\theta_{linka}$  may be reduced and  $\theta_{linkd}$  may be increased.

#### Fiber Post-Processing for FEA model

$\lambda$	0.01	regularization constant for fitting interpolation function to fiber subunits
$s_{maxspace}$	2	maximum spacing between vertices along a fiber (in microns)
$s_{boundthick}$	15	thickness of boundaries (in pixels)

The purpose of these parameters is to reduce the computational complexity of the network for the finite element calculation. If the user finds that more nodes are necessary,  $s_{maxspace}$  can be reduced. If the user finds that the interpolated curve is overfit to the data, then  $\lambda$  should be reduced, though we have not found this to be an issue

#### Appendix B: Persistence Length Calculation

The persistence length is calculated for fiber  $\vec{f}^k = \{v_1^k, v_2^k, \dots, v_p^k\} = \{v_i^k\}$  by finding the best least squares fitting

exponential for the points  $c_{i,s}^k(l_{i,s}^k)$ , as defined below.

$$\begin{aligned} l_{i,s}^k &= \sum_{j=1}^s \|\vec{u}_{i+j}^k - \vec{u}_{i+j-1}^k\|_2 \\ \hat{t}_i^k &= (\vec{u}_{i+1}^k - \vec{u}_i^k) / \|\vec{u}_{i+1}^k - \vec{u}_i^k\|_2 \\ c_{i,s}^k &= \hat{t}_{i+s}^k \cdot \hat{t}_i^k \end{aligned}$$

In words,  $l_{i,s}^k$  is the contour length of fiber  $\vec{f}^k$  connecting vertex  $v_i^k$  to vertex  $v_{i+s}^k$ ,  $\hat{t}_i^k$  is the tangential direction of a fiber at vertex  $v_i^k$  and  $c_{i,s}^k$  is the cosine of the angle difference between vertex  $v_i^k$  and vertex  $v_{i+s}^k$ . The best exponential least squares fit, in combination with Equation 1, gives the persistence length for fiber  $\vec{f}^k$ .

### Appendix C: Post Processing for Mechanical Modeling

Once the network is extracted, two additional post processing steps are necessary before solving for the micromechanical response: interpolation of fibers and removal of the floppy portions of the network. To reduce the computational complexity in solving for the equilibrium configuration of the network, it is efficient to use fewer vertices for each fiber than those found by FIRE. To do this, a fiber is transformed in the following way. Each vertex of a fiber is classified as either a fixed vertex if it lies at a fiber end or on a cross-link, or a free vertex otherwise. The fiber is then broken into  $s$  fiber subunits, where  $s$  is the number of fixed vertices minus one. For example, if  $\vec{f}^k = (v_1^k, v_2^k, v_3^k, v_4^k, v_5^k, v_6^k)$  and the fixed vertices are  $(v_1^k, v_4^k, v_6^k)$ , then  $f^{k,1} = (v_1^k, v_2^k, v_3^k, v_4^k) \equiv (v_1^{k,1}, v_2^{k,1}, v_3^{k,1}, v_4^{k,1})$  and  $f^{k,2} = (v_5^k, v_6^k) \equiv (v_1^{k,2}, v_2^{k,2})$ , where the first superscript indicates the fiber number and the second superscript indicates the subfiber index. Each fiber subunit is then rotated into a local coordinate system, such that the end points lie along the  $x$  axis. The subfiber is then fit by an interpolation function. The interpolation function is sampled at interval  $s_{\text{maxspace}}$  and the interpolated positions are then rotated back to the global coordinate system to become the new vertices in the network. If there are only two vertices  $(v_i, v_j)$  in a subfiber, then a linear interpolation is used.

$$h_0(\xi) = ((1 - \xi)x_i + \xi x_j)\hat{x}$$

where  $\xi = (x - x_i) / |x_j - x_i| \in [0, 1]$ . However, if the subfiber contains more than two vertices, then that subfiber is represented by a total of five basis functions.

$$g(\xi) = h_0(\xi) + \sum_{l=0}^4 c_l h_l(\xi)$$

$$h_1(\xi) = (\xi^3 - 2\xi^2 + \xi)\hat{y}, \quad h_2(\xi) = (-\xi^3 + \xi^2)\hat{y}$$

$$h_3(\xi) = (\xi^3 - 2\xi^2 + \xi)\hat{z}, \quad h_4(\xi) = (-\xi^3 + \xi^2)\hat{z}$$

The four additional basis functions are some of the cubic Hermite polynomials that are used when constructing the linearized stiffness matrix for a network of beams. The

constants  $c_l$  are chosen to minimize the regularized least squared error between the interpolated function and the actual data points. For a particular fiber:

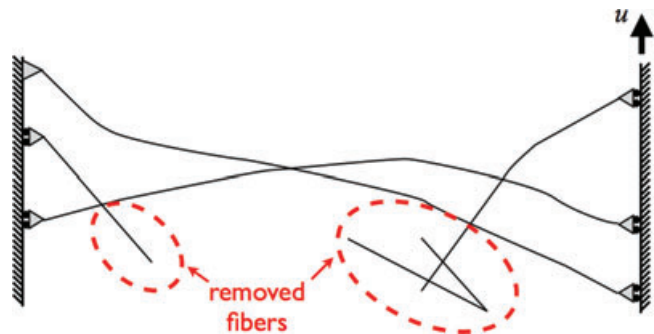
$$\begin{aligned} E^{k,i} &= \sum_{l=1}^{p_{k,i}} (y_l - c_1 h_1(\xi_l) - c_2 h_2(\xi_l))^2 \\ &+ \sum_{l=1}^{p_{k,i}} (z_l - c_3 h_3(\xi_l) - c_4 h_4(\xi_l))^2 + \lambda \sum_{l=1}^4 c_l^2 \end{aligned}$$

The regularization by  $\lambda$  (typically equal to 0.01) is necessary to prevent overfitting. We can directly solve for the constants  $c_l$  by differentiating the error with respect to the constants. In the case where  $l = 1$ :

$$\begin{aligned} \frac{dE^{k,i}}{dc_1} &= 2 \sum_{l=1}^{p_{k,i}} (y_l - c_1 h_1(\xi_l) - c_2 h_2(\xi_l)) h_1(\xi_l) + 2\lambda c_1 = 0 \\ &\left( \sum_{l=1}^{p_{k,i}} h_1(\xi_l)^2 + \lambda \right) c_1 + \left( \sum_{l=1}^{p_{k,i}} h_1(\xi_l) h_2(\xi_l) \right) c_2 \\ &= \sum_{l=1}^{p_{k,i}} y_l h_1(\xi_l) \end{aligned}$$

Differentiating with respect to other constants leads to 4 linear equations and 4 unknowns and these are solved to find the constants. The interpolation function is then rotated back to the global coordinate system and sampled at intervals of  $s_{\text{maxspace}}$  to generate an interpolated subfiber. The subfibers are then stitched back together to create an entire interpolated fiber. It should be noted that differentiability is not imposed at the point where two fibers are connected.

In general, when the extracted network above is used to generate a stiffness matrix, the matrix is singular and cannot be solved simply by Gaussian elimination. When a deformation is imposed upon opposing faces, an edge will only contribute to the network stiffness if it lies on a non-repeating path connecting the two faces. A non-repeating path is one where



**Fig. 8.** The dashed circles indicate the edges in the network that do not contribute to the stiffness of the network undergoing a displacement  $u$ . These edges can easily be identified and removed because they do not lie on a non-repeating path that connects the two opposing faces of the network.



each edge in the path is visited only once. An example network with non-contributing edges is shown in Figure 8. To reduce the computational complexity of the finite element problem for estimating the small strain stiffness of the networks, the non-contributing edges are removed from the network. It should be noted that there may exist nonrepeating paths

that lie outside of the confocal volume that contribute to the network stiffness. To address this issue, we have done calculations for the network stiffness of smaller networks (12.8  $\mu\text{m}$  cubes) in addition to the 25.6  $\mu\text{m}$  cubes that were analyzed in Figures 5 and 7. The two results were similar (data not shown).