Biopolymer network geometries: Characterization, regeneration, and elastic properties

Stefan B. Lindström,[1,*] David A. Vader,[2] Artem Kulachenko,[3] and David A. Weitz[2,†]

[1]*Department of Fiber and Polymer Technology, Royal Institute of Technology, SE-100 44 Stockholm, Sweden*
[2]*School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA*
[3]*Department of Solid Mechanics, Royal Institute of Technology, SE-100 44 Stockholm, Sweden*

We study the geometry of biopolymer networks and effects of the geometry on bulk mechanical properties. It is shown numerically that the physical network geometry can be quantified statistically and regenerated from its statistical description, so that the regenerated network exhibits the same network mechanics as the physical network in the elastic regime. A collagen-I biopolymer network is used for validation. The method enables parametric studies of the network geometry, whose parameters are often difficult to vary independently in experiments.

I. INTRODUCTION

Biopolymers are the building blocks of living tissues. Within the cell, microtubules form stiff rods [1], while actin and intermediate filaments assemble into flexible or semi-flexible structures [1,2]. In the extracellular matrix, collagen and fibrin self-assemble into stiff bundles, which are themselves part of a complex hierarchical structure [3,4]. These diverse building blocks cross link into networks, which display a wide range of mechanical properties [5]. Particularly, biopolymer networks exhibit strain stiffening and negative normal stresses when subjected to simple shear. The bulk mechanical properties of a biopolymer network depend on the constitutive relations of filaments [6,7] and branch points [8], as well as the network architecture [9,10].

In this paper, we are concerned with effects of the network geometry on bulk mechanical properties. We show that the architecture of collagen networks can be quantified using at most five independent parameters, we present an algorithm for generating network geometries from those parameters, and we show that the regenerated network geometries are mechanically equivalent to physical collagen network geometries by comparing their response in simple shear using the finite element method (FEM), in which each filament is represented by a three-node Timoshenko beam element. Finally, we show an example of how the network generation algorithm can be employed to study the effects of the network geometry on bulk mechanical properties.

II. NETWORK CHARACTERIZATION

We use reconstituted collagen type I networks as a model system in our study. Bovine collagen type I is labeled with the fluorescent molecule tetramethylrhodamine isothiocyanate (TRITC) using a protocol similar to that previously described [11]. Networks are polymerized at a final collagen concentration of 1.0 mg/ml [12]. A voxel-based representation of the network is obtained using confocal microscopy [Fig. 1(a)] and converted into a three-dimensional Euclidean graph representation using the skeletonization package in Amira (TGS, San Diego, CA). Each edge of the graph represents a filament and each node is a branch point [9,10,13].

What features of the geometry of a biopolymer network govern its bulk mechanical properties? By visual inspection of Fig. 1(a), we conclude that the network is essentially homogeneous at length scales much larger than the mesh size. Wyart and co-workers [10] showed that the coordination number Z, which is the average number of filaments incident to each branch point, has an order of magnitude effect on bulk stiffness in two dimensions. Heussinger and Frey [14] reported that the number of branch points per fiber—fiber means several connected filaments with a direction correlation—significantly affects the low strain response in shear. It is thus clear that the macroscale homogeneity, the number of filaments at each branch point, and the direction correlations of those filaments must be captured by the network description, at the very least.

Guided by these arguments we propose a statistical network description (see Fig. 2), which includes three probability densities normalized to unity: (a) The filament length distribution $P(l)$, which enforces macroscale homogeneity by precluding too many short or long filaments. Here, l denotes filament length normalized by $n^{-1/3}$, where n is the node
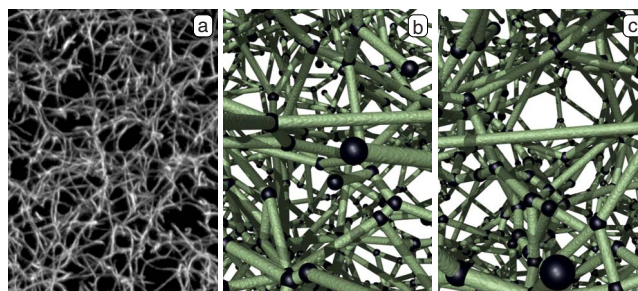

FIG. 1. (Color online) (a) Confocal microscope image of collagen-I, 1.0 mg/ml, with image width of 100 μm and depth of 5 μm. (b) Euclidean graph representation of collagen obtained from confocal microscopy data. (c) Realization of a statistical description of the geometry. Dangling ends are due to cropping of the network for the purpose of visualization.

*stefan.lindstroem@gmail.com
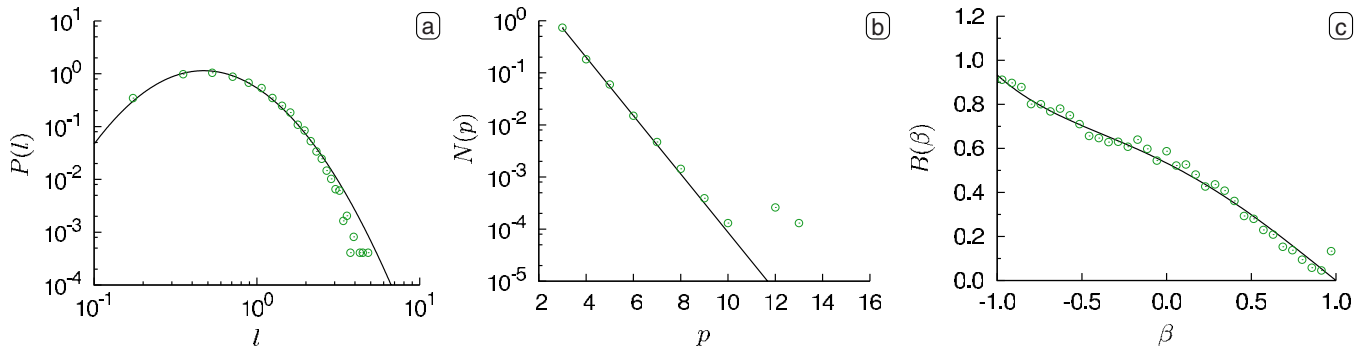†weitz@seas.harvard.edu

FIG. 2. (Color online) Geometry statistics of 1.0 mg/ml collagen-I with each probability density normalized to unity. (a) Measured normalized filament length distribution with a fitted logarithmic-normal distribution. (b) Measured valency distribution with a fitted geometric distribution. (c) Measured direction cosine distribution with a fitted two-parameter power-series distribution.

density. In physical units, the average filament length is $L = 2.0$ μm. (b) The valency distribution $N(p)$, which enforces the correct topology; the valency $p \in [3, \infty)$ is the number of incident edges to a node. (c) The direction cosine distribution $B(\beta)$, which is the probability density of cosines β between edges incident to the same node and enforces direction correlations within the network.

For our collagen-I networks, we observe that the filament length distribution is logarithmic-normal:

$$P(l) = \frac{1}{ls\sqrt{2\pi}} \exp\left[-\frac{(\mu - \ln l)^2}{2s^2} \right],$$ (1)

where μ and s are the mean and standard deviation of $\ln l$. The valencies follow a shifted geometric distribution,

$$N(p) = q(1 - q)^{p-3},$$ (2)

where $q = 1/(Z - 2)$ and the coordination number Z is the average valency. The distribution of direction cosines is well represented by a truncated power series,

$$B(\beta) = \sum_{k=1}^{m} b_k (1 - \beta)^{2k-1}.$$ (3)

Choosing $m = 3$ and normalizing $B(\beta)$ to unity leave two parameters b_1 and b_2. In total, five independent parameters μ, s, Z, b_1, and b_2 are used for the network description.

III. EUCLIDEAN GRAPH GENERATION

We propose a Euclidean graph generation (EGG) algorithm for regenerating three-dimensional network geometries on a periodic cubic domain Ω of size Λ using simulated annealing. Simulated annealing is a general heuristic method for global optimization. Previously, simulated annealing methods have been employed to generate voxel-based representations of two-phase materials [15,16], for instance, rock structures [17], glass bead assemblies [18], and fiber networks slightly deviating from a regular grid [19]. However, there has been no previous attempt to generate collagen fiber networks using simulated annealing.

An initial graph configuration $\mathcal{H}_0$ is generated by placing nodes drawn from a randomly uniform distribution inside the

domain Ω. Then, each node is attributed a valency drawn from N. Finally, pairs of randomly chosen nodes are connected by edges, so that the assigned valency is not exceeded for any node. Obviously, this initial configuration does not follow the length distribution or direction cosine distribution.

We define a neighbor $\mathcal{H}'$ of a graph $\mathcal{H}$ as a graph obtained by either (a) removing two edges from $\mathcal{H}$ and then adding two edges without changing the valency of any node or (b) moving the position of one node in $\mathcal{H}$ a random distance in the range $[0, \rho]$ in a random direction (see Fig. 3). Here, ρ is a parameter which may be tuned to enhance convergence of the algorithm (in this work, $\rho/\Lambda = 5 \times 10^{-2}$). Note that the node valency distribution is unchanged subject to (a) or (b) modification. We define a non-negative "energy" function to be $E(\mathcal{H}) = A_P(\mathcal{H}) + A_B(\mathcal{H})$, where A_P and A_B are the Cramér-von Mises test statistics for P and B, respectively (see the Appendix). With this definition, $E(\mathcal{H})$ has a global minimum for graphs with the target length distribution and direction cosine distribution. Simulated annealing minimizes $E(\mathcal{H})$, and thus finds the target structure, by iteratively attempting to modify the current graph $\mathcal{H}_m$ into a randomly chosen neighboring graph $\mathcal{H}_{m+1} = \mathcal{H}'_m$. Type (a) and (b) transition attempts are equally probable. A transition is accepted when $E(\mathcal{H}'_m) \leq E(\mathcal{H}_m)$ or accepted with probability $\exp\{[E(\mathcal{H}_m) - E(\mathcal{H}'_m)]/T\}$ when $E(\mathcal{H}'_m) > E(\mathcal{H}_m)$. Here, the "temperature" T is chosen to be an exponentially decreasing function of the number of accepted transitions, ensuring pro-
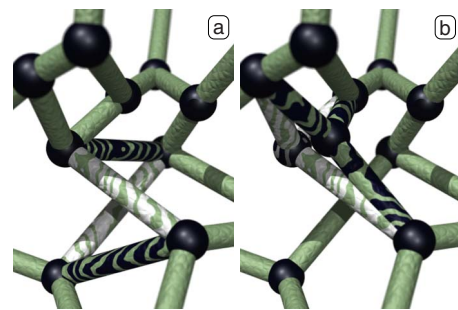

FIG. 3. (Color online) Neighboring graph configurations (black stripes) of the present configuration (white stripes). Two types of elementary modifications are allowed: (a) shifting a pair of edges while maintaining node valency and (b) change of one node position.
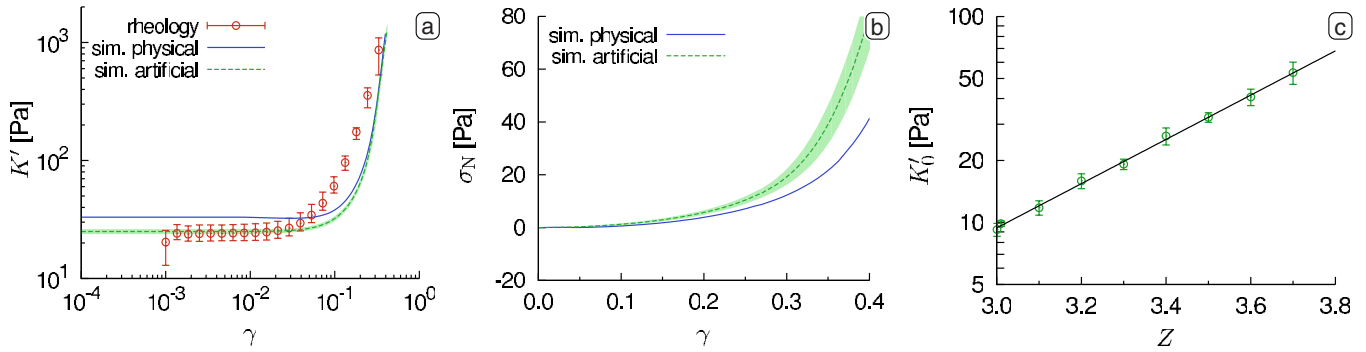
FIG. 4. (Color online) (a) Bulk shear modulus K' as a function of applied shear strain γ from experiments (circles) and from FEM analyses of a graph representation of a collagen-I network (solid line) and regenerated network (dashed line) with standard deviation (shaded region). (b) Normal stress σ_N as a function of applied shear strain γ, obtained from FEM analyses of a network graph representation (solid line) and regenerated network (dashed line) with standard deviation (shaded region). (c) Low strain shear modulus K'_0 of regenerated networks as a function of coordination number Z (circles), with a least-squares fitted exponential function (line).

gressive reduction in energy-increasing transitions. Its decay rate may be modified to enhance convergence.

IV. EXPERIMENTS AND DISCUSSION

With the EGG algorithm, we generate artificial networks which follow the description of the geometry $\{P, N, B\}$. It is left to prove that these artificial networks have the same mechanical properties as the physical network geometries. Artificial networks of 5000 nodes are generated on a cubic domain whose side measures approximately 20 mesh sizes. A physical network and its regenerated counterpart are compared in Figs. 1(b) and 1(c).

The physical and artificial network geometries are analyzed in a shear cell geometry using FEM with affine deformations prescribed to the boundaries. Collagen strands are modeled by three-node quadratic Timoshenko beam elements, allowing for tensile deformation, bending, torsion, and transverse shear. This element allows for linearly varying bending moments and takes large deformations and rotations into account. The cross section of the beam is assumed to vary with the axial stretch to maintain a constant volume. The beam elements are coupled in all degrees of freedom at the nodes, that is, the branch points are assumed to be rigid. Material parameters are measured through analyses of the thermal vibrational modes of the fibers [12]: the bending stiffness of the fibers is $EI = 5 \times 10^{-22}$ N m^2 and the transverse shear stiffness is 6×10^{-11} N. Assuming that fibers have a circular cross section of diameter 1.0×10^{-7} m, as obtained from transmission electron microscopy measurements [12], the elongational stiffness is computed to $EA = 8 \times 10^{-7}$ N. There are thus no fitting parameters in the constitutive model. Each numerical experiment is repeated for five different network realizations.

The FEM simulations render shear stress τ and normal stress σ_N as functions of shear strain γ. The bulk shear moduli $K' = d\tau/d\gamma$ of real and artificial networks are compared in Fig. 4(a) over a range $\gamma \in [0, 0.4]$. Within this range, they are equal within $\pm 40\%$. Similarly, the development of the normal stress σ_N is plotted in Fig. 4(b). The normal stress is more challenging to predict, because it measures the

difference of stresses in the principal strain directions. At large strains, artificial networks exhibit 50% greater normal stress than physical networks.

We also polymerize collagen gels in a plate-plate geometry of an AR-G2 stress-controlled rheometer (TA Instruments, New Castle, DE). The measured development of K' with γ is compared with simulation results in Figs. 4(a). The low strain behavior is predicted by simulations with one significant digit. This level of accuracy is consistent with the 10–20 % error of each measured parameter. The observed low strain shear modulus $K'_0 = 24$ Pa is on the order of the modulus of networks of beams deforming in bending $EI/L^4 = 30$ Pa, while it is inconsistent with the low strain modulus of entropic networks $nk_B T = 300$ μPa and networks of beams loaded in tension $EA/L^2 = 200$ kPa.

The onset of strain stiffening occurs at a lower strain in experiments than in FEM simulations. At the same time, the stiffening gradient of the experiment is smaller than that in simulations. These discrepancies are seen at large strains for both the physical and the artificial networks. Hence, they are not due to failure of the EGG algorithm, but to neglected or overseen effects in the finite element model. Since the simulated physical and artificial network responses are in good agreement, we can still conclude that the suggested statistical representation of the networks is adequate. The difference between experimental results and the numerical analysis at large deformations remains to be explained. Neglected effects include the nonstraight shape of the filaments [Fig. 1(a)], contact interactions between filaments, the nonlinear constitutive relation [20,21], and failure.

To show the predictive power of our approach, we proceed with a sample parametric study. The coordination number, which is $Z \approx 3.4$ in the physical collagen network, is varied in the range $Z \in (3.0, 3.7]$, while the line density and parameters μ, s, b_1, and b_2 are kept constant. The generated networks are analyzed for low strain shear modulus K'_0, as shown in Fig. 4(c). There is an exponential dependence of the modulus on the coordination number: $K'_0 = ke^{\alpha(Z-3)}$, where $k = 9.4$ Pa and $\alpha = 2.5$ in the present case. This shows that the mechanics of networks of different topologies, but with the same line density, can differ significantly.

There are many materials, and many material properties, for which a Euclidean graph representation of the geometry is suitable, for instance, conductivity or permeability of colloidal gels and granular or porous media. For each type of heterogeneous material, and for each material property of interest, a different characterization of the geometry may be required. Since parameters describing the geometry of these materials depend on complicated or even unknown forming processes, it is often difficult to systematically vary these parameters in experiments. The EGG algorithm may thus be applicable to many materials with a hierarchical structure, for parametric studies as well as for *objectively* determining whether a given statistical description of the geometry is sufficient to capture a given material property.

V. CONCLUSIONS

With this work, we have shown that the EGG algorithm, if used with an appropriate statistical description of a physical network geometry, generates artificial networks which capture the bulk mechanical properties of the corresponding physical network, at least in the elastic regime. This framework makes possible systematic studies of both effects of the network geometry and effects of the constitutive relations of the filaments on bulk mechanical properties. Important outstanding questions concern the origin of strain stiffening, a prediction of its onset, and the effects of plastic deformations and relaxation of the filaments on network mechanics.

ACKNOWLEDGMENTS

APPENDIX: HISTOGRAM OF THE CRAMÉR-VON MISES TEST

According to Anderson [22], the hypothesis that a set of observed variables $x_1 < x_2 < \cdots < x_n$ originate from a random process with distribution f can be rejected if

$$A_f = \frac{1}{12n} + \sum_{j=1}^{n} \left[\frac{2j-1}{2n} - F(x_j) \right]^2 \qquad \text{(A1)}$$

exceeds some tabulated value. Here, F is the cumulative distribution function of f. Thus, of two different sets of data, the set producing the smallest value of A_f is in a sense the better fit to the distribution f.

The evaluation of Eq. (A1) takes $\mathcal{O}(n)$ operations and also requires that one maintains a sorted set of variables. We can, however, make an approximate evaluation by first computing a histogram of p bins for the observed variables and reduce summation complexity to $\mathcal{O}(p)$.

The histogram of a set $X = \{x_1, x_2, \ldots, x_n\}$, which need not be ordered, is defined as a set of integers $m_j = |\{x_k : x_k \in I_j\}|, j = 1, 2, \ldots, p$, where $|\cdots|$ denotes cardinality and $I_j = [y_j, y_{j+1})$ are disjoint intervals whose union covers at least the range of X. We make the estimate $A_f(X) \approx A_f(X^*)$, where X^* is some set with the same histogram m_j as X. For simplicity, we choose X^* so that it contains exactly m_j instances of the same value x_j^* being the midpoint of I_j. The sum of Eq. (A1) thus becomes

$$\frac{1}{n^2} \sum_{j=1}^{p} \sum_{k=1}^{m_j} \left[k + M_{j-1} - \frac{1}{2} - nF(x_j^*) \right]^2$$

$$= \frac{1}{n^2} \sum_{j=1}^{p} \left[\sum_{k=1}^{m_j} k^2 + 2S_j \sum_{k=1}^{m_j} k + S_j^2 \sum_{k=1}^{m_j} 1 \right]$$

$$= \frac{1}{n^2} \sum_{j=1}^{p} m_j \left[\frac{1}{6}(m_j+1)(6S_j + 2m_j + 1) + S_j^2 \right],$$

where $M_j = \Sigma_{k=1}^{j} m_k$ is the cumulative histogram and $S_j = M_{j-1} - nF(x_j^*) - \frac{1}{2}$ for brevity.

[1] C. P. Brangwynne, G. H. Koenderink, E. Barry, Z. Dogic, F. C. MacKintosh, and D. A. Weitz, Biophys. J. **93**, 346 (2007).

[2] J. E. Eriksson, T. Dechat, B. Grin, B. Helfand, M. Mendez, H.-M. Pallari, and R. D. Goldman, J. Clin. Invest. **119**, 1763 (2009).

[3] L. Yang, K. O. van der Werf, C. F. C. Fitie, M. L. Bennink, P. J. Dijkstra, and J. Feijen, Biophys. J. **94**, 2204 (2008).

[4] H. Kang, Q. Wen, P. A. Janmey, J. X. Tang, E. Conti, and F. C. MacKintosh, J. Phys. Chem. B **113**, 3799 (2009).

[5] C. Storm, J. J. Pastore, F. C. MacKintosh, T. C. Lubensky, and P. A. Janmey, Nature (London) **435**, 191 (2005).

[6] F. C. MacKintosh, J. Kas, and P. A. Janmey, Phys. Rev. Lett. **75**, 4425 (1995).

[7] A. Kabla and L. Mahadevan, J. R. Soc., Interface **4**, 99 (2007).

[8] K. E. Kasza, G. H. Koenderink, Y. C. Lin, C. P. Broedersz, W. Messner, F. Nakamura, T. P. Stossel, F. C. MacKintosh, and D. A. Weitz, Phys. Rev. E **79**, 041928 (2009).

[9] P. R. Onck, T. Koeman, T. van Dillen, and E. van der Giessen, Phys. Rev. Lett. **95**, 178102 (2005).

[10] M. Wyart, H. Liang, A. Kabla, and L. Mahadevan, Phys. Rev. Lett. **101**, 215501 (2008).

[11] A. Baici, G. Cohen, K. Fehr, and A. Böni, Anal. Biochem. **108**, 230 (1980).

[12] David A. Vader, Ph.D. thesis, Harvard University, 2010.

[13] A. M. Stein, D. A. Vader, L. M. Jawerth, D. A. Weitz, and L. M. Sander, J. Microsc. **232**, 463 (2008).

[14] C. Heussinger and E. Frey, Phys. Rev. Lett. **97**, 105501 (2006).

[15] C. L. Y. Yeong and S. Torquato, Phys. Rev. E **57**, 495 (1998).

[16] Y. Jiao, F. H. Stillinger, and S. Torquato, Phys. Rev. E **76**, 031110 (2007).

[17] Y. Jiao, F. H. Stillinger, and S. Torquato, Phys. Rev. E **77**, 031135 (2008).

[18] H. Lee, M. Brandyberry, A. Tudor, and K. Matous, Phys. Rev. E **80**, 061301 (2009).

[19] F. S. A. Cavalcante, S. Ito, K. Brewer, H. Sakai, A. M. Alencar, M. P. Almeida, J. S. Andrade, Jr., A. Majumdar, E. P. Ingenito, and B. Suki, J. Appl. Physiol. **98**, 672 (2005).

[20] J. A. J. van der Rijt, K. O. van der Werf, M. L. Bennink, P. J. Dijkstra, and J. Feijen, Macromol. Biosci. **6**, 697 (2006).

[21] Z. L. Shen, M. R. Dodge, H. Kahn, R. Ballarini, and S. J. Eppell, Biophys. J. **95**, 3956 (2008).

[22] T. W. Anderson, Ann. Math. Stat. **33**, 1148 (1962).