

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Biopolymer Networks: Image Analysis, Reconstruction and Modeling

*A thesis presented in partial fulfilment of the requirements
for the degree of*

*Doctor of Philosophy
in
Physics*

*at Massey University, Manawatū,
New Zealand*

Pablo Hernandez-Cerdan

2018



It's not about the end, it's all about the path.

¡Buen camino!

Abstract

The aim of this work is to extract the architecture of biopolymer networks from 3D images. This is motivated to further understand non-affine regimes, found during network formation and in low-density biopolymer networks, where the geometry of a network has a fundamental role in defining its mechanical properties.

Firstly, developed image analysis tools were extended to 3D and contributed to high-performance open-source libraries for image analysis. These developments in isotropic wavelets will help in extracting realistic networks by removing spurious noise generated during image acquisition.

Secondly, images of biopolymer gels from transmission electron microscopy (TEM), were used to reliably extract the network architecture. The imaged material was also studied with small-angle x-ray scattering (SAXS), and the comparison showed a strong agreement for network-size features.

Thirdly, spatial graphs were extracted from the image. A one-to-one map is provided between image and graph, keeping all the geometric information from the image. This then opened the door to using analytical tools from the complex networks field to characterize images.

Finally, statistical distributions extracted from three graph properties were used to reconstruct a completely in-silico network using a simulated annealing technique to generate new networks. This can be used as a computational exploration tool of how network behavior depends on network architecture.

Acknowledgements

A special thank to my supervisor Bill Williams whose support, openness, and vision of science is encouraging. As Brad would say, I will be forever indebted to your supervision and I add: and from all the stuff that I have learned from you beyond academia, specially your humane values and the love for dumplings. Thanks to Brad Mansel, Leif Lundin and Andrew Leis for the collaboration comparing scattering and imaging techniques, Allan Raudsepp for providing the confocal microscopy images for actin, and to all my group mates from the [Biophysics and Soft Matter group](#), for all the shared moments, and how easy they made my adaptation to a new country. Finally, I am grateful to the [MacDiarmid Institute](#) for constantly providing the enrichment opportunity to meet other peer students and researchers, and to the staff of the Institute of Fundamental Sciences for its kindness dealing with all sort of issues.

In the personal side, I would like to thank my family, from who I have inherited, directly and indirectly, all the principles I stand for and made me realize how privileged I am to have them. And Martica! Without you here in New Zealand this quest would have been impossible, or faster, who knows, but way less enjoyable for sure. It has been a pleasure habernos re-encontrado por kiwi-land. And even we got married, in English!!, crazy, but made me feel extremely fortunate. Thanks really for all your support.

This research was supported by the MacDiarmid Institute, the Riddet Institute, and Massey University, New Zealand.

Contents

Abstract	ii
Acknowledgements	iii
Contents	iv
List of Figures	vii
List of Publications	x
Abbreviations	xi
1 Introduction	1
1.1 Soft matter and biopolymers	1
1.2 Semiflexible single chains: WLC model	2
1.3 Networks of semiflexible chains:	5
1.4 Bulk Rheology	8
1.4.1 Linear viscoelasticity	8
1.4.2 Nonlinear behaviour: strain stiffening	9
1.5 Network Structure	11
1.5.1 Mikado Networks	11
1.5.2 Lattice Models	12
1.5.3 Eight-chain Models	13
1.6 Motivation	14
2 Isotropic and Steerable Wavelets in N Dimensions. A multiresolution analysis framework	16
2.1 Motivation	16
2.2 Abstract	16
2.3 Wavelet Multiresolution Analysis	17
2.3.1 Introduction	17
2.3.2 Motivation: spatial and frequency resolution	18
2.3.3 Wavelet transformation	19
2.3.4 Wavelet Pyramid	21
2.3.5 Isotropic wavelets	22
2.4 Riesz Transform	25
2.4.1 Monogenic Signal	26

2.4.2	Generalized Riesz Transform	27
2.4.3	Generalized Steerable Framework	27
2.5	Implementation Details	28
2.5.1	Summary	28
2.5.2	Frequency Iterators	29
2.5.3	Wavelet Transform	30
2.5.4	Riesz Transform	31
2.5.5	Structure Tensor	31
2.5.6	Phase Analysis	32
2.6	A guided example:	32
2.6.1	Input in the frequency domain.	33
2.6.2	Choosing an isotropic wavelet	33
2.6.3	Forward / Analysis	33
2.6.4	Inverse / Reconstruction	34
2.6.5	PhaseAnalyzer	36
2.7	Conclusion and future work	41
2.8	Acknowledgements	42
3	Validation of Transmission Electron-Microscopy Imaging: Comparison with SAXS	43
3.1	Motivation	43
3.2	Abstract	43
3.3	Introduction	44
3.4	Materials and methods	45
3.5	Results and Discussion	48
3.5.1	Qualitative Structural Features	49
3.5.2	Consistency Between Techniques at Different Lengthscales	50
3.5.3	Image Processing	52
3.5.4	Extracting Persistence Lengths	54
3.6	Conclusions	55
4	Spatial Graph Extraction: Analysis of microscopy images to obtain network architecture	58
4.1	Motivation	58
4.2	Introduction	58
4.2.1	Notation	60
4.3	Denoising	61
4.4	Segmentation	62
4.4.1	Binarization	63
4.4.1.1	Region Growing Segmentation	64
4.4.2	Hole Filling	65
4.5	Skeletonization	66
4.5.1	Distance maps	68
4.5.2	Critical Kernels Framework	68
4.6	Spatial Graph Extraction	72
4.7	Statistical distributions of graph properties	74
4.8	Testing existing thinning algorithms	75

4.8.1	Fiber Extraction -FIRE-	75
4.8.1.1	Results in FIRE	77
4.8.2	Avizo: XSkeleton	79
4.8.2.1	Avizo results:	81
4.9	Application of the Network Analysis Pipeline to Different Biopolymer Networks.	84
4.9.1	Confocal Light Microscopy: Actin	84
4.9.2	TEM preparation	90
4.9.3	Transmission Electron Microscopy: Carrageenan	91
4.9.4	Transmission Electron Microscopy: Pectin	94
4.10	Conclusions	97
5	Reconstructing networks from statistical distributions	99
5.1	Motivation	99
5.2	Methods	99
5.3	Euclidean Graph Generation Algorithm	102
5.4	Results	103
6	Conclusions and Future Work	106
6.1	Scope of Thesis	106
6.2	Summary	106
6.3	Conclusions	108
6.4	Future work	109
6.4.1	Network Formation	109
6.4.2	Long time behaviour: network quakes and aging	109
6.4.3	ITKBoostGraph	110
6.4.4	Complex Networks Tools	110
6.4.5	Public Database	110
A	Publications and Open Source Contributions	111
A.1	Open Source Contributions	111
A.1.1	ITKIsotropicWavelets	111
A.1.2	Radial Intensity FFT	111
A.1.3	FFT Radial Intensity	112
A.1.4	DGtal: Critical Kernels Framework	112
A.1.5	Spatial Graph Extraction	113
A.1.6	Open Science	113
	Glossary	114
	References	116

List of Figures

1.1	Optical Tweezers	5
1.2	Force extension curve: CEWLC	6
1.3	Strain-stiffening in semiflexible polymers	9
1.4	Affine and non-affine deformations	10
1.5	Affine and non-affine phases	11
1.6	Mikado network used as a starting scaffold for dynamics simulations, from Ref. [31].	12
1.7	Unit cell of a face-centered-cubic lattice, from Ref. [40]	13
1.8	Eight-chain model under shear, from left to right: initial state, a rotation, and a simple shear, from Ref. Palmer and Boyce [30]	14
2.1	The changes over time/space of a signal are not captured by the FFT.	18
2.2	Simultaneous spatial and frequency resolution of different transformations. $\Delta t, \Delta f$	19
2.3	Forward 2.3a and Inverse wavelet 2.3b two-level pyramid with two high pass sub-bands.	21
2.4	Forward wavelet pyramid with a classic two level filter bank and only one sub-band, HP is the high pass filter, and LP low pass	22
2.5	Tiling of the frequency domain by isotropic wavelets when the dilation factor is 2. All the wavelets fulfill the conditions from Proposition 2.3. The mother wavelets are represented at $i = 0$	24
2.6	Shape of SubBands when HighPassSubBands = 5, sub-bands increase the frequency resolution, but can generate extra artifacts in the spatial domain due to the sharp frequency cut-offs.	25
2.7	The original image looks like a regular check-board, but it isn't. Pixels in the regions A and B have the same intensity value (129), however our vision system performs local phase analysis that allows us to treat A,B regions as different, keeping a global checker-board structure. 2.7b uses a non-linear map of intensity-color to enhance the irregular checker board structure.	37
2.8	(Results of the phase analysis (with soft threshold) for different number of scales in the wavelet pyramid. The input image is a checker-board of size 512×512).	38
2.9	Using six scales (Level: 6), results for different number of high frequency sub-bands.	39
2.10	The optical illusion generated by 2.10a the Hermann grid is generated by the local phase analysis of our vision. 2.10b is the result of a phase analysis with the Monogenic signal for a wavelet pyramid of 8 levels and only one high pass sub band. 2.10c is the same pyramid but with 10 high pass bands to increase the frequency resolution. Both results use Simoncelli wavelet, but there is not much difference using Held, or Vow mother wavelets.	40

2.11	Biopolymer network (actin) 3D image enhanced and denoised using phase analysis with the Monogenic signal, with a wavelet pyramid of 4 levels and 4 high pass sub-bands using Simoncelli isotropic wavelet.	41
3.1	The Graphical User Interface (GUI) of the image analysis software developed herein. A non-interactive interface is also provided for batch processing, along with python scripts for plotting and manipulation of the generated data. . . .	48
3.2	Montages of TEM images obtained from different polysaccharides, as described in the text.	49
3.3	SAXS results for the three samples shown in the TEM micrographs of figure 3.2, in the absence of TEM sample preparation. Black lines signify power law fitting, as is described in section 3.1 of the text.	50
3.4	Comparison of data obtained from TEM best practice and SAXS. Good agreement can be seen in the large network scale (at low qs). The shaded area shows the upper q range (smaller than 20 nm or equivalently 25 pixels). The top x-axis shows the corresponding real-space length scale, where $d = 2\pi/q$. .	53
3.5	Results obtained when applying two widely-used de-noising techniques: total variation regularization (TV) and multi-scale wavelet denoising using a BLS-GSM algorithm.	53
3.6	Visual comparison of the de-noised pectin TEM images (Fig 3.2a). (a) shows a region from the unmodified original image, (b): de-noised image using a total variation method [78] with $\lambda = 0.15$, (c) shows the same region after applying a multi-scale wavelet approach using the BLS-GSM algorithm [58] with $\sigma = 28.4$. Only high-frequency regions are affected by de-noising algorithms, as reflected in the generated I-q plots shown in Fig 3.5a.	54
3.7	A TEM micrograph with FibreApp tracking output shown in blue. The insert shows the internal contour length vs mean-square end-to-end distance calculated from the corresponding tracking parameters.	55
4.1	Visualization of a Spatial Graph (4.1a) and the corresponding adjacency list (4.1b). The graph is represented in blue, formed by nodes (voxels with id) and edges connecting them. The voxel positions (in brackets in 4.1a) of nodes and edge points are stored as labels in the adjacency list.	59
4.2	Pre-processing the image (actin): denoising with wavelets and total variation methods.	62
4.3	Different threshold values on networks	63
4.4	Thin image of a pectin network after skeletonization performed with the critical kernels framework (see subsection 4.9.4) and visualized with a 3D Viewer. The skeleton keeps the topology of the original object, is thin, centered, and robust against rough and noisy boundaries.	67
4.5	Distance map - Actin	68
4.6	Asymmetric thinning algorithm with persistence, from Ref. [91]	70
4.7	Asymmetric thinning of a volume with different parameters. (a): Ultimate skeleton, only voxels that conserve topology are kept. (b): keeps 1-isthmus as part of the skeleton. (c): 1-isthmus with persistence parameter $p = 10$	71
4.8	Skeletonization	72
4.9	Extracting the raw graph from the thin image and post processed to remove the largest edges between tri-connected nodes with degree 3. Numbers show the number of edges (degree) of each node.	74

4.10	Spatial Graph where vertices with degree 2 are deleted, and its position are stored as edge points in the new edges. Numbers in circles represent degree of each node.	74
4.11	FIRE algorithm - Find LMP	77
4.12	Fire: Step by step for $T_P = 0.12$	78
4.13	Distributions of length and degree in FIRE	79
4.14	Avizo image: Actin network visualization and workspace	80
4.15	Distributions of length, degree, and cosines with Avizo for actin with $T=30,P=5$	82
4.16	Distributions of length, degree, and cosines with Avizo for actin with $T=21,P=5$	83
4.17	Actin spatial graph extraction steps.	88
4.18	Statistical distributions (PDF) of actin network. Computed properties of the graph are represented in histograms with bins and normalized by area to obtain the PDF. Orange line shows the function with parameters obtained from a non-linear least squares fit to the data. Green lines show the same function with fixed parameters: 4.18a: mean degree (Z) of the network, 4.18c: mean (μ_l) and standard deviation (s_l) of logarithmic end-to-end distances, 4.18d mean (μ_l) and standard deviation (s_l) of logarithmic contour lengths. Histogram bins: end-to-end distances (30), direction cosines (21) contour lengths (30)	90
4.19	Potassium carrageenan spatial graph extraction steps.	93
4.20	Statistical distributions (PDF) of Potassium Carrageenan network. Computed properties of the graph are represented in histograms with bins and normalized by area to obtain the PDF. Orange line shows the function with parameters obtained from a non-linear least squares fit to the data. Green lines show the same function with fixed parameters: 4.20a: mean degree (Z) of the network, 4.20c: mean (μ_l) and standard deviation (s_l) of logarithmic end-to-end distances, 4.20d mean (μ_l) and standard deviation (s_l) of logarithmic contour lengths. Histogram bins: end-to-end distances (50), direction cosines (21) contour lengths (50)	94
4.21	Pectin spatial graph extraction steps.	96
4.22	Statistical distributions (PDF) of Pectin network. Computed properties of the graph are represented in histograms with bins and normalized by area to obtain the PDF. Orange line shows the function with parameters obtained from a non-linear least squares fit to the data. Green lines show the same function with fixed parameters: 4.22a: mean degree (Z) of the network, 4.22c: mean (μ_l) and standard deviation (s_l) of logarithmic end-to-end distances, 4.22d mean (μ_l) and standard deviation (s_l) of logarithmic contour lengths. Histogram bins: end-to-end distances (50), direction cosines (21) contour lengths (50)	97
5.1	Collagen distributions	100
5.2	Testing with theoretical collagen data	104
5.3	Collagen: comparing target and simulated distributions	105
6.1	Scheme of the thesis	108

List of Publications

- [1] Pablo Hernandez-Cerdan. “Isotropic and Steerable Wavelets in N Dimensions. A Multiresolution Analysis Framework for ITK.” In: *arXiv:1710.01103 [cs]* (Oct. 2017). arXiv: [1710.01103 \[cs\]](https://arxiv.org/abs/1710.01103).
- [2] Pablo Hernandez-Cerdan, Bradley W. Mansel, Andrew Leis, Leif Lundin, and Martin A.K. Williams. “Structural Analysis of Polysaccharide Networks by Transmission Electron Microscopy: Comparison with Small-Angle X-Ray Scattering.” In: *Biomacromolecules* (Jan. 2018). ISSN: 1525-7797. DOI: [10.1021/acs.biomac.7b01773](https://doi.org/10.1021/acs.biomac.7b01773).
- [3] Hina Shah, Pablo Hernandez, Francois Budin, Deepak Chittajallu, Jean-Baptiste Vimort, Rick Walter, André Mol, Asma Khan, and Beatriz Paniagua. “Automatic Quantification Framework to Detect Cracks in Teeth.” In: *Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging*. Vol. 10578. International Society for Optics and Photonics, Mar. 2018, 105781K. DOI: [10.1117/12.2293603](https://doi.org/10.1117/12.2293603).

Abbreviations

AFM	Atomic force microscopy.
BGL	Boost Graph Library. c++ library.
CEWLC	Clickable extensible wormlike chain.
DFT	Discrete Fourier Transform, representation of data in frequency domain.
DGtal	Digital Geometry Tools and Algorithms. c++ library.
EtE	End to end.
EWLC	Extensible wormlike chain.
FE	Force extension (curve).
FFT	Fast Fourier Transform, specific algorithm of DFT.
ITK	Insight ToolKit. Image Analysis c++ library specialized to work with ND images.
LAOS	Large amplitude oscillatory shear (rheology).
LMP	Local maximum point.
MRA	Multiresolution framework.
MSD	Mean square displacement.
NP	Nucleation point.
PME	Pectin-methylesterase.
SAXS	Small angle X-ray scattering .
SEM	Scanning electron microscopy.
TEM	Transmission electron microscopy.
TV	Total Variation (regularization, denoising method).
WLC	Wormlike chain.