



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Regulation of topological entanglement in ring polymers

Andrea Bonato



Doctor of Philosophy
The University of Edinburgh
February, 2023

Abstract

Entanglement abundance and complexity can be both beneficial and detrimental to the biological and mechanical function of polymers. In living organisms, for instance, DNA entanglement is so impactful that its proliferation might have catastrophic consequences, such as mutation and death.

Even though entanglement regulation is therefore necessary to keep the behaviour of both biological and synthetic polymer systems under control, how it is practically achieved is currently not well understood for many systems. To fill this void, by formulating analytic predictions and performing computer simulations, we study the equilibrium properties of sets of geometrically and topologically constrained ring polymers, and model how entanglement abundance and complexity is regulated in both biological and synthetic polymer systems.

We find that a complex of rings undergoing recombination under confinement is distinguished by the presence of a topological gelation transition, which can be controlled by the stiffness or the concentration of the rings. Furthermore, we show that an efficient and controlled way to resolve entanglement by transient cross stranding is having it compete entropically with a slip-linked polymer network. Finally, we study multicomponent polymer links where the monomers can be distributed among the components in all possible ways, and show that, asymptotically, due to entropy maximisation, one of the rings grows at the expense of the others, which behave as roots sliding along the contour of the growing component.

Lay Summary

Polymers are long chains of repeated molecular units held together by covalent bonds. They are ubiquitous in both nature and industry, biological polymers range from DNA to proteins, and synthetic material composed of polymers such as fibres, plastics and gels are part of our everyday life.

Due to their structure and flexibility, it is common for long cables and strings to be entangled, and untying a knot is often necessary to ensure the proper functioning of a laptop charger or a fishing line. The same holds true for polymers: for instance, the inadvertent emergence of DNA knots hinders biological processes such as DNA transcription. For this reason, managing entanglement abundance and complexity in polymer systems is often essential.

Motivated by this, by formulating analytic predictions and performing computer simulations, we study the properties and regulation of entangled polymers modelling real biological and synthetic systems.

Declaration

I declare that this thesis was composed by myself, that it has not been submitted in any other applications for a degree, and that the work presented is entirely my own, with the exceptions listed below.

- The problem tackled in chapter 3 was posed by Dr. D. Marenduzzo and Dr. E. Orlandini. The LAMMPS script for the molecular dynamics simulations from chapter 3 was first written by Dr. E. Orlandini and then modified by myself. I developed the mean field theory in collaboration with Dr. D. Michieletto.
- The idea behind the work presented in chapter 4 was suggested by Dr. D. Michieletto. The project was supervised by Dr. D. Marenduzzo and Dr. D. Michieletto.
- The project reported in chapter 5 was proposed by Dr. E. Orlandini, who also contributed in writing the code for the Monte Carlo simulations. Dr. S. G. Whittington proved Ths. 5 and 6, and, together with Dr. E. Orlandini, supervised the work.

Part of the results included in chapters 3, 4 and 5 have been published in peer-reviewed journals [1, 2, 3, 4].

I contributed in writing the manuscripts listed above.

Andrea Bonato, March 2022

Acknowledgments

During my PhD I had the fortune of being supervised by great researchers and people, without the help of whom finishing this work would have not been possible.

First, I express my gratitude for my main supervisor, Dr. Davide Marenduzzo, who not only guided me during my first experience in research, but also supported and encouraged me to come up with my own ideas and solutions, without skimping on his advice when needed. I extend my heartfelt thanks to my co-supervisor, Dr. Davide Michieletto, who has always been ready to give me a hand and guide me, and whose expertise greatly helped me taking my first steps in this line of work.

I could never forget Dr. Enzo Orlandini, who counselled me and directed many of the projects I worked on since before graduation in Padova. Thank you sincerely, Enzo. I also want to thank Dr. S. G. Whittington for overseeing and contributing to the work presented in chapter 5.

Finally, I want to say thank you to my family, my parents, who always supported me, in particular, my new and old friends at university, my flatmate Costi, and all my friends in Edinburgh, with whom I shared this fantastic experience.

Contents

1	Introduction	15
2	Critical behaviour of polymers in a good solvent, theory and simulations	21
2.1	Scaling behaviour of topologically constrained polymers	21
2.1.1	The self-avoiding walk and critical phenomena	21
2.1.2	Polymer networks	23
2.1.3	Ring polymers with fixed topology	25
2.2	Markov chain Monte Carlo	28
2.2.1	Ergodic Markov chains	28
2.2.2	Sampling	29
2.2.3	BFACF algorithm for linked loops	30
2.2.4	Ergodic classes of the BFACF algorithm for polydis- perse linked polygons	32
2.2.5	Limit distribution	34
2.3	Molecular dynamics simulations	35
2.3.1	NVE and velocity-Verlet	35
2.3.2	NVT	36
2.3.3	Optimisation	38
2.3.4	MD simulations of coarse-grained polymers	38
3	Topological gelation of confined recombinant ring	41
3.1	Model	43
3.2	Geometric properties of confined recombinant rings	44

3.2.1	Free energy of recombinant rings at equilibrium	46
3.2.2	Radial positioning and mixing of recombinant rings . . .	49
3.3	Topological gelation is accompanied by the formation of a percolating network of linked loops	51
3.4	Topological gelation traps the recombinant rings inside a permeabilised sphere	56
3.5	Discussion and conclusion	58
4	Simplifying topological entanglements by entropic competition of slip-links	63
4.1	Model	65
4.1.1	Sampling of the system	66
4.2	Competition of SLs	68
4.2.1	Equations for SLs vs SLs using Duplantier scaling for polymer networks	69
4.2.2	Scaling behaviours for competing SLs	71
4.3	SLs versus knots	74
4.4	Kinetics of topological simplification	76
4.5	Conclusions	78
5	Asymptotics of linked polygons	79
5.1	Rigorous results about the exponential growth rate	80
5.2	Growth rate	82
5.2.1	Critical exponents	87
5.2.2	Entropic competition	90
5.3	Discussion	99
6	Conclusions	101

List of Figures

1.1	Hermann Staudinger and polymers	16
1.2	DNA knots and links	16
1.3	DNA replication and site-specific recombination	18
2.1	Random and self-avoiding walks	21
2.2	Self-avoiding walks and loops	22
2.3	Self-avoiding polymer network	24
2.4	Topologically constrained ring polymers – knots and links . . .	25
2.5	Linking number	26
2.6	Knots localisation	27
2.7	Scheme of the BFACF algorithm	30
2.8	BFACF algorithm	32
2.9	Lattice link diagram	33
3.1	Sketch of ring polymers undergoing recombination under confinement	42
3.2	Scheme of the reconnection algorithm	43
3.3	Geometric properties of confined recombinant rings	45
3.4	Number of rings vs confinement	46
3.5	Distribution of length	48
3.6	Mixing and radial positioning of recombinant rings	50
3.7	Topological gelation	52
3.8	Distribution of linking number	53
3.9	Critical length for gelation	54

3.10	Network structure of linked rings in the gel phase	56
3.11	Escape dynamics of recombined rings from a permeabilised sphere	58
3.12	Escape dynamics of recombinant rings from a permeabilised sphere	59
4.1	Sketch of the entropic competition of SL and entanglements .	64
4.2	Conformational de-correlation time	67
4.3	Grafted round table and necklace networks	68
4.4	Entropic competition between symmetric SL networks	73
4.5	Entropic competition between SLs and a trefoil Knot	74
4.6	Kinetics of topological simplification	77
5.1	Connected sum of knots	81
5.2	Table of two-component links	83
5.3	Table of multi-component links	84
5.4	Extrapolation of the growth rate of two-components link . . .	86
5.5	Extrapolation of the growth rate of multi-component links . .	87
5.6	Extrapolation of the critical exponent of the Hopf link	88
5.7	Extrapolation of the critical exponent of two-component and multi-component link	89
5.8	Snapshots	91
5.9	Mean length of the largest component	92
5.10	Mean length of the components of prime two-component links	93
5.11	Mean length of the components of prime multi-component links	93
5.12	Mean length of the components of two-component, at least one of which is a knot, links	94
5.13	Mean length of the components of composite multi-component links	95
5.14	Length distributions of the components of two-component links	96
5.15	Length distributions of the components of multi-component links	97

Chapter 1

Introduction

The synthesis of polymer materials such as vulcanised rubber and artificial fibre gradually achieved enormous commercial success in the early 20th century. Nobel prize Hermann Staudinger (Fig. 1.1A) first understood the molecular nature of such materials. In the 1920s, he proposed that polymers were not colloids, as it was believed at the time, but long chains of repeated molecular units (called monomers) held together by covalent bonds (Fig. 1.1B). Staudinger's ground breaking idea was eventually accepted by the scientific community and ignited the long lasting interest of many branches of scientific research, including Chemistry, Biology, Physics and Mathematics, in the properties of natural and synthetic polymers. Nowadays, biopolymers such as DNA, RNA and proteins have been recognised to play a fundamental role in the lifecycle of living organisms and industrial polymeric materials, such as fibres, plastics, elastomers and gels are part of our everyday life.

When the degree of polymerisation (i.e. the number of monomers) is large, entanglement in polymer systems is highly likely. It was recognised experimentally, as in the case of cyclic DNA under confinement [5], but also proved mathematically: for instance, it was first conjectured by Frisch, Wasserman and Delbruck [6, 7] in the 1960s, and later proved rigorously [8, 9], that the knotting probability of sufficiently long ring polymers is close to one.

Close to inevitable, entanglement strongly impacts the biological [10, 11, 12, 13] and mechanical function [14] of polymers. In living organisms, for example, entangled DNA structures, such as knots (Fig. 1.2A) and catenanes (i.e. linked ring polymers, Fig. 1.2B), inadvertently form during the lifecycle of a cell, and their proliferation, if left untreated, would be catastrophic, and would eventually lead to cell mutation and death. This is what happens,

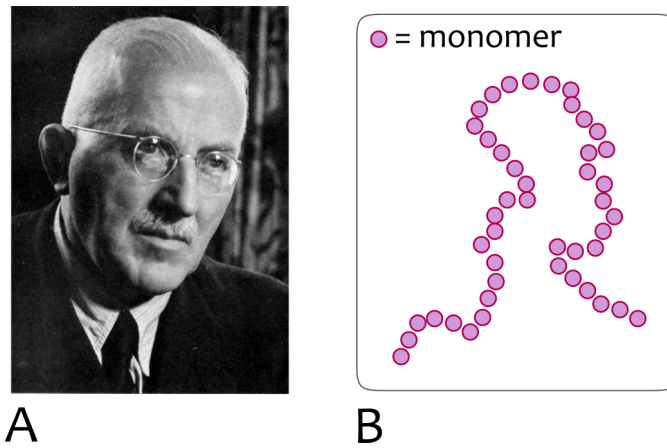


Figure 1.1: **A** Picture of the German chemist Hermann Staudinger, the first to understand the molecular structure of polymers and Nobel prize winner in 1953. Image from wikipedia.org. **B** Sketch of the molecular structure of a polymer, a long chain of repeated units called monomers.

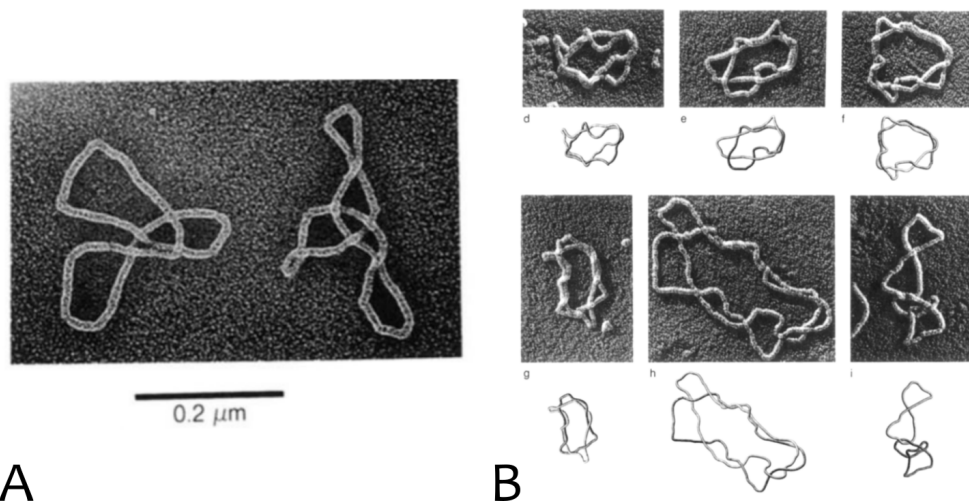


Figure 1.2: **A C** Electron microscopy of DNA (trefoil) knots made by topo II (type 2 topoisomerase) from bacteriophage T4. Image from Ref. [21]. **B** Electron microscopy of DNA catenanes accumulated in plasmid DNA after topo IV inhibition. Image from Ref. [10].

for instance, during the replication of circular DNA, examples of which are plasmids in bacteria and mitochondrial DNA (mtDNA) and extrachromosomal circular DNA (eccDNA) in eukaryotic cells. As shown in Fig. 1.3A, during replication, two identical copies of a piece of double stranded DNA are created by progressively separating its strands, and by bonding each of

them to a newly synthesised complementary strand. When this process is applied to circular DNA, it might happen, in some circumstances (e.g. if the original DNA is supercoiled), that the two produced replicas are linked to each other (Fig. 1.2B); their segregation, in this case, requires the intervention of specific enzymes. The malfunctioning of these enzymes, and the resulting difficulty in separating the components of replicated DNA catenanes, comes with severe consequences: topoisomerase IV (TopoIV) mutants in bacteria are conditionally lethal [10], TopoIII mutants are cause of human mitochondrial diseases [15], and it was recently speculated that TopoII, due to the lack of tension, might be unable to separate double minutes (i.e. fragments of eccDNA found in a large number of human tumours) catenanes during mitosis (eukaryotic cells reproduction) [16], and this might impact the heterogeneity and resistance to therapies of a wide range of tumours. Topoisomerases are not the only enzymes capable of resolving DNA entanglement, recombinases are also recognised to be capable of keeping DNA entanglement under control: it was observed, for instance, that Xer mediated site-specific recombination (Fig. 1.3B) can effectively separate catenated *Escherichia coli* plasmids formed during replication [17] *in vitro*.

DNA catenation is of fundamental importance not only to circular DNA, but also to linear DNA during mitosis, as it helps keeping together sister chromatids (identical copies of a chromosome) and must be resolved to allow their segregation [13]. Furthermore, contrary to what happens in the replication of circular DNA, DNA catenation seems to be advantageous to trypanosomes, protozoan parasites causing human sleeping sickness, as their mitochondrial DNA, for currently unclear reasons, is organised into massively interlocked networks of circular DNA (kinetoplast DNA), and knotting may be beneficial to viruses, as it regulates the ejection rate of viral genome [18].

Since entanglement may be both beneficial and detrimental, its regulation is essential to living organisms. How it is practically achieved, however, is currently not well understood. Motivated by this, by formulating analytic predictions and performing computer simulations, we model how entanglement is controlled in biological systems. Specifically, in chapter 3 we investigate the geometric and topological features of a system of polymer rings undergoing recombination under confinement and, in chapter 4, we show that entropic competition with slip-links, topological constraints physically found in both biological and synthetic polymer systems, provides a pathway to efficiently resolve entanglement. Since entanglement abundance determines the elastic response of polymer solutions [19], we speculate that our findings not only might impact our understanding of genome topology regulation, but also suggest potential strategies to control the rheological properties of synthetic

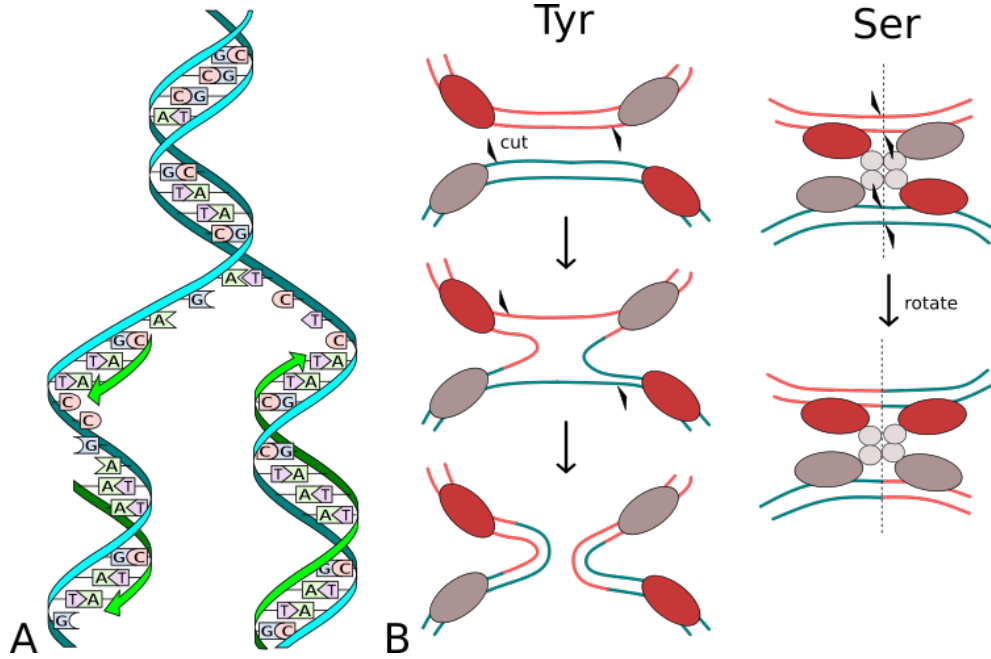


Figure 1.3: **A** Double stranded DNA (dsDNA) replication. A single stranded DNA monomer (nucleotide) is composed of a backbone element (sugar+phosphate) and a base, which can be of 4 different types, A, C, G and T. dsDNA is assembled by having the bases of a first strand bind to the bases of a second strand. Since A binds exclusively to T (and vice versa) and G to C, the two strands of dsDNA are complementary. During replication, the two strands of the original dsDNA are separated and paired to newly synthesised complementary strands. Image from wikipedia.org. **B** The two types of site-specific recombination: Tyrosine (Tyr) and Serine (Ser) site-specific recombination. The operation site on DNA is identified by looking at the local sequence of base types.

materials.

Not only the abundance of topological constraints, but also their complexity quantitatively determines the properties of entangled polymer systems. For example, it is known that ring polymers knotted with different topologies behave differently when translocating through small pores [20], and it is likely that topological complexity affects entanglement simplification in living organisms. For this reason, determining the statistical properties of polymers subject to fixed topological constraints is essential to understand entanglement regulation.

Comparing the paths described by a random walker to the spatial configu-

rations of a polymer has been, perhaps, the most successful way to describe the thermal fluctuations of polymer solutions. In particular, the connection between the self-avoiding walk and critical phenomena made it possible to formulate scaling predictions for the entropy, radius of gyration, looping probability, etc., of polymers in a good solvent (also called polymers in the dilute regime) [22], which, in turn, shed light on the structure and behaviour of polymers in thermal equilibrium.

The scaling behaviour of the entropy, for instance, not only helps characterise the geometric properties, such as the number of monomers involved in the knotted portion of a loop [23], of polymer knots and links, but also quantifies how likely it is for a system of ring polymers to be of a given link type. Whereas the critical behaviour of ring polymers with fixed knot type has been studied extensively [23, 24, 25], understanding how a fixed topological constraint affects the scaling behaviour and critical exponent of the entropy of a set of linked loops is an unexplored issue in the literature. This is investigated in chapter 5 of this work.

The random walk and self-avoiding walk models and their continuum limit are powerful tools which successfully describe the behaviour of polymers in a good solvent, but often fail to provide a satisfactory analytical description when topological and geometrical constraints are included [26]. For this reason, in this work both scaling analysis derived from the self-avoiding walk model and computer simulations are employed. In chapter 2, we briefly recall the connection between polymers and critical phenomena, focussing on the scaling behaviour of the entropy of polymers in the dilute regime – this is essential to understand the theoretical predictions of chapter 4 and the entirety of chapter 5 – and we detail the computational methods employed in the study reported in chapters 3, 4 (molecular dynamics) and 5 (Monte Carlo), as well as the theory on which they are built upon.

Chapter 2

Critical behaviour of polymers in a good solvent, theory and simulations

2.1 Scaling behaviour of topologically constrained polymers

2.1.1 The self-avoiding walk and critical phenomena

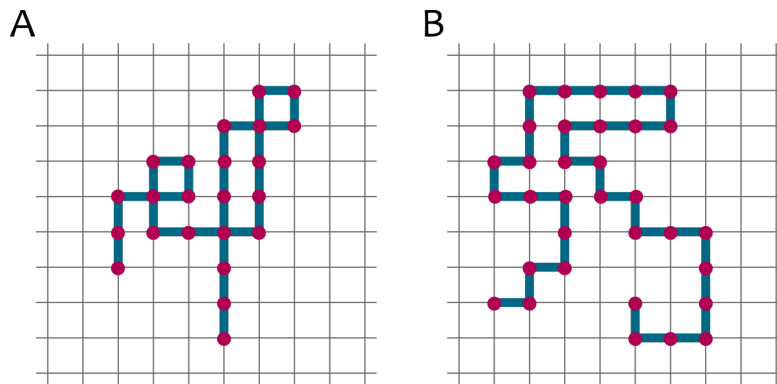


Figure 2.1: Realizations of **A** a random and **B** a self-avoiding walk on the square lattice. The path travelled by a self-avoiding walker never intersects itself.

Polymers have been studied theoretically for decades [27]. It is recognised that the configurations occupied by a polymer in a good solvent, for which

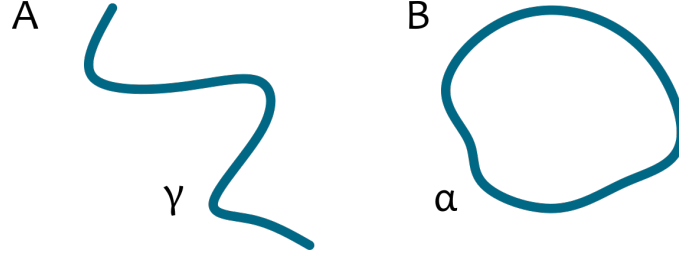


Figure 2.2: Sketches of **A** a self-avoiding walk and **B** a self-avoiding loop. The number of self-avoiding rings and chains, describing ring and linear polymers in a good solvent, correspond, respectively, to the magnetic susceptibility and specific heat of the n -vector model in the limit $n \rightarrow 0$.

interactions between polymer segments and solvent molecules are energetically favourable, are well described by the paths traversed by a self-avoiding walker (Fig. 2.1B) [22]. Whereas exact calculations are possible for an ideal chain (i.e. freely jointed chain or random walk 2.1A. See, e.g. [22]), formulating analytical and rigorous predictions based on the self-avoiding walk model is much more challenging. De Gennes first recognised the connection between polymers and critical phenomena by proving that the self-avoiding walk can be obtained as the limit for $n \rightarrow 0$ of $O(n)$, the n -vector spin model for ferromagnets [22]. He showed that the mean end-to-end length R , the number of self-avoiding walks of length (i.e. number of monomers) N , $\mathcal{Z}_N$ (Fig. 2.2A), and the number of self-avoiding loops of N monomers, $\mathcal{Z}_N(o)$ (Fig. 2.2B), are associated, respectively, to the correlation length, the magnetic susceptibility and the specific heat of the $O(n)$, $n \rightarrow 0$, model. In this correspondence, the temperature of the spin model is the conjugate variable of the length of the walk, thus, in the limit $N \rightarrow \infty$, scaling laws as

$$R^2 \propto N^{2\nu}, \quad (2.1)$$

$$\mathcal{Z}_N \propto \mu^N N^{\gamma-1}, \quad (2.2)$$

and

$$\mathcal{Z}_N(o) \propto \mu^N N^{-2+\alpha}, \quad (2.3)$$

are expected. Whereas α , ν and γ are universal critical exponents, μ , known as connectivity constant, is non-universal. The upper critical dimension is 4 [28]: starting from $d = 5$, a self-avoiding walk has so many ways to avoid itself that it is effectively described by an ideal chain. This consideration can be partially extended to $d = 4$, for which it is predicted that the value of the critical exponents of the self-avoiding walk and random walk (mean

field) coincide¹ (i.e. $\nu = \frac{1}{2}$ and $\gamma = 1$), provided the introduction of multiplicative logarithmic corrections [30, 31] in Eqs. 2.1, 2.2 and 2.3. Whereas the scaling limit of the self-avoiding walk can be derived rigorously only for $d \geq 5$ [28], field theory or, equivalently, direct renormalisation [32] of the Edwards model [33], allow to approximatively derive the value of the critical exponents, expressed as expansions in powers of $\epsilon = 4 - d$, in any dimension d . In our applications, we are mainly interested in the scaling of the number of configurations in $d = 3$, hence in the ϵ expansion of the exponents γ and α with $\epsilon = 1$. On the other hand, since α is related to ν by the hyperscaling relation $\alpha = 2 - d\nu$, the expansion of the exponent ν is equally interesting. To the third order in ϵ , ν and γ read [32]

$$\gamma = 1 + \frac{\epsilon}{8} + \frac{13}{4} \left(\frac{\epsilon}{8}\right)^2 + \left[\frac{97}{8} - 33\zeta(3)\right] \left(\frac{\epsilon}{8}\right)^3 + \mathcal{O}(\epsilon^4) \quad (2.4)$$

and

$$\nu = \frac{1}{2} \left\{ 1 + \frac{\epsilon}{8} + \frac{15}{4} \left(\frac{\epsilon}{8}\right)^2 + \left[\frac{135}{8} - 33\zeta(3)\right] \left(\frac{\epsilon}{8}\right)^3 + \mathcal{O}(\epsilon^4) \right\}. \quad (2.5)$$

For $\epsilon = 1$, to the second order,

$$\begin{aligned} \gamma &= 1.176 + \mathcal{O}(\epsilon^3), \\ \nu &= 0.592 + \mathcal{O}(\epsilon^3), \end{aligned} \quad (2.6)$$

values close to the numerical estimates [34, 35]

$$\begin{aligned} \gamma &= 1.157, \\ \nu &= 0.588. \end{aligned} \quad (2.7)$$

2.1.2 Polymer networks

Direct renormalisation can also be applied to predict the scaling of the number of configurations of a polymer network with fixed topology [36]. This is done by factorising the hyperscaling contribution² of the building blocks of the network, namely the loops and the vertices (Fig. 2.3). The contribution of each vertex of multiplicity (i.e. the number of connected branches) $\mathcal{M}$ is quantified by $\sigma_{\mathcal{M}}$, which is the exponent associated with the number of configurations of a star polymer with $\mathcal{M}$ branches, and the contribution of

¹The famous Flory formula [29], for example, predicts $\nu = \frac{3}{2+d} = \frac{1}{2}$ in $d = 4$

²i.e. how much they contribute to the scaling of the entropy of the network.

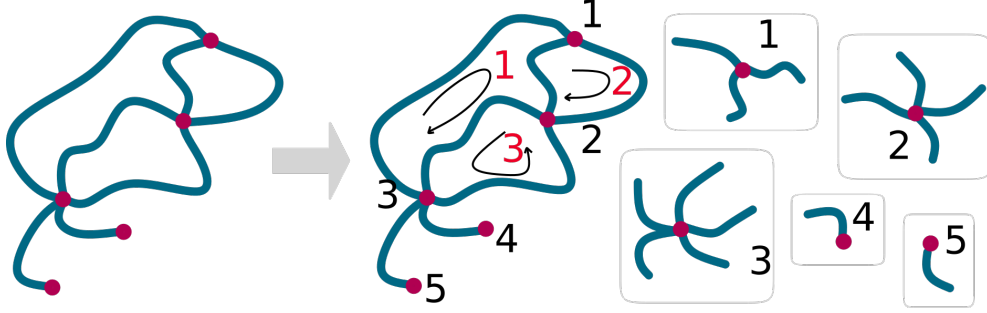


Figure 2.3: Sketch of a self-avoiding polymer network with $\mathcal{V} = 5$, $\mathcal{L} = 3$, $\mathcal{M}_1 = 3$, $\mathcal{M}_2 = 4$, $\mathcal{M}_3 = 5$, $\mathcal{M}_4 = 1$ and $\mathcal{M}_5 = 1$ and decomposition into star polymers.

each loop is gauged by $-\nu d$, which is the scaling exponent of ring polymers. Overall, the number of configurations of a network of topology $\mathcal{G}$, composed of $\mathcal{N}$ branches of length $N_i = N$, $i = 1, \dots, \mathcal{N}$, $\mathcal{L}$ loops and $\mathcal{V}$ vertices of multiplicity $\mathcal{M}_i$, $i = 1, \dots, \mathcal{V}$ scales as

$$\mathcal{Z}_N(\mathcal{G}) \propto \mu^{\mathcal{N}N} N^{\gamma_{\mathcal{G}}-1}, \quad (2.8)$$

where

$$\gamma_{\mathcal{G}} = -\nu d \mathcal{L} + \sum_{\mathcal{M} \geq 1} n_{\mathcal{M}} \sigma_{\mathcal{M}}, \quad (2.9)$$

where $n_{\mathcal{M}}$ is the number of vertices of multiplicity $\mathcal{M}$. The exponents $\sigma_{\mathcal{M}}$ can be expanded in powers of ϵ ; to the second order [36],

$$\sigma_{\mathcal{M}} = \frac{\epsilon}{8}(2 - \mathcal{M})\frac{\mathcal{M}}{2} + \left(\frac{\epsilon}{8}\right)^2 \frac{\mathcal{M}}{8}(\mathcal{M} - 2)(8\mathcal{M} - 21) + \mathcal{O}(\epsilon^3). \quad (2.10)$$

In our applications, the most relevant exponent is σ_4 , associated with vertices which can be formed, for instance, by bridging or slip-linking (see chapter 4) two polymer segments. According to Eq. (2.10), for $d = 3$,

$$\sigma_4 = -0.328125 + \mathcal{O}(\epsilon^3). \quad (2.11)$$

Given the polynomial nature in both $\mathcal{M}$ and $\epsilon = 1$ of the expansion in Eq. (2.10), a resummation would be needed to yield accurate results, however, to our knowledge, this has not yet been attempted. Anyway, the best numerical estimates of σ_4 [35, 37] are not far from the analytical prediction:

$$\sigma_4 = -0.44, -0.48, \quad (2.12)$$

Exponent γ can also be obtained from Eq. (2.10) by using the relation $\gamma = 2\sigma_1 + 1$ [36], in this way the first two orders of Eq. (2.4) would be recovered.

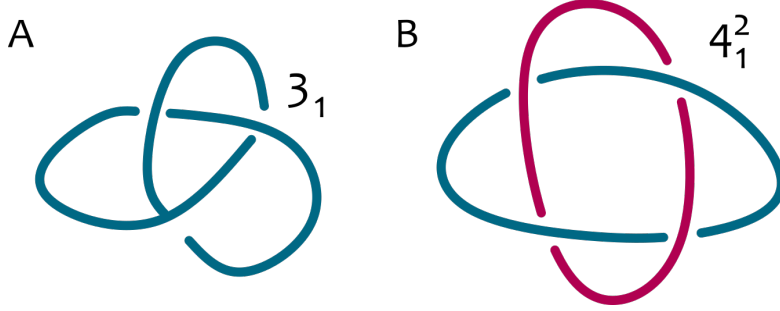


Figure 2.4: Example of topologically constrained ring polymers. **A** A trefoil (3_1 in the Alexander-Briggs notation) knot and **B** a Solomon (4_1^2) link.

Eqs. (2.8) and (2.9) were derived for monodisperse [36] polymer networks, i.e. networks satisfying $N_i = N$, $i = 1, \dots, \mathcal{N}$. For polydisperse networks, if we define $N = \sum_{i=1}^{\mathcal{N}} N_i$, it can be shown that [36]

$$\begin{aligned} \mathcal{Z}_N^{\text{poly}}(\mathcal{G}) &\propto \mu^N N^{\gamma_{\mathcal{G}}^{\text{poly}} - 1}, \\ \gamma_{\mathcal{G}}^{\text{poly}} - \mathcal{N} &= \gamma_{\mathcal{G}} - 1. \end{aligned} \quad (2.13)$$

2.1.3 Ring polymers with fixed topology

The mapping of the self-avoiding walk problem to field theory has been successful in predicting the asymptotic behaviour of polymer networks. The formulation of the statistical mechanics of polymers subject to a generic set of topological constraints (e.g. knots and links, see Fig. 2.4), on the other hand, is challenging. The general approach is writing down a theory for given fixed values of some topological invariants of interest; however, the analytical complexity of these invariants is a major practical obstacle to overcome. For instance, one of these complicated invariants is the linking number of two closed chains $\gamma_1 : t \in S^1 \rightarrow \gamma_1(t)$ and $\gamma_2 : s \in S^1 \rightarrow \gamma_2(s)$, which can be defined and computed from the link diagram of the chains as shown in Fig. 2.5. Whereas it is intuitively clear that the linking number counts the number of times the curves wraps around each other, its analytical formulation requires a double line integral:

$$\begin{aligned} Lk(\gamma_1, \gamma_2) &= \frac{1}{4\pi} \oint_{\gamma_1} \oint_{\gamma_2} \frac{\mathbf{r}_1 - \mathbf{r}_2}{|\mathbf{r}_1 - \mathbf{r}_2|^3} \cdot (d\mathbf{r}_1 \times d\mathbf{r}_2) \\ &= \frac{1}{4\pi} \int_{S^1 \times S^1} \frac{\det(\dot{\gamma}_1(s), \dot{\gamma}_2(t), \gamma_1(s) - \gamma_2(t))}{|\gamma_1(s) - \gamma_2(t)|^3} ds dt, \end{aligned} \quad (2.14)$$

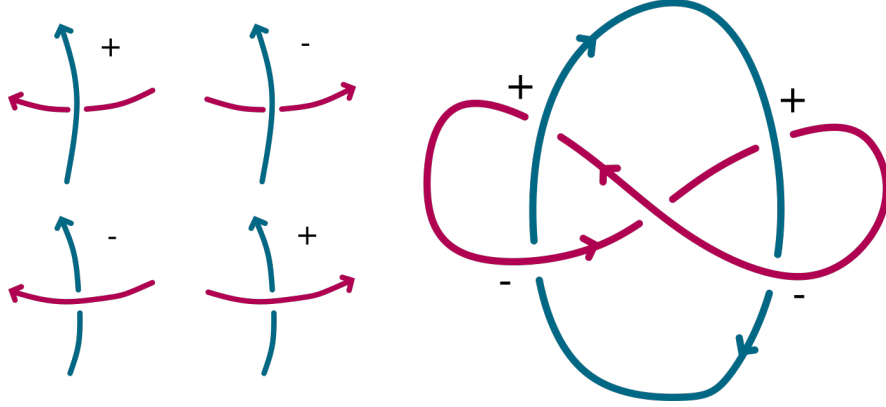


Figure 2.5: The total linking number of a link is given by the number of positive (+) crossings minus the number of negative (-) crossings of its link diagram (for a definition of link diagram, see, Fig. 2.9). Only crossings between different curves are considered and their sign is determined by arbitrarily assigning an orientation to each curve. The linking number is a topological invariant.

where $\dot{\gamma}_1 = \frac{d\gamma_1}{ds}$ and $\dot{\gamma}_2 = \frac{d\gamma_2}{dt}$. Eq. (2.14) is known as Gauss linking integral³. This expression computes the (signed) number of times the image ($\in$ the unit sphere) of the Gauss map $\Gamma(s, t) : (s, t) \in S^1 \times S^1 \rightarrow \frac{\gamma_1(s) - \gamma_2(t)}{|\gamma_1(s) - \gamma_2(t)|}$ covers the unit sphere⁴ (which area is 4π). This coincides with the linking number, since it also coincides with the signed number of times the Gauss map covers a given value v , which is also the signed number of crossings in the projection of the link (γ_1 and γ_2) to the plane perpendicular to v .⁵

Whilst strategies to write down a partition function for any fixed topological invariant do exist, the calculation of the expectation values of generic observables and the prediction of asymptotic scaling laws such as Eq. (2.8) remain challenging [26].

³Note that the second equality in Eq. (2.14) follows from the definition of the double line integral over γ_1 and γ_2 ($\oint_{\gamma} \mathbf{F}(\mathbf{r}) d\mathbf{r} = \int_{S^1} \mathbf{F}(\gamma(t)) \cdot \dot{\gamma}(t) dt$). $\det(\dot{\gamma}_1(s), \dot{\gamma}_2(t), \gamma_1(s) - \gamma_2(t))$ is the triple product $(\gamma_1(s) - \gamma_2(t)) \cdot (\dot{\gamma}_1(s) \times \dot{\gamma}_2(t))$.

⁴This number is also called degree of the Gauss map.

⁵It is interesting to point out that it is conjectured that Eq. (2.14) was first formulated by Gauss while he was studying Earth's magnetic field [38]. The connection between the geometry of two curves and electromagnetism comes to the surface, for instance, while taking the line integral along a closed curve γ_2 of the magnetic field generated by a constant current flowing through another closed path γ_1 : the Ampère-Maxwell equation infers that this integral is proportional to the linking number of γ_1 and γ_2 . Computing the magnetic field using the Biot-Savart law gives Eq. (2.14).

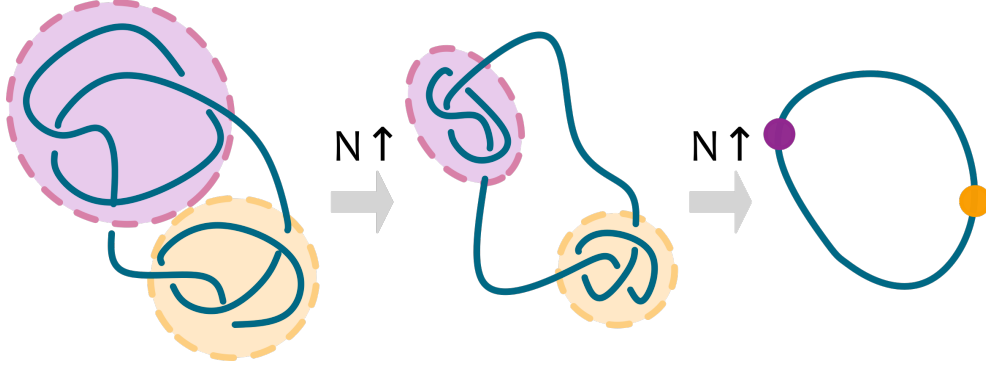


Figure 2.6: Sketch of how the monomers of a $3_1\#3_1$ knot (composition of two trefoil knots, $n_f = 2$) are distributed between the knotted and unknotted portions. The number of monomers involved in each knotted portion grows sublinearly with N , the total number of monomers.

A scaling law for $\mathcal{Z}_N(\tau)$, the number of configurations of ring polymers (i.e. knots) with fixed topology τ is hard to deduce from both Edwards theory, for the aforementioned difficulties, and the De Gennes $O(n)$, $n \rightarrow 0$ equivalence, as polymers with fixed topology do not have a magnetic counterpart. However, it can be rigorously proved [39] that $\mathcal{Z}_N(\tau)$ satisfies⁶

$$\liminf_{N \rightarrow \infty} \frac{\ln \mathcal{Z}_N(\tau)}{N} = k_\tau \leq \limsup_{N \rightarrow \infty} \frac{\ln \mathcal{Z}_N(\tau)}{N} = K_\tau \leq \lim_{N \rightarrow \infty} \frac{\ln \mathcal{Z}_N(o)}{N} = k, \quad (2.15)$$

where $\mathcal{Z}_N(o)$ is the number of configurations of ring polymers with arbitrary topology. $\mathcal{Z}_N(\tau)$ will then grow exponentially with N , and it is consequently assumed [23]

$$\mathcal{Z}_N(\tau) = A(\tau) N^{\alpha(\tau)-3} \mu(\tau)^N \left(1 + \frac{B(\tau)}{N^\Delta} + \dots \right). \quad (2.16)$$

Eq. (2.16) includes the leading scaling term and the first order correction, which functional form is set to be the same as that of the two leading terms of the scaling of $\mathcal{Z}_N(o)$ (i.e. Eq. (2.3) plus the first order correction-to-scaling) [40]. The difference from Eq. (2.3) is that, in Eq. (2.16), the critical exponents α and Δ , the amplitudes A and B and the connectivity constant

⁶Given a succession of real numbers (x_n) , the limit inferior and limit superior are defined, respectively as $\lim_{n \rightarrow \infty} \inf x_n = \lim_{n \rightarrow \infty} (\inf_{m \geq n} x_m)$ and $\lim_{n \rightarrow \infty} \sup x_n = \lim_{n \rightarrow \infty} (\sup_{m \geq n} x_m)$, where $\sup_{m \geq n} x_m$ and $\inf_{m \geq n} x_m$ are the largest and smallest values of $(x_m)_{m \geq n}$. The limit superior and limit inferior always exist (as opposed to the standard limit for, e.g., oscillating successions) and they are equal if and only if $\lim_{n \rightarrow \infty} x_n$ exists. In this case, their common value coincides with $\lim_{n \rightarrow \infty} x_n$.

μ are *a priori* topology dependent. Note that Eq. (2.15) implies $\mu(\tau) \leq \mu$ for every τ . Orlandini *et al.* [23] provided numerical evidence that, for $d = 3$, $\alpha(\tau) = \alpha(\emptyset) + n_f(\tau)$, where $\emptyset$ is the unknot and $n_f(\tau)$ is the number of prime factors of τ (see Figs. 2.6 and 5.1). This is explained by noting that the number of monomers contained in each physically knotted portion of a ring, N_k , grows sublinearly with N , the total number of monomers of the ring, $N_k \propto N^t$, $t \simeq 0.75$ [25]. As a knotted loop grows in size, its knotted portions involve a progressively smaller fraction of monomers, therefore, an infinitely long knot looks like an unknot crowned with $n_f(\tau)$ sliding roots [23] (Fig. 2.6).

2.2 Markov chain Monte Carlo

Analytical predictions of the scaling laws and critical exponents for topologically constrained polymers are difficult to formulate. A numerical approach to the problem is essential and, in many cases, to date, is the only one possible. Markov chain Monte Carlo (MCMC) algorithms have been successful in exploring the asymptotic behaviour of knotted loops, as they allow to sample from the configuration space according to a known distribution. For example, in the aforementioned work by Orlandini *et al.* [23] polygons (i.e. piece-wise straight knots) with fixed topology τ on the cubic lattice are sampled, and the exponent $\alpha(\tau)$ (see Eq. (2.16)) and the connective constant $\mu(\tau)$ are then estimated by extrapolation in the limit $\langle N \rangle \rightarrow \infty$.

2.2.1 Ergodic Markov chains

As detailed below, MCMC algorithms are built upon the properties of Markov processes [41, 42]. Given a discrete time Markov process X_t on a countable state space Ω , a probability distribution π on Ω is said to be stationary for X_t if it satisfies

$$\sum_{\omega \in \Omega} \pi(\omega) p_{\omega\omega'} = \pi(\omega') \quad \forall \omega' \in \Omega, \quad (2.17)$$

where $p_{\omega\omega'}$ is the transition probability from state ω to state ω' ⁷. A Markov chain X_t is irreducible if

$$\forall \omega, \omega' \in \Omega, \forall t \in \mathbb{N}, \exists n \in \mathbb{N} : Pr(X_{t+n} = \omega' | X_t = \omega) > 0; \quad (2.18)$$

⁷Here we assumed p to be time independent, namely X_t is homogeneous.

in simple words, X_t is irreducible if there is a positive probability of ever reaching any given state starting from any other. An irreducible Markov chain is aperiodic if

$$\forall \omega \in \Omega, \gcd \{n \geq 1 : \Pr(X_{t+n} = \omega | X_t = \omega) > 0\} = 1, \quad (2.19)$$

where $\gcd$ indicates the greatest common divisor and $\Pr(X_{t+n} = \omega | X_t = \omega)$ is the probability of returning to state ω after n steps. The asymptotics of aperiodic irreducible Markov chains on a finite space state Ω are well known:

Theorem 1 *For any aperiodic and irreducible Markov chain on a finite state space, there exist a unique stationary probability distribution.*

Theorem 2 *Any aperiodic and irreducible Markov chain converges, for $t \rightarrow \infty$, to its stationary distribution π , regardless of the starting initial distribution: $\lim_{t \rightarrow \infty} \Pr(X_t = \omega | X_0 = \omega') = \pi(\omega) \forall \omega, \omega' \in \Omega$*

Theorems 1 and 2 hold true for infinite but countable state spaces (this is actually the case we are interested in), provided that the Markov chain is positive recurrent⁸, which is guaranteed if there exists a stationary distribution. They assure that, if we run a Markov chain with the desired properties for long enough, the distribution will be close to its stationary distribution. In this case, it is often said that the Markov chain is ergodic and that it approaches equilibrium for long times.

2.2.2 Sampling

The strategy to build a MCMC algorithm to sample a space Ω is then: devising a way of sampling chains of stochastically generated states $\omega_t \in \Omega$ which are as uncorrelated as possible, proving that the underneath process X_t is irreducible and finding a stationary distribution, which uniqueness is guaranteed by Th. 1 (e.g. Fig. 2.7). It was proved that no ergodic local (i.e. performing only local transformation) and length-conserving Monte Carlo algorithm for sampling self-avoiding walks exists [43], hence local MCMC methods must be grand canonical to have a chance to be ergodic.

MCMC methods are often characterised by long correlation and equilibrium approach time. To improve their mobility, specific sampling techniques can be introduced, such as the multiple Markov chain method [44], which consists in running Markov chains $\{X_t^i\}$ in parallel and periodically swapping their realisations at time t , namely $X_t^i = \omega_t^i \longleftrightarrow \omega_t^j = X_t^j$.

⁸i.e. if the average time it takes to return to any given state $\omega \in \Omega$ is finite.

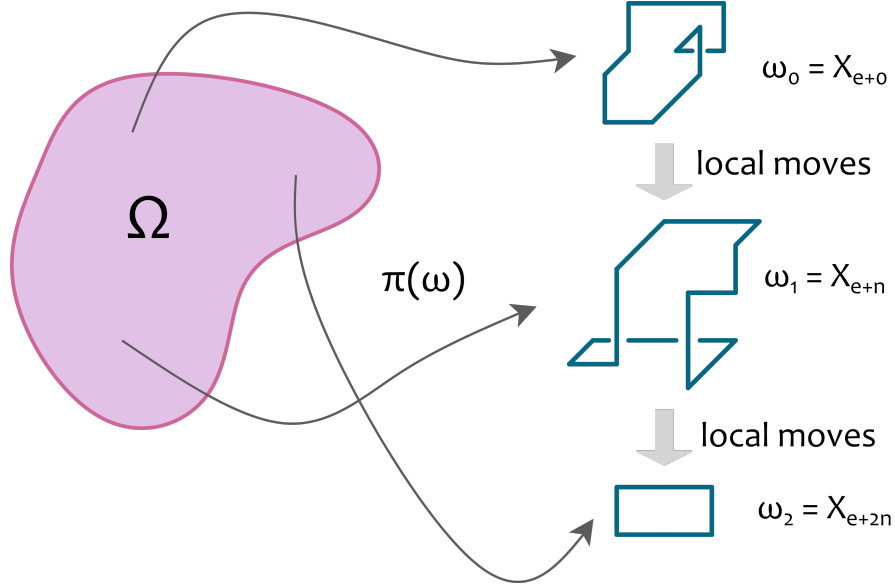


Figure 2.7: Scheme of the BFACF algorithm for sampling polygons on the cubic lattice. A sequence of configurations $\{X_i\}_{i \in \mathbb{N}}$ is generated by performing local transformations. After e equilibration steps, every n iterations, the configuration $\omega_a = X_{e+an}$, $a \in \mathbb{N}$ is sampled. At equilibrium, the sampled configurations are distributed according to the stationary distribution π .

2.2.3 BFACF algorithm for linked loops

A well known local and grand canonical MCMC method capable of ergodically [45] sampling polygons with fixed given topology on the cubic lattice is the BFACF (**B**erg-**F**oerster-**A**ragao-**C**aracciolo-**F**röhlich) [46, 47, 48] algorithm. In this work we extended this algorithm to sample sets of polygons linked with any given topology (see chapter 5), and proved that its ergodic classes are the link types. We now describe how the algorithm works and provide a sketch of the proof that it is ergodic.

Starting from a given initial configuration, a sequence of sets of linked polygons is generated by iteratively performing local modifications (Fig. 2.7) as detailed below. We will use the notation $\omega = (\omega_1; \omega_2; \dots; \omega_k) = (\omega_1(1), \omega_1(2), \dots, \omega_1(n_1); \omega_2(1), \omega_2(2), \dots, \omega_2(n_2), \dots, \omega_k(1), \omega_k(2), \dots, \omega_k(n_k))$, to indicate the configuration of a set of k polygons defined by the sequence of vertices $\omega_l(1), \omega_l(2), \dots, \omega_l(n_l)$, $l = 1, \dots, k$, and $[\omega_l(i), \omega_l(j)]$ to indicate the segment of the l -th polygon defined by the vertices $\omega_l(i), \dots, \omega_l(j)$ if $i < j$ or $\omega_l(i), \dots, \omega_l(n_l), \omega_l(1), \dots, \omega_l(j)$ if $i > j$. Note that periodic boundary conditions are required for polygons: $\omega_l(n_l + j) = \omega_l(j)$ and $\omega_l(1 - j) =$

$$\omega_l(n_l - (j - 1)) , j > 0.$$

1. Given a configuration

$$\omega = (\omega_1; \omega_2; \dots; \omega_k) = (\omega_1(1), \dots, \omega_1(n_1), \dots, \omega_k(1), \dots, \omega_k(n_k)),$$

one edge of the polygon set, together with one of its orthogonal vectors, $\vec{e}$, is randomly chosen. Without loss of generality, let us suppose that this edge, $[\omega_1(i), \omega_1(i + 1)]$, $i \in [1, n_1]$ belongs to the first, labelled by 1, polygon⁹.

2. Let us call $\vec{r}(\omega_1(i))$ and $\vec{r}(\omega_1(i + 1))$ the positions of the vertices $\omega_1(i)$ and $\omega_1(i + 1)$. If both $\vec{x}$ and $\vec{y}$ defined as $\vec{x} = \vec{r}(\omega_1(i)) + \vec{e}$ and $\vec{y} = \vec{r}(\omega_1(i + 1)) + \vec{e}$ are not already occupied (by vertices of any of the k polygons), a transformation of type $\Delta = +2$ along the direction $\vec{e}$ (Figure 2.8A), is attempted on $[\omega_1(i), \omega_1(i + 1)]$ with probability

$$p^\Delta(n) = \min \left(1, \left(\frac{n + \Delta}{n} \right)^{q - \frac{\Delta}{2}} F^\Delta \right), \quad (2.20)$$

where $n = \sum_l n_l$ is the total number of edges of the polygon set before the transformation and F and q are parameters. The new configuration would be $\omega' = (\omega_1(0), \dots, \omega_1(i), x, y, \omega_1(i + 1), \dots, \omega_1(n_1), \dots)$, with the new vertices x and y in the positions $\vec{x}$ and $\vec{y}$. Its number of edges is $n + 2$.

3. If $\vec{x} = \vec{r}(\omega_1(i - 1))$ and $\vec{y} = \vec{r}(\omega_1(i + 2))$, a move of type $\Delta = -2$ along the direction $\vec{e}$ (Figure 2.8A) is attempted on $[\omega_1(i), \omega_1(i + 1)]$ with probability defined by Eq. (2.20). The new configuration would be $\omega' = (\omega_1(0), \dots, \omega_1(i - 1), \omega(i + 2), \dots, \omega_1(n_1), \dots)$; its number of edges is $n - 2$.
4. If instead $\vec{x} = \vec{r}(\omega_1(i - 1))$ and $\vec{y}$ is not occupied, or if $\vec{x}$ is not occupied and $\vec{y} = \vec{r}(\omega_1(i + 2))$, then a flip move (Figure 2.8B) is performed on $[\omega_1(i - 1), \omega_1(i + 1)]$ or $[\omega_1(i), \omega_1(i + 2)]$ respectively. The new configurations would be

$$\omega' = (\omega_1(0), \dots, \omega_1(i - 1), y, \omega_1(i + 1), \dots, \omega_1(n_1), \dots)$$

or

$$\omega' = (\omega_1(0), \dots, \omega_1(i), x, \omega_1(i + 2), \dots, \omega_1(n_1), \dots)$$

respectively, with the vertices x and y in the positions $\vec{x}$ and $\vec{y}$.

⁹Otherwise, the procedure described below applies to the selected polygon.

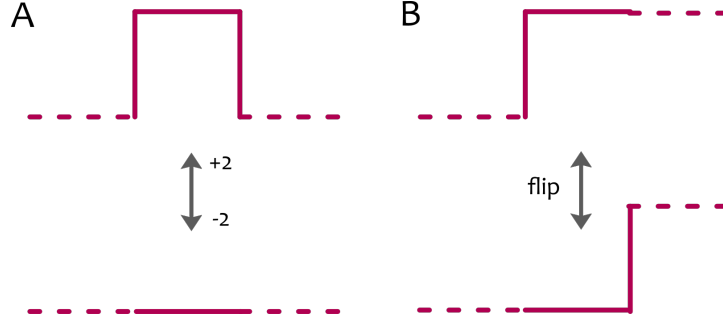


Figure 2.8: Local deformations performed by an iteration of the BFACF algorithm for linked polygons. **A** Moves of types ± 2 change the lengths of a polygon by $\Delta = \pm 2$. **B** Flips modify the configuration of a polygon without changing its length.

2.2.4 Ergodic classes of the BFACF algorithm for poly-disperse linked polygons

Ω , collection of sets of linked polygons, is an ergodic class for the BFACF algorithm if

$$\forall \omega, \omega' \in \Omega, \forall t \in \mathbb{N}, \exists n \in \mathbb{N} : Pr(X_{t+n} = \omega' | X_t = \omega) > 0, \quad (2.18)$$

where t and n are integers, and X_l is the configuration generated after l iterations of the algorithm (each iteration consists of moves 1-4 detailed above).

We now give a sketch of the proof that

Theorem 3 *Ergodic classes for the BFACF algorithm for linked polygons on the cubic lattice are the link types.*

Since, by construction, the transformations implemented by the BFACF algorithm described in the previous section preserve the link type, ergodic classes for the BFACF algorithm for linked polygons are made of configurations of the same link type.

Theorem 3 can hence be proved by showing that all possible sets of polygons linked with a given link type are part of the same ergodic class. This is achieved here by proving that two arbitrary set of polygons linked with the same link type can be mapped one into the other by performing BFACF moves (flips or ± 2 moves; see Fig. 2.8).

To this end, of crucial importance is the connection between links and their projections on a plane.

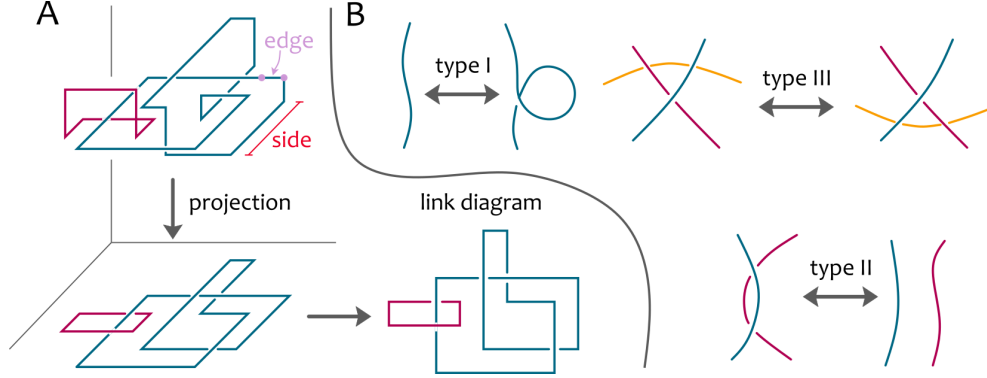


Figure 2.9: **A** Example of two linked polygons with lattice link diagram. Not all linked polygons admit a well defined lattice diagram. **B** Reidemeister moves.

Definition 1 (link diagram) *A link diagram is a projection of a set of linked loops decorated with underpasses and overpasses at each crossing point (see Fig. 2.9A).*

Definition 2 (equivalent link diagrams) *Two link diagrams are equivalent if there exist a finite sequence of transformations, consisting in either ambient isotopies on the plane or the Reidemeister moves (Fig. 2.9B), mapping one diagram into the other.*

Theorem 4 *Links are equivalent –i.e. of the same link type– if and only if their link diagrams are equivalent [49, 50].*

These concepts can be extended to polygons on a lattice. In a lattice link diagram (Fig. 2.9A), multiple points belonging to edges perpendicular to the projection plane are projected to the same point. Even discounting this, the projection of a set of linked polygons might not be a well defined diagram. This is the case, for instance, if more than one edge parallel to the projection plane is projected to the same edge. Theorem 4 applies to polygons on a lattice, but only to those provided with link diagram. By extending the construction detailed in Rensburg and Whittington [45] to a set of polygons, it can be showed that

- given a side (see Fig. 2.9) of a set of linked polygons on the cubic lattice and enough space, it is possible, by performing BFACF moves, to shift it one unit along a perpendicular direction and

- BFACF moves can be used to add an extra edge to each of the sides of a set of linked polygons intersecting a given plane perpendicular to one of the axis.

Using these tools, it can then be shown that

Lemma 1 *All sets of linked polygons can be modified, by performing BFACF moves, into sets of polygons provided with lattice link diagrams,*

and

Lemma 2 *BFACF moves can be used to map two arbitrary sets of linked polygons with the same lattice link diagram one into the other.*

Theorem 3 is therefore proved by adding one last ingredient:

Lemma 3 *BFACF moves can be used to map the link diagrams of two equivalent sets of linked polygons (i.e. with the same link type) one into the other.*

Given the correspondence between the equivalence classes of links provided with link diagrams and those of link diagrams (theorem 4), this is proved by showing that Reidemeister moves (on a diagram) can be performed by using a combination of BFACF moves on the cubic lattice and that two isotopic link diagrams can be mapped one into the other with BFACF moves. To this end, the procedure for lattice knots [45] can once again be extended to linked polygons, noting that

- by using the tools mentioned above, the space required to perform a Reidemeister move on a diagram can still be created, and that
- to map isotopic knot diagrams, a comparison of the areas delimited by the edges of each diagram is required. In the case of link diagrams, the perimeter of these areas can be the projection of edges belonging to different polygons; this, however, does not prevent the identification of corresponding areas of two isotopic diagrams.

2.2.5 Limit distribution

Theorem 3 assures that the BFACF algorithm does not change the link type, thus it is well suited to sample sets of linked polygons with fixed link type.

Furthermore, if the sampled configurations are uncorrelated (sampled at sufficiently large time –i.e. number of iterations– intervals), then it also assures that, if it exists, the sampling probability tends to a unique stationary distribution (see theorem 1). By proving it satisfies detailed balance¹⁰, it is straightforward to show that

$$\pi_{\tau_0}(\omega) = \frac{1}{\Phi} |\omega|^q F^{|\omega|} \chi(\tau(\omega), \tau_0), \quad (2.21)$$

where τ_0 is the link type of the starting configuration, Φ is a normalisation constant, $\chi(a, b)$ is 1 when a and b are the same link type and zero otherwise, and $|\omega|$ is the number of edges of ω , is the limit distribution. F , the step fugacity, controls the range of $|\omega|$ (see Ch. 5), and q can be chosen to optimize the sampling [51].

2.3 Molecular dynamics simulations

Monte Carlo algorithms allow the investigation of the statistical properties of a physical system by sampling the state space according to a known probability distribution. They do not, however, explicitly¹¹ follow the evolution of the system; in other words, they are typically unable to offer insight regarding the dynamics of a physical system [52].

2.3.1 NVE and velocity-Verlet

The purpose of Molecular dynamics (MD) simulations is numerically solving the equations of motion. In the classical microcanonical ensemble, often referred to as *NVE* in MD simulations, the total energy, E , the number of particles N and the volume V are conserved. The conservative forces acting on the particles are then described by a potential U and the equations of motion of the particles read

$$m_i \ddot{\mathbf{x}}_i = -\nabla_i U(\mathbf{x}_1, \dots, \mathbf{x}_N), \quad (2.22)$$

where m_i and $\mathbf{x}_i$ are the mass and position of particle i , $\nabla_i U(\mathbf{x}_1, \dots, \mathbf{x}_N)$ are the derivatives with respect to $\mathbf{x}_i$ of the potential. The solution of Eqs. (2.22), not only allows to follow the evolution of the system starting from an initial

¹⁰ $\pi(\omega)p_{\omega\omega'} = \pi(\omega')p_{\omega'\omega}$. This is a stronger condition than stationarity (Eq. (2.17)).

¹¹i.e. without simplifying assumptions.

configuration, but also to estimate ensemble averages, as stated by the ergodic hypothesis: if we assume that the trajectory of the system eventually passes through every point of the phase space, then the time average of any observable O asymptotically coincides with its ensemble average, namely

$$\lim_{T \rightarrow \infty} \int_{t_0}^T O(\mathbf{x}(t), \dot{\mathbf{x}}(t)) dt = \int_{\Omega} \pi(\omega) O(\omega) d\omega, \quad (2.23)$$

where Ω is the phase space and $\omega \in \Omega$.

Computing long trajectories is the aim of MD simulations; this makes the implementation of "good" numerical integrators for the equations of motion an interesting and complex problem [52, 53]. Geometric integrators, characterised by the remarkable property of preserving certain geometrical properties of the flow of a differential equation, are generally deemed superior. Symplectic integrators, in particular, preserve the structure of Hamilton's equations, and it can be shown that their predicted Hamiltonian flow is governed by a slightly perturbed Hamiltonian which tends to the unperturbed one when the integration timestep tends to 0 [54]. Although they possess the property of nearly conserving the total energy of conservative systems, which makes them ideal to simulate the NVE ensemble, the popularity of symplectic, and more generally geometric, integrators is not based on mathematical reasons, but on the reliability evinced from numerical tests and applications [52]. Among the geometric integrators, widely used and well studied is the Verlet algorithm [55], which is symplectic and time-reversible; its velocity variant (velocity-Verlet) reads:

$$\begin{aligned} \mathbf{v}_i \left(t + \frac{1}{2} \Delta t \right) &= \mathbf{v}_i(t) + \frac{1}{2m_i} \mathbf{f}_i(t) \Delta t, \\ \mathbf{x}_i(t + \Delta t) &= \mathbf{x}_i(t) + \mathbf{v}_i \left(t + \frac{1}{2} \Delta t \right) \Delta t, \\ \mathbf{v}_i(t + \Delta t) &= \mathbf{v}_i \left(t + \frac{1}{2} \Delta t \right) + \frac{1}{2m_i} \mathbf{f}_i(t + \Delta t) \Delta t, \end{aligned} \quad (2.24)$$

where $\mathbf{x}_i(t)$ and $\mathbf{v}_i(t)$ are, respectively, the position and the velocity of particle i at time t , $\mathbf{f}_i(t)$ is the force experienced by particle i at time t and Δt is the integration timestep.

2.3.2 NVT

Since fixing the temperature of an experimental system is usually easier than controlling its total energy, simulating the dynamics or computing equilibrium averages of a system in the canonical ensemble (NVT) is often the target

of MD simulations. The statistics of the NVT ensemble can be simulated by thermostating the NVE ensemble, namely by introducing a device to fix the mean value of the temperature, which is associated to the mean velocity of the simulated particles through the equipartition theorem. A commonly employed thermostat is the Langevin heat bath. The (mesoscopic) particles of the system are immersed in a bath of microscopic particles, which is implicitly accounted for by modifying the deterministic equations of motion (2.22). Each particle experiences a dissipative frictional drag and a stochastic noise due to colliding with the particles of the bath; the motion of particle i is thus described by the famous Langevin equation

$$m_i \frac{d^2 \mathbf{x}_i}{dt^2} = -\nabla_i U(\mathbf{x}_1, \dots, \mathbf{x}_N) - \xi m_i \mathbf{v}_i + \sqrt{2k_B T \xi m_i} \mathbf{h}(t), \quad (2.25)$$

where ξ is the experienced friction, k_B is the Boltzmann constant, T is the temperature and $\mathbf{h}$ is stochastic noise satisfying

$$\begin{aligned} \langle \mathbf{h}(t) \rangle &= \mathbf{0}, \\ \langle h_i(t) \cdot h_j(t') \rangle &= \delta_{ij} \delta(t - t'). \end{aligned} \quad (2.26)$$

A simple way to numerically integrate Eq. (2.25) is by introducing in Eq. (2.24) a modified force defined as follows [56]:

$$\mathbf{f}_i^{tot}(t) = -\nabla_i U(\mathbf{x}_1, \dots, \mathbf{x}_N) - \xi m_i \dot{\mathbf{x}}_i(t) + \sqrt{\frac{2k_B T \xi m_i}{dt}} \hat{\mathbf{h}}_i(t), \quad (2.27)$$

where each component of $\hat{\mathbf{h}}(t)$ is a Gaussian random variable with mean 0 and variance 1^{12} . The resulting implementation of the velocity-Verlet algorithm, accounting for the thermal bath, is

$$\begin{aligned} \mathbf{v}_i \left(t + \frac{1}{2} \Delta t \right) &= \mathbf{v}_i(t) - \frac{\Delta t}{2} \left(\frac{\nabla_i U(t)}{m_i} + \xi \mathbf{v}_i(t) \right) + \sqrt{\frac{\Delta t k_B T \xi}{m_i}} \hat{\mathbf{h}}_i, \\ \mathbf{x}_i(t + \Delta t) &= \mathbf{x}_i(t) + \mathbf{v}_i \left(t + \frac{1}{2} \Delta t \right) \Delta t, \\ \mathbf{v}_i(t + \Delta t) &= \mathbf{v}_i \left(t + \frac{\Delta t}{2} \right) - \frac{\Delta t}{2} \left(\frac{\nabla_i U(t + \Delta t)}{m_i} + \xi \mathbf{v}_i \left(t + \frac{\Delta t}{2} \right) \right) + \sqrt{\frac{\Delta t k_B T \xi}{m_i}} \hat{\mathbf{h}}_i, \end{aligned} \quad (2.28)$$

where $\nabla_i U(t)$ and $\nabla_i U(t + \Delta t)$ are the derivatives of the potential computed by using the positions of the particles at time t and $t + \Delta t$ and $\hat{\mathbf{h}}_i$ is a random vector, each component of which is, as described above, normally distributed.

¹² $\frac{\hat{\mathbf{h}}}{\sqrt{dt}} = \mathbf{h}$ is derived by imposing that the solution of Eq. (2.25), which is a stochastic differential equation, describes Brownian motion.

2.3.3 Optimisation

Integrating the equations of motion for many timesteps in a reasonable time, especially for systems of many particles, requires enormous computational power. For this reason, many optimisation techniques have been devised [57]. Most of the time of a MD simulation is actually spent on computing the potential in Eqs. (2.22) and (2.25), whereas the integration time (e.g. using Eq. (2.24)) generally takes only 2 – 3% of the total workload [57]. This is because the potential is a sum of many terms, each describing a many-body interaction between a group of atoms: even considering only two-bodies interactions between N particles, there would be N^2 terms to compute at each timestep. Building neighbour lists is a popular method, first proposed by Verlet [55], to greatly lighten the burden of computing the forces. If all the interactions are short-ranged, then there is no need to consider every pair or group of atoms. As the simulation proceeds, for each particle, a dynamic list of atoms within a certain cut-off range is compiled and updated; only interactions between neighbours are then evaluated. Compiling neighbour lists can be done efficiently by first binning the simulated domain and then looking for neighbours only within nearby bins.

Another way to significantly improve the efficiency of MD algorithms is parallelisation, namely assigning tasks, carried out simultaneously, to multiple processors (CPUs). Parallelisation of the workload can be achieved through spatial-decomposition: each CPU is assigned to a specific portion of the simulation domain, and computes the forces acting on the atoms inside its sub-domain exclusively. To do its job, a processor will generally require the acquisition of the position and velocity of atoms belonging to other (adjacent for short-range interactions) sub-domains, hence a communication algorithm is to be implemented as well.

2.3.4 MD simulations of coarse-grained polymers

Optimisation is certainly required, but it is often not sufficient. Since the number of atoms composing macroscopic and mesoscopic systems is enormous, the space and time scales which can be explored by performing all-atoms molecular dynamics simulations are severely restricted. These shackles can be broken by coarse-graining the system under investigation, namely by describing it with a greatly reduced number of effective degrees of freedom, obtained by integrating over the microscopic states of the system. This technique is reminiscent of the theoretical treatment of critical phenomena. Perhaps the most crucial prescription of coarse-graining modelling, the in-

interactions governing the motion of the effective degrees of freedom are to be defined in such a way they reproduce the correct behaviour of the studied system.

While the random and self-avoiding walks are popular coarse-grained models for the analytical and Monte Carlo description of polymers in a good solvent, a chain of beads connected by springs is the most suited model for performing molecular dynamics simulations. If $\mathbf{x}_i$ in the Langevin equations (2.25) are the positions of the beads, each modelling an entire segment of a real polymer, of a bead-spring chain, then their time evolution describes the thermic fluctuation of a polymer in contact with a heat bath, provided the properties of the simulated polymer are accurately captured by the potential U . For instance, the molecular structure, elasticity [58] and finite-extensibility of polymers are accounted for by the springs connecting the beads, which are introduced in the model via the pairwise FENE (finitely extendible nonlinear elastic) potential [59]

$$U_{\text{FENE}}(r) = \begin{cases} -0.5kR_0^2 \ln(1 - (r/R_0)^2) & r \leq R_0 \\ \infty & r > R_0 \end{cases}, \quad (2.29)$$

where r is the distance between two bonded beads, $R_0 = 1.5\sigma$ is the maximum extension of the bond, $k = 30\epsilon/\sigma^2$ [60, 61], ϵ is the energy unit and σ is the bead size. Excluded volume interaction and stiffness are also typically accounted for in MD simulations of coarse-grained polymers. The former are described by the shifted and truncated, so to keep the repulsive part exclusively, Lennard-Jones potential [62], the Weeks-Chandler-Andersen potential

$$U_{\text{WCA}}(r) = \begin{cases} 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 + \frac{1}{4} \right] & r \leq r_c \\ 0 & r > r_c \end{cases}, \quad (2.30)$$

where r is the distance between two interacting beads and $r_c = 2^{1/6}\sigma$. The latter is included via a Kratky-Porod potential [63, 64] acting on triplets of neighbouring beads:

$$U_{\text{bend}} = \frac{\epsilon l_p}{\sigma} (1 + \cos \theta), \quad (2.31)$$

where θ is the angle between consecutive bonds and l_p is the persistence length¹³.

In this work we perform molecular dynamics simulations of bead-spring polymers coupled to an implicit heat bath as described above. We integrate the Langevin equations (2.25) by using the velocity-Verlet scheme (Eq. (2.28)). This is carried out by the LAMMPS engine [65], which implements the construction of neighbour lists and the spatial-decomposition of the workload.

¹³The higher it is, the stiffer the polymer.

Chapter 3

Topological gelation of confined recombinant ring

Recombination is a ubiquitous process in biology: it involves the transient disruption of two DNA segments that are spatially proximate in 3D – although not necessarily adjacent in 1D – followed by a biochemical process which rejoins the DNA ends in a different way (Fig. 3.1A; the latter process is typically referred to as alternative end-joining). This reaction is often regulated by proteins when it occurs between specific DNA sites (see Fig. 1.3B), however it can also happen spontaneously [66]. Recombination underlies the formation of gametes in eukaryotes and ensures genetic diversity in prokaryotes [67]. Additionally, DNA recombination has been speculated to regulate gene expression [68], is known to aid resolution of complex DNA topology in cell division [67], and is required during immunological response [69].

Since the outcome of random recombination operations may yield linked or knotted DNA products, which, in turn, would pose a topological problem to the cell, recombination is a tightly regulated process that yields precisely controlled topologies [70]. Enzyme-mediated recombination has been well studied in the dilute regime by using short plasmids with specific recombination sites *in vitro* [17] and tangle theory [71, 69]. However, it is far less understood how recombination is regulated, and how certain target topologies are achieved, for long chromosomes within the entangled and crowded intracellular milieu. Thus, understanding the biophysical principles underpinning the topological regulation of the genetic material of living cells under recombination remains an outstanding open question in DNA biophysics and genome biology.

To shed light on this issue, here we study a system of ring polymers under-

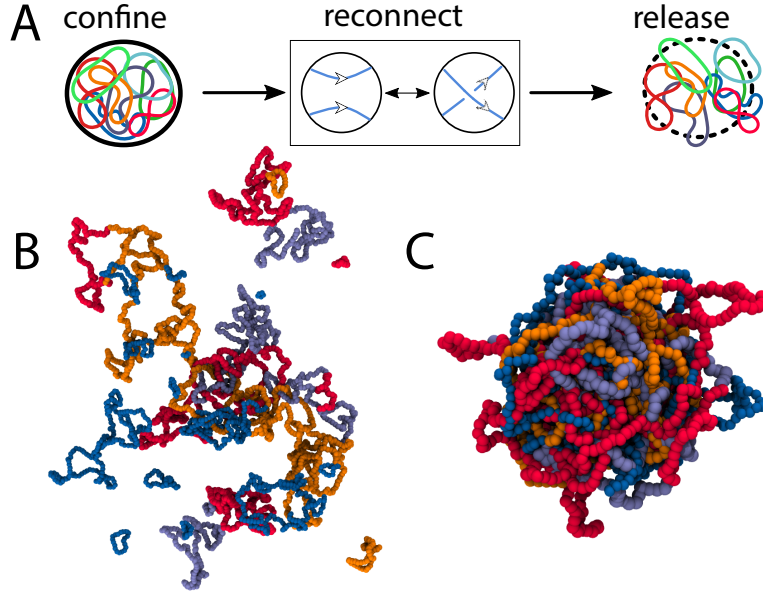
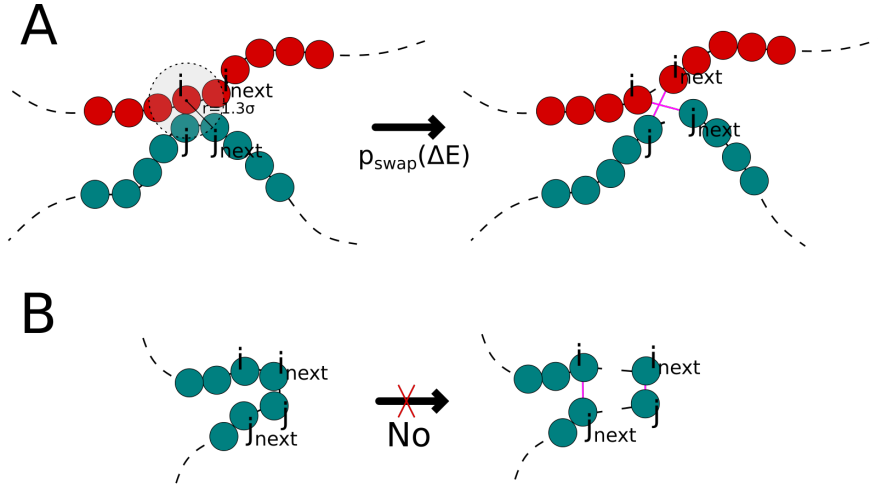


Figure 3.1: **A** We study ring polymers confined within a sphere of radius R and undergoing recombination events anywhere along their contour. **B-C** The two panels show two possible states of the system at equilibrium and after relaxing the confinement. **B** depicts an ensemble of many, small unlinked rings, whereas **C** shows an ensemble of few, long and linked rings. The confinement release step is done here for ease of visualisation.

going recombination reactions (hereafter called “recombinant rings”) under spherical confinement 3.1. We discover that depending mainly on DNA stiffness and radius of the confining sphere, this system harbours a topological transition between a regime with many short, unlinked and segregated rings and another one with few long, mixed and linked rings. The transition that we find can be characterised in both geometrical and topological terms. Geometrically, it is naturally explained as the result of the competition between the bending energy of the loops and the total entropy of the system. Topologically, it can be seen as a gelation transition and understood in terms of the critical overlap concentration c^* , above which linking is expected to be thermodynamically favoured. We discuss our results in the context of DNA recombination *in vivo* and suggest potential strategies to exploit this gelation transition to design DNA rings with desired length and topological features *in vitro*. We note our system can be viewed, more generally, as a confined melt of living polymer loops, which can form spontaneously following polymerisation of monomers given an appropriate choice of the reaction



kinetics [72].

3.1 Model

The system is initialised as one bead-spring ring polymer (see chapter 2) with $N = 1000$ beads of size σ inside a large sphere that is slowly compressed to $R = 7\sigma$, allowing the confined polymer to equilibrate. Confinement is achieved by adding a restoring radial force F acting on the beads of the polymer,

$$F(r) = \begin{cases} -30\epsilon (r - R)^2 & r \geq R \\ 0 & r < R \end{cases}, \quad (3.1)$$

where r is the distance from the centre of the sphere and ϵ the energy unit. Consecutive beads are strung together by FENE bonds (Eq. (2.29)), and excluded volume interactions are modelled via the WCA potential (Eq. (2.30)). The stiffness of the chains is a key parameter of our system, and is accounted for by imposing a bending penalty (the Kratky-Porod potential, Eq. (2.31))

acting on triplets of neighbouring beads. It is quantified by the parameter K , proportional to the persistence length l_p of the polymers ($K = \frac{\epsilon l_p}{\sigma}$, see Eq. (2.31)). After the initial equilibration, we allow the rings to undergo recombination – i.e., transient chain breakage followed by alternative end-joining (or bond reconnection), Figs. 3.1A and 3.2 – between any two segments that are proximal in 3D (effectively we consider only segments closer than $r_c = 1.3\sigma$, see Fig. 3.2). Recombination moves that would change the configuration of the system from ω to ω' are accepted or rejected according to a standard Metropolis test based on $\Delta E = E(\omega) - E(\omega')$, the energy difference after and before a trial recombination event. The acceptance probability reads

$$p_{\text{swap}}(\omega'|\omega) = \begin{cases} \exp(-\Delta E/\epsilon) & \Delta E \geq 0 \\ 1 & \Delta E < 0 \end{cases} . \quad (3.2)$$

If the distance condition is satisfied, and the Metropolis test successful, recombination moves can occur at every integration timestep ($\Delta t = 0.001\tau_b$, with $\tau_b = \sigma^2\gamma/k_B T$ the Brownian time). In practice, recombination rates depend strongly on the stiffness parameter K : for instance, the rate ranges between 0.235 per Brownian time τ_b at $K = 0$, to 0.004 per τ_b at $K = 5$. Given Brownian times for a fibre of size 30 nm as chromatin, in a medium with viscosity ~ 100 cP as the nucleoplasm, these rates can be converted to $0.5 - 30 \text{ s}^{-1}$. This is faster than the 1/min rates in biological systems [73], but of the same order as recombination events in wormlike micelles [74, 75]. The precise value of this rate does not modify our conclusions as we are interested in the long-time, steady state behaviour of the system, rather than the transient dynamics.

Every $10^3\tau_b = 10^5\Delta t$ we take a snapshot of the system and record the number of rings at time t , $N_r(t)$, as well as their length $L_r(n, t)$. The total number of beads in the system is kept constant and given by $N = 1000 = \sum_{n=0}^{N_r(t)} L_r(n, t)$ so that the monomer density is $\rho = 3N/(4\pi R^3) \simeq 0.7\sigma^{-3}$ – equivalently, the volume fraction is $\phi = 0.364$.

3.2 Geometric properties of confined recombining rings

We begin by characterising the geometric properties of the system as a function of K , for a fixed value of R . Specifically, we monitor the number of rings and their lengths: in our simulations, both these quantities evolve at short time to reach a well-defined steady state value (see Fig. 3.3A,B).

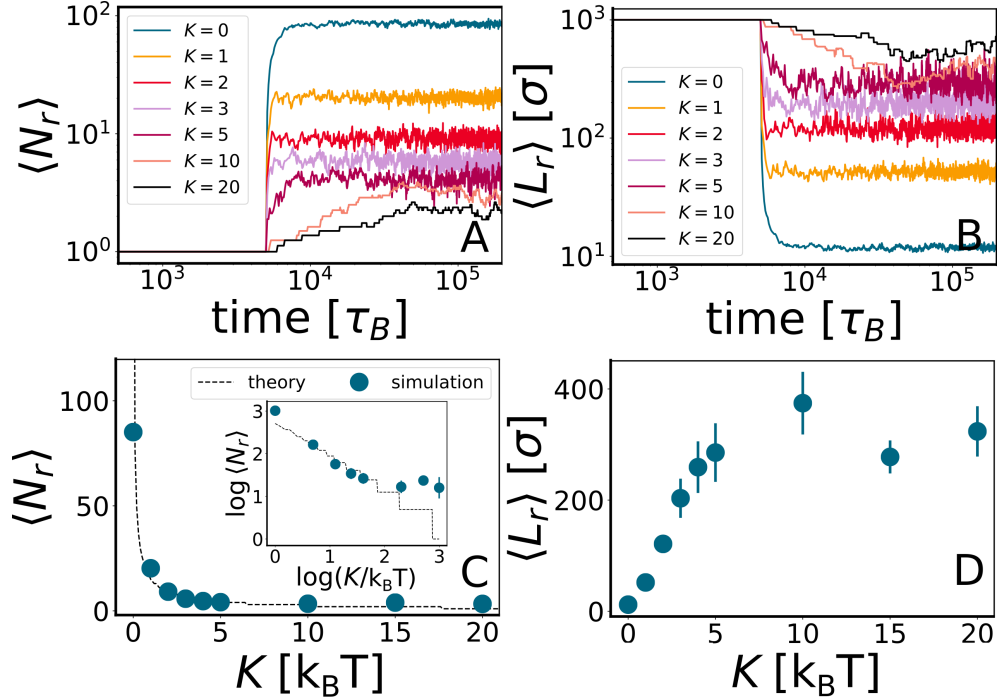


Figure 3.3: Time dependence of the number of recombinant rings N_r **A** and their average contour length L_r **B**. Different curves refer to different values of the stiffness parameter K (proportional to the persistence length L_p). All the data refer to a spherical confinement of radius $R = 7\sigma$ i.e. to a volume fraction $\phi = 0.364$ and the initial condition $N_r(0) = 1, L_r(0) = 1000$. In panels **C** and **D** we report the average steady-state values of the observables of panels **A** and **B** respectively as a function of K . The black line in panel **C** is predicted by minimizing the free energy (3.9) with $\lambda = 5.3$.

Our data are suggestive of a smooth transition or crossover between a regime in which many short rings populate the sphere (at low K) and another one in which few long rings remain in steady state (at high K). We temporarily refer to these as the short ring and the long ring regimes respectively. A typical snapshot of the system in the two regimes is shown in Fig. 3.1B (where the confining sphere is removed for ease of visualisation). This crossover can be understood in terms of the competition between the bending energy of the loops – regulated by K – and their combinatorial and translational entropy.

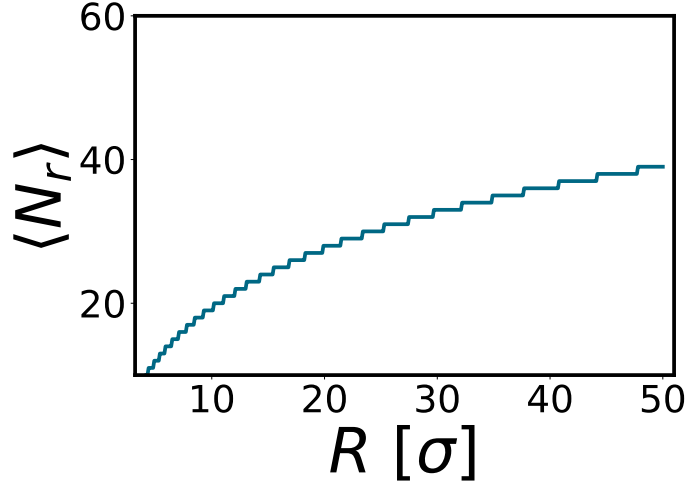


Figure 3.4: Average number of rings N_r , as a function of the confinement radius R , predicted from Eq. (3.9) with $\lambda = 5.3$ and $K = 1$.

3.2.1 Free energy of recombinant rings at equilibrium

For simplicity, let us suppose that in the long time limit there are m rings of equal contour length $l_i = l = N\sigma/m = L/m$ where N is the total number of monomers in the system and σ is the step length (i.e. the distance between two consecutive monomers). If l_p is the persistence length, the corresponding bending energy is

$$U = k_B T \sum_{i=1}^m \frac{2l_p \lambda \pi^2}{l_i} \simeq m k_B T \frac{2l_p \lambda \pi^2 m}{L}, \quad (3.3)$$

while an estimate of the entropy is given by $S = k_B \ln \mathcal{Z}$, where $\mathcal{Z}$ is the number of configurations of the system. We point out that, in Eq. (3.3), we assumed the bending energy of each ring of the system to be $k_B T \frac{2l_p \lambda \pi^2 m^2}{l_i}$. This is an approximation of the bending energy of ring polymers of any contour length l_i , which was set such that its functional form coincides with the bending energy of short rings (with length comparable to l_p). λ is a fitting parameter. To estimate $\mathcal{Z}$ we suppose that the rings are independent. This gives

$$\mathcal{Z} = \sum_{l_1, \dots, l_m} f(l_1, \dots, l_m) \prod_{i=1}^m \mathcal{Z}^1(l_i), \quad (3.4)$$

where the sum is over all the possible values of the m lengths l_i (which are multiples of σ compatible with $\sum_{i=1}^m l_i = L = N\sigma$ and $l_i \geq 3\sigma \forall i$), f accounts for all the possible ways N monomers are distributed between the m rings

(i.e. reconnection), and $\mathcal{Z}^1(l_i)$ is the number of configurations for a ring with length l_i . This number is not known exactly but if we restrict ourselves to the case of freely-jointed linear chains with Kuhn length $2l_p$ and confined in a space of volume $V = R^3$ we have [76]

$$\mathcal{Z}^1(l_i) = \exp \left(-\frac{l_p \pi^2}{V^{2/3}} l_i + \ln(V) + \ln(2^9/\pi^6) \right). \quad (3.5)$$

From (3.5) we then have

$$\prod_{i=1}^m \mathcal{Z}^1(l_i) = \exp \left(-\frac{l_p \pi^2}{V^{2/3}} L + m \ln(V) + m \ln(2^9/\pi^6) \right). \quad (3.6)$$

On the other hand,

$$\sum_{l_i, \dots, l_m} f(l_i, \dots, l_m) \approx \{N, m\} \left[\frac{1}{2} \left(\frac{N}{m} \right)! \right]^m, \quad (3.7)$$

where $\{N, m\}$, is the Stirling number of the second kind and counts all the possible ways to partition N monomers into m groups while $\left[\frac{1}{2} \left(\frac{N}{m} \right)! \right]$ is the number of permutations of the monomers of a linear chain of length $l = \frac{N\sigma}{m}$. This gives

$$\frac{S}{k_B} = \ln(\{N, m\}) + m \ln \left[\left(\frac{N}{m} \right)! \right] - \frac{l_p \pi^2}{V^{2/3}} L + m \ln \left(\frac{V 2^8}{\pi^6} \right). \quad (3.8)$$

The free energy of the system $F = U - TS$ is then given by

$$\frac{F}{k_B T} = \frac{2l_p \lambda \pi^2 m^2}{L} + m \ln(m) - m - m \ln \left(\frac{V 2^8}{\pi^6} \right) + \text{const}. \quad (3.9)$$

By minimising the free energy as a function of m , we can find the average number of rings as a function of K and R . As can be seen in Fig. 3.3C, the theory captures the numerical results accurately for $l_p \lesssim R$. This is not the case when the persistence length of the system is comparable to or greater than the confinement radius. It is likely because the theory fails to include the effect of the geometry of stiff polymers under strong confinement [77]. Since the total number of monomers is conserved, Eq. (3.9) also describes the behaviour of the expected mean length $\langle L_r \rangle$ (Fig. 3.3D). Our theory shows that increasing K leads to a transition (or crossover) between the short ring and the long ring regimes and predicts that a similar transition can be observed by decreasing R for fixed K (Fig. 3.4).

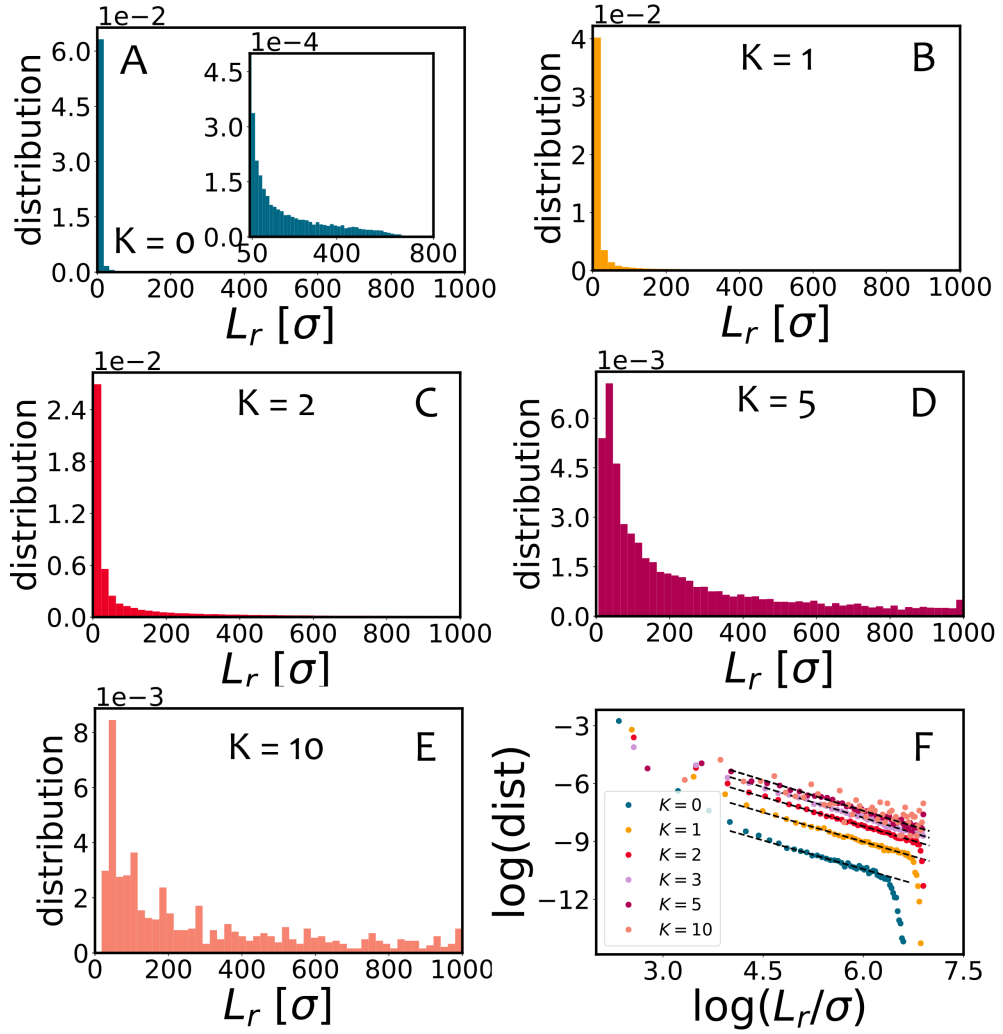


Figure 3.5: **(A-E)** Distribution of length of recombinant rings for $K = 0$, $K = 1$, $K = 2$, $K = 5$ and $K = 10$, and $N = 1000$. **(F)** Log-log plot of the length distribution for $K = 0$, $K = 1$, $K = 2$, $K = 3$, $K = 5$ and $K = 10$. The fitting curves are lines with slope -1 .

An intriguing feature of our recombinant polymer system is that the rings obtained in steady state have very broad size distributions (see Fig. 3.5). Notably, this is true for both small and large K , and the size distribution is a power law, $P(L_r) \sim 1/L_r$, for all cases. As recombination changes polymer size but does not violate detailed balance in our model, the system is effectively in thermodynamic equilibrium, and the size distribution should be linked to the Boltzmann weight of rings of different sizes. Neglecting the dependence on K , which is expected to be a fair approximation for sufficiently

long rings, we expect the size distribution to be $\sim L_r^{-c}$, i.e., the probability of forming a loop of size L_r . We note that a power-law size distribution of rings is also found, for analogous reasons, in living polymers in the phase where loops are favoured over linear chains [72]. Deviations from the power law $P(L_r) \sim 1/L_r$ (the lines in Fig. 3.5F) are arguably due to finite size corrections for small loops and to the absence of a single ring state for low values of K .¹

3.2.2 Radial positioning and mixing of recombinant rings

Inside the cell, chromosomes segregate into territories, and position themselves non-randomly with respect to the nuclear lamina. For instance, in lymphocytes gene-rich, or active, chromosomes tend to be positioned internally, while gene poor, or inactive/heterochromatic, chromosomes tend to be closer to the lamina [78, 79]. While the interaction between heterochromatin and lamina is likely key to achieve the correct positioning *in vivo* [80, 81], entropic effects are also important [79]. In our synthetic systems, we find that entropy and stiffness are sufficient to induce a non-random positioning of our recombinant rings (toy chromosomes) inside the confining sphere (a toy nucleus). Additionally, ring stiffness also controls mixing, so that the two regimes found for different values of K also differ in terms of interloop mixing.

To characterise the radial positioning of our loops, we compute the normalised radial density of monomers, $G(r)$, for each of these. Uniform positioning of loops within the sphere corresponds to a constant $G(r)$. Instead, we see that smaller loops (small L_r , dark blue region in Fig. 3.6B, Fig. 3.6C,D) are depleted in the interior of the sphere. This is because smaller soft spheres can approach the surface more frequently with respect to larger ones. There is also a layering of loop monomers (yellow stripe in Fig. 3.6B, peaks in Fig. 3.6C), which is reminiscent of the density profile at an interface between a solid wall and a concentrated colloidal fluid [82].

To study mixing, we construct a parameter, similar to the conditional entropy used to measure the entropy of mixing [83], by computing the probability that a sphere of fixed radius centred at a monomer in one ring of length s contains beads belonging to other rings. It is the ensemble average of the observable $d^m(s)$, which, for a configuration of N_r rings with lengths l_i , $i = 1, \dots, N_r$, is

¹i.e. the distribution goes abruptly to 0 at a certain max length $L_r^{max}(K)$.

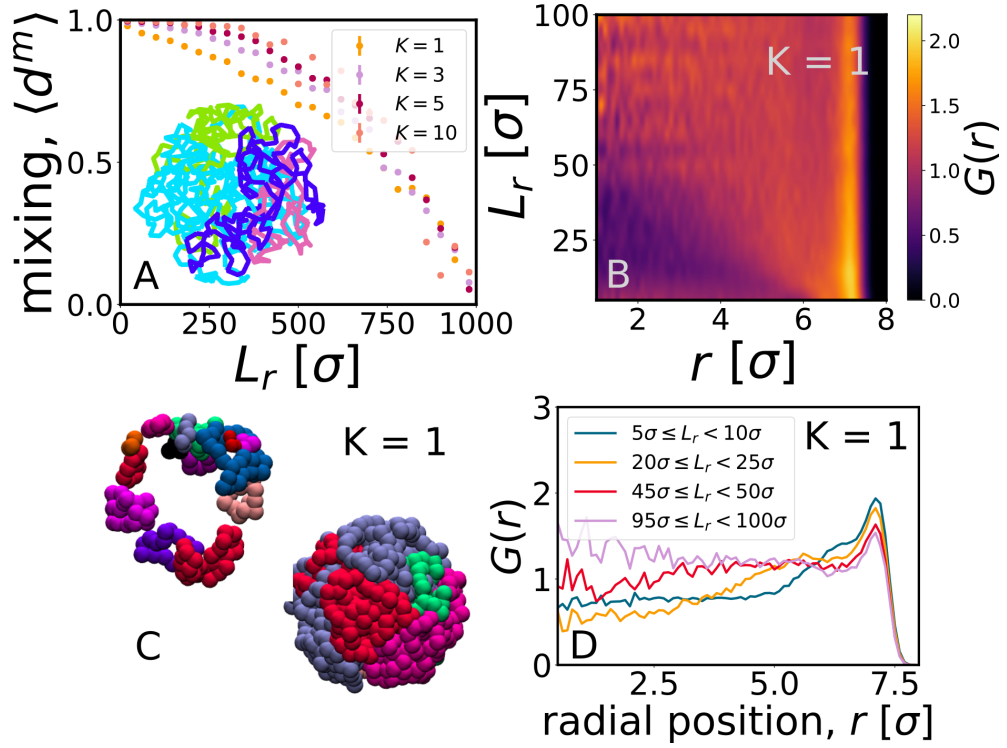


Figure 3.6: **A** Probability that a ring of length L_r is mixed with other rings as a function of L_r and persistence length K (Eq. (3.10) with $R_{sp} = 1.8\sigma$). **B** Heatmap of the radial distribution ($G(r)$) of monomers in rings of length L_r as a function of r , the distance from the centre of the confinement sphere, and L_r . **C** Snapshots from a configuration sampled with $K = 1$. Rings with $L_r < 50\sigma$ are shown on the left, rings with $L_r \geq 50\sigma$ on the right. **D** $G(r)$ for 4 different length windows, $5\sigma \leq L_r < 10\sigma$, $20\sigma \leq L_r < 25\sigma$, $45\sigma \leq L_r < 50\sigma$ and $95\sigma \leq L_r < 100\sigma$.

defined as

$$d^m(s) = \frac{\sum_{i=1}^{N_r} \delta_{l_i s} \sum_{j=1}^{l_i} \Theta(n_{R_{sp}}(i, j))}{\sum_{i=1}^{N_r} l_i \delta_{l_i s}}, \quad (3.10)$$

where

$$n_{R_{sp}}(i, j) = \sum_{\substack{m=1 \\ m \neq i}}^{N_r} \sum_{n=1}^{l_m} \Theta(R_{sp} - |\mathbf{r}_{ij} - \mathbf{r}_{mn}|) \quad (3.11)$$

is the number of beads belonging to rings $m = 1, \dots, N_l$, $m \neq i$, inside a sphere of radius R_{sp} centered in $\mathbf{r}_{ij}$, the position of the j -th monomer of the

i -th ring. In Eqs. 3.10 and 3.11, δ is the Kronecker delta,

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} . \quad (3.12)$$

and Θ is defined as

$$\Theta(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases} . \quad (3.13)$$

This mixing probability increases with stiffness, and decreases with ring length, so that longer and more flexible rings tend to segregate more efficiently. The inset of Fig. 3.6A shows an example of one such configuration with small mixing probability. In our system, increasing K leads to few longer rings in steady state (Fig. 3.3D), so that the two effects compete with each other. For our parameter range, the dominant one is that of stiffness, so that for larger K the fewer rings present in steady state mix more with each other.

3.3 Topological gelation is accompanied by the formation of a percolating network of linked loops

We now discuss more in depth the topological and physical nature of the transition shown in Fig. 3.3. Importantly, we find that, as the recombinant polymers get stiffer for a given value of R , the typical topologies change qualitatively. The topological entanglement of the confined rings can be quantified by looking at their mutual linking. This is done by computing the Gaussian linking number for each pair of rings γ_i and γ_j , which is given by

$$Lk(\gamma_i, \gamma_j) = \frac{1}{4\pi} \oint_{\gamma_i} \oint_{\gamma_j} \frac{\mathbf{r}_1 - \mathbf{r}_2}{|\mathbf{r}_1 - \mathbf{r}_2|^3} \cdot (d\mathbf{r}_1 \times d\mathbf{r}_2) . \quad (2.14)$$

At time t , the number of linked pairs $N_{Lk}(t)$, and the absolute value of the total linking number $|Lk|(t)$ are obtained by summing Eq. (2.14) over each pair of rings (i, j) , namely

$$N_{Lk}(t) = \frac{1}{2} \sum_{i,j}^{N_r(t)} \chi(i, j)(t) , \quad (3.14)$$

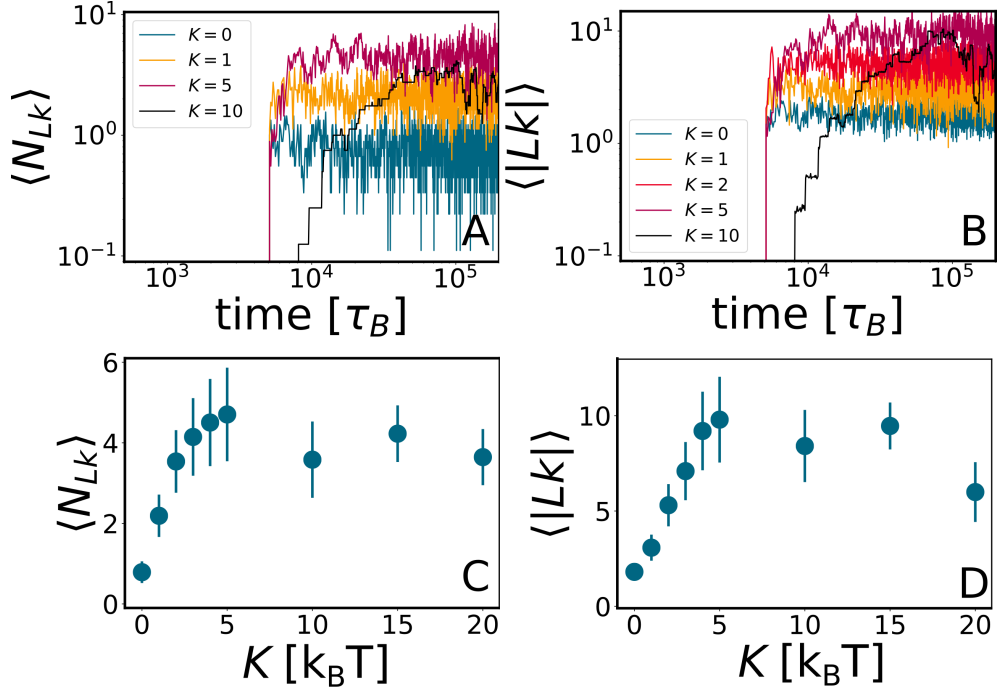


Figure 3.7: Time dependence of the number of linked pairs of rings N_{Lk} **A** and average of the absolute value of the total linking number, $\langle |Lk| \rangle$, **B** produced during digestion. Different curves refer to different values of the stiffness parameter K . In panels **C** and **D** we report the average steady-state values of the observables of panels **A** and **B** respectively as a function of K .

where $N_r(t)$ is the number of rings at time t and

$$\chi(i, j)(t) = \begin{cases} 1 & |Lk(\gamma_i, \gamma_j)|(t) > 0.5 \\ 0 & |Lk(\gamma_i, \gamma_j)|(t) \leq 0.5 \end{cases}, \quad (3.15)$$

and

$$|Lk|(t) = \frac{1}{2} \sum_{i,j} |Lk(\gamma_i, \gamma_j)|(t). \quad (3.16)$$

Fig. 3.7A,B shows that the number of linked pairs of rings and the average total absolute value of the linking number both increase with the stiffness K . Flexible rings are therefore typically short and unlinked, whereas stiffer loops are typically longer and linked. This trend in link complexity is highlighted by the distribution of the absolute value of the pairwise linking number reported in Fig. 3.8 for different values of the flexibility K . Notably, for $K = 15$ there are configurations in which a pair of recombinant rings are linked with

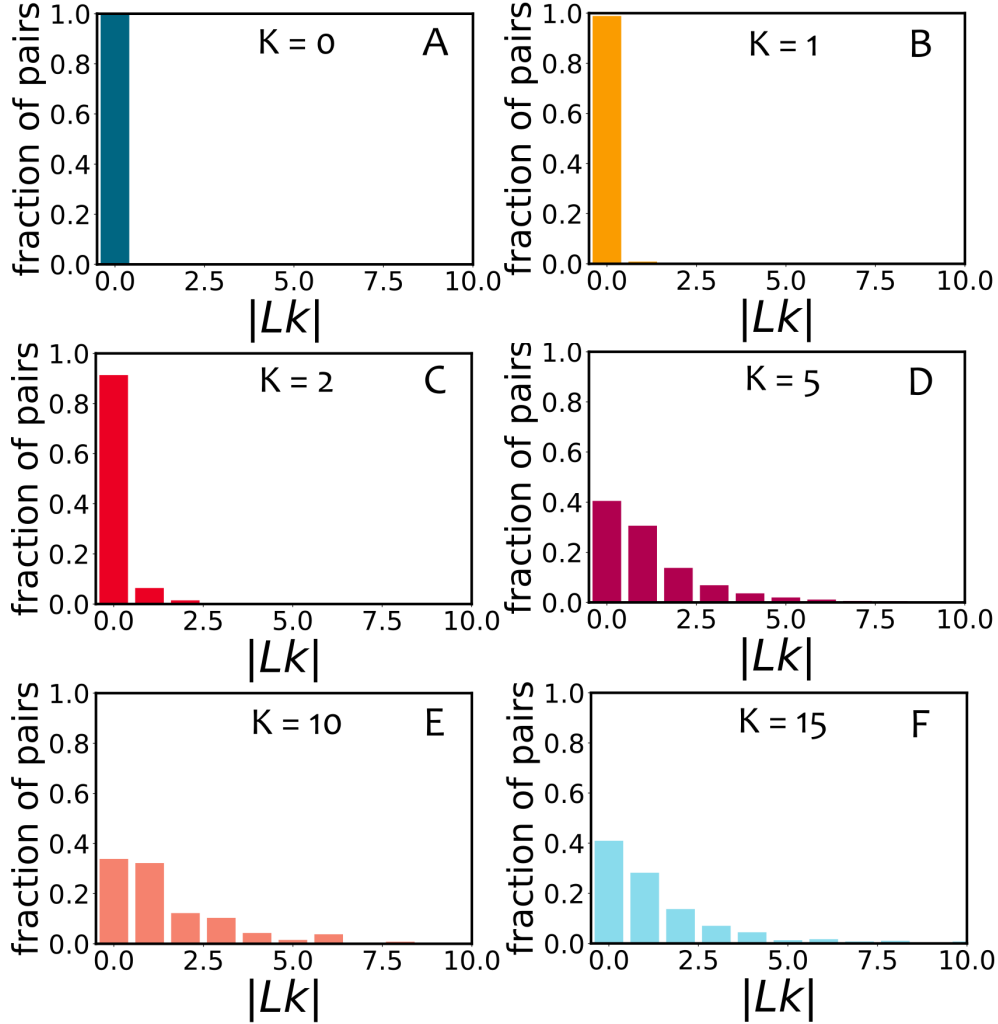


Figure 3.8: **(A-F)** Distribution of the absolute value of the linking number of pairs of recombinant rings for $K = 0$, $K = 1$, $K = 2$, $K = 5$, $K = 10$ and $K = 15$.

$|Lk| = 10$. The transition or crossover which we observe by increasing stiffness is therefore associated with an increase in topological entanglement and in linking between chains. Because linking between the rings should eventually endow our system with a non-zero elastic modulus, we refer to this transition as *topological gelation*. We note that the increase in linking between loops is the topological manifestation of the increase in mixing noted in the previous Section.

Up till now, we have mostly investigated the effect of increasing the stiffness of the system. As discussed in the geometric characterisation of the transition

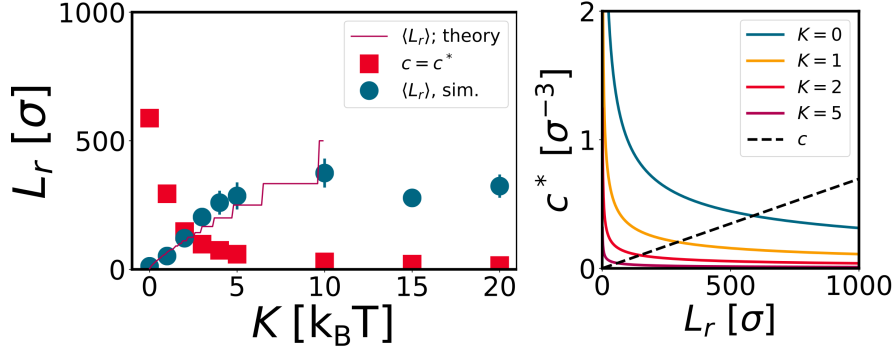


Figure 3.9: Critical length L^* for which the concentration c of a ring confined in a sphere of radius $R = 7\sigma$ equals its critical concentration c^* vs mean length of the rings as a function of K . At L^* , the curves (c^*) and black line (c) shown in the panel to the right intersect. c^* is predicted for an ideal polymer ring (3.17).

or crossover between a short ring and long ring regime, gelation can also be achieved by increasing the ring density at a fixed value of the stiffness. More insight into this transition, and its underlying physical mechanism, can be gained by analysing the overlap between recombinant rings of different stiffness. Given a certain value for the average length of the rings, which we call L , we can define a critical overlap concentration as

$$c^* = \frac{3L}{4\pi R_g^3}, \quad (3.17)$$

with R_g the radius of gyration of a ring of length L ,

$$R_g^2 = \frac{1}{L} \sum_{i=1}^L (\mathbf{r}_i - \langle \mathbf{r} \rangle)^2, \quad (3.18)$$

where $\mathbf{r}_i$ is the position of the i -th bead. This is the concentration above which rings of size L start to overlap with each other, and can be computed as a function of L by assuming, for instance, $R_g^2 = (L/2K)(2K)^2/12$ for $K \geq 1$, $R_g^2 = L/12$ for $K = 0$. The value of the critical length, for which the concentration c of a ring confined inside a sphere of radius $R = 7\sigma$ equals c^* , is shown in Fig. 3.9A (red squares). It is a decreasing function of K : in other words, increasing K makes it easier for rings to feel and overlap with each other. Given that the average length of the recombinant rings (blue circles and purple line in Fig. 3.9A) displays the opposite behaviour, there will be a value of the stiffness parameter K for which the average length and the

critical overlapping length coincide. This value qualitatively marks the point separating the short ring from the long ring regime. In more detail, when the red curve (filled squares) in Fig. 3.9A is above the blue curve (filled circles), then $c < c^*$, and we expect rings should not overlap and therefore segregate, leading to the short ring phase (Fig. 3.1B). Instead, when the red curve is below the blue one, then $c > c^*$, and we expect rings to mix and recombine significantly with each other, leading to the long ring phase (Fig. 3.1C).

This is akin to a gelation transition, as it is reasonable to assume that the system behaves as a liquid of soft particles (the rings) for $c < c^*$, and as a gel (or a solid-like structure) for c sufficiently larger than c^* . The main difference with respect to colloidal gels of soft particles is that here direct attractive interactions between the rings (particles) are replaced by topological interactions, as above c^* rings can link with each other. For this reason we refer to the transition between the short and long ring phases as *topological gelation*, and we shall reinforce this interpretation in the analysis described in the following sections.

Finally, note that, in our system, increasing K leads to a decrease in c^* , whereas decreasing R leads to an increase of c . This is the reason why, as anticipated above, the system can be made to gel either by increasing K or by decreasing R : both variations increase the value of c/c^* , hence favour the long ring phase.

As topological gelation is approached, the nature of the typical network of linked rings forming in steady state changes qualitatively, as shown in Fig. 3.10. For low K (in the liquid phase), the network has low connectivity, and clusters have typically one or very few nodes (Fig. 3.10C). In sharp contrast, for large K (in the gel phase), there is a large connected component which accounts for a substantial fraction of all rings. The formation of a giant connected component can be quantified by measuring the probability that the network condenses into a single connected cluster of nodes, which appears to depart from 0 abruptly at the critical value of $K \simeq 2$, which is where $c/c^* \simeq 1$ according to our theory. In our system, rings outside the largest connected component are typically unlinked, so that the average fraction of rings in the largest component is approximately equal to the linking probability for any ring in the network. Fig. 3.10A shows that gelation occurs when this linking probability proxy is about $1/2$.

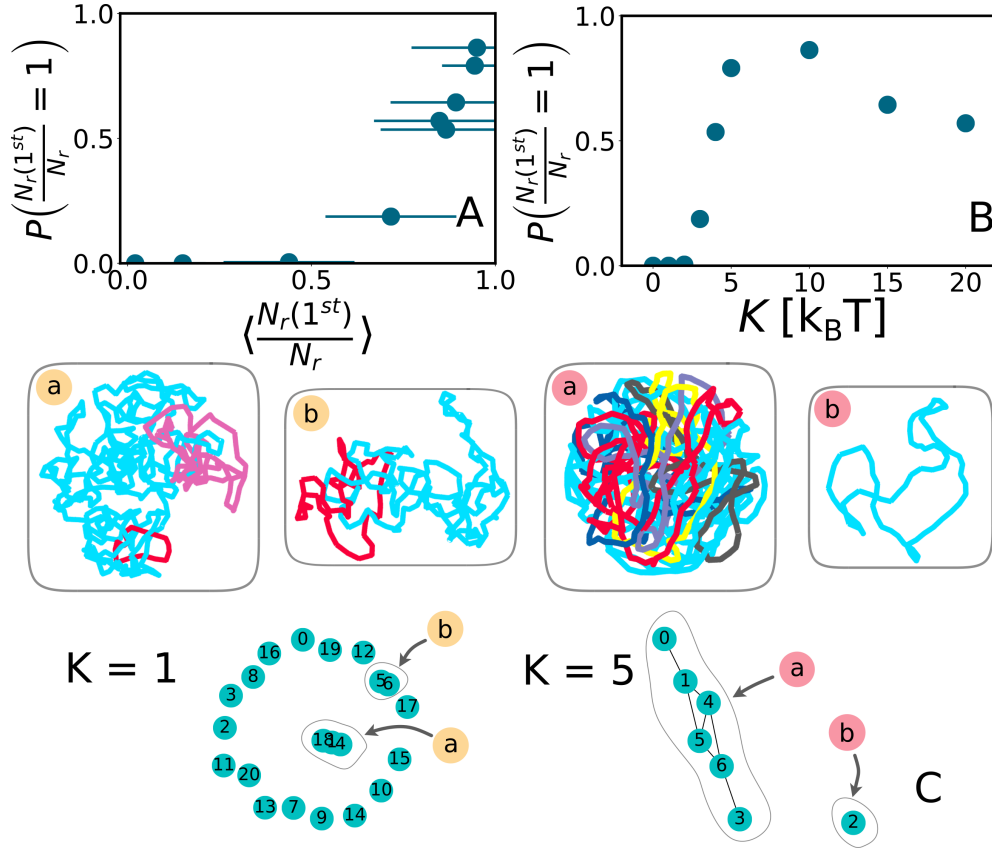


Figure 3.10: Probability of observing a single cluster as a function of the linking probability **A** and the stiffness parameter K **B**. **C** Snapshots of clusters of linked rings in simulations with $K = 1$ and $K = 5$. Each sampled configuration is associated with a network; linked rings (vertices) are connected by edges. A connected component of the network represents a cluster of linked rings.

3.4 Topological gelation traps the recombinant rings inside a permeabilised sphere

The topological gelation transition can be vividly shown by the following procedure, which is inspired by DNA filter elution [84, 85], a technique which proved to be useful, e.g., in studies of DNA strand breaks and crosslinks induced by carcinogenic agents [84], consisting in isolating the DNA contained within a cell by trapping it inside a membrane filter. After a steady state is reached in our simulations, we permeabilise the confining sphere by converting it into a spherical mesh with pores of controlled size. We then dis-

allow recombination and appreciate its products by monitoring the number of monomers still inside the sphere as a function of time, $n(t)$. Permeabilisation is implemented by using a rigid structure made of beads of size σ , whose interaction with the polymer beads is described by the WCA potential (Eq. (2.30)). The permeabilised sphere, which resembles a spherical grate, is built by overlapping rigid rings lying on planes perpendicular to one of the three axes of an orthogonal set (Fig. 3.11). The curves $n(t)$ are reported in Fig. 3.11A,B respectively for the liquid ($K = 0$) and gel ($K = 3$) regimes. Note that, to single out the effect of topology, immediately after permeabilisation, we set $K = 1$ for both cases. It is apparent that the two regimes lead to very different dynamics.

In the liquid phase, the “powder” of overwhelmingly small and unlinked rings rapidly diffuses out of the sphere, as their translational entropy increases if confinement is removed, and rings are small enough to translocate through the pores (see snapshots in Fig. 3.11A). It may take longer, but eventually also larger rings, which are typically unlinked (Fig. 3.8), find their way out the confinement sphere. This is qualitatively equivalent to experiments in permeabilised cells which show elution of sufficiently small molecules, such as diffusing proteins or DNA fragments of small size [85].

In sharp contrast, when the system is in the gel phase, the network of linked loops which emerges after recombination is too large and topologically too complex to translocate through the pores and most of the system is kinetically trapped inside the permeabilized sphere. This is apparent from the very slow decrease of $n(t)$ that, at very large times, reaches a plateau suggesting a new long lived steady state of the system, with a substantial proportion of rings still inside the sphere (see Fig. 3.11B). This second scenario is also reminiscent of elution experiments, where large superstructures, like microphase separated protein droplets or chromatin-protein aggregates such as transcription factories, resist elution and remain inside permeabilised nuclei [85].

A different scenario is observed if the recombination process is still active during elution. In Fig. 3.12A we report $n(t)$ for the liquid regime ($K = 0$). Strikingly, the number of monomers inside the sphere decreases until it reaches ~ 14 , which is the value expected in the case of monomers uniformly distributed inside the simulation box. The decay is well fitted through the whole range of time by a stretched exponential. This is mostly due to the fast escape dynamics of the short rings produced by the fragmentation of the recombinant polymers left inside the sphere (see snapshots in Fig. 3.12A).

In the gel regime ($K = 5$, Fig. 3.12B), even with the recombination process still active, very few rings are small enough to rapidly pass through the pores

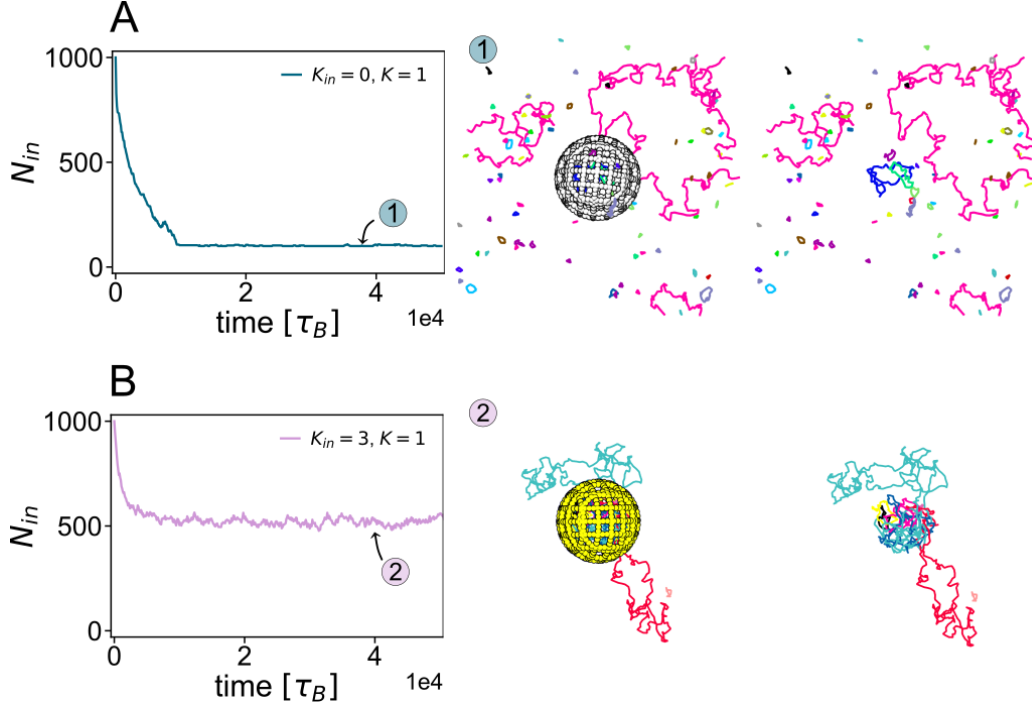


Figure 3.11: Results from simulations of recombined polymer rings escaping from a permeabilised sphere. Immediately after permeabilisation, recombination is deactivated and K set to 1. Number of monomers trapped inside the sphere as a function of time for $K = 0$ **A** and $K = 3$ **B**. The curves are decorated with snapshots from the simulations.

of the permeabilised sphere and the elution dynamics is mostly governed by the translocation of long and possibly linked rings through a pore. This is in general a very slow process, but in this case we note that the pores may act in synergy with recombination by forcing two segments of the same ring to be proximate in space, increasing the probability of occurrence of recombination events that eventually simplify the long rings while they exit the sphere (see snapshots in Fig. 3.12B).

We remark that, in this second set of simulations, the stiffness of the recombinant ring is not altered after permeabilisation.

3.5 Discussion and conclusion

In summary, here we have used coarse-grained molecular dynamics simulations to study the behaviour of a solution of polymer rings undergoing recom-

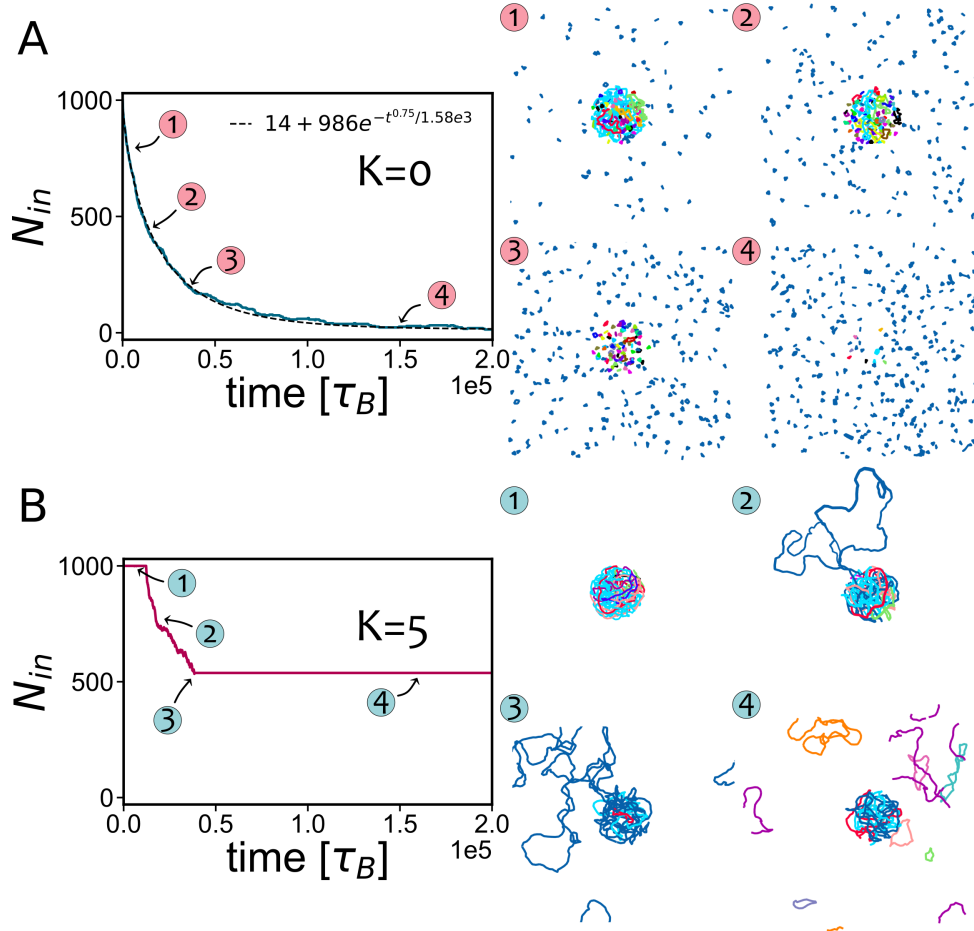


Figure 3.12: Simulations of recombinant rings escaping from a permeabilised sphere. Recombination events can occur throughout the elution process. **(A)** Number of monomers still inside the sphere as a function of time for a system of fully flexible rings ($K = 0$), and snapshots taken at different times. Rings which escaped confinement are coloured in cyan, different confined rings are drawn with different colours. **(B)** Number of confined monomers as a function of time for $K = 5$ and corresponding snapshots.

bination reactions inside a spherical container. Our theoretical predictions could be tested *in vitro*, by using a mixture of DNA rings and recombinase enzymes inside a vesicle. The system is also relevant as a simplified framework to predict the possible behaviour of recombinant DNA *in vivo*.

Our main finding is that this recombinant polymer mixture harbours a transition, or crossover, between two fundamentally distinct regimes, which can

be triggered either by stiffening the polymer rings, or by increasing their density. For flexible or sufficiently dilute polymer rings, recombination results in the production of a gas, or fluid, or short segregated and unlinked loops. For polymers with a sufficiently large persistence length, or for a sufficiently dense ring solution, a network of long linked loops emerges in steady state. This transition is a topological analogue of the gelation transition observed for sufficiently dense suspensions of sticky colloids, where the formation of force chains is substituted by topological linking. Like gelation and vitrification, our topological transition is accompanied by a dramatic slowdown in the system dynamics, which can be vividly quantified, for instance, by measuring the rate of escape from the confining sphere when the latter is pierced by appropriate-size pores to permeabilise it, as in electroelution experiments with DNA or chromatin.

Topological gelation can be seen as the inverse of topological digestion by topoisomerase, as the topological complexity increases after gelation, whereas it decreases after digestion. We anticipate that topological gelation can also be found in the absence of spherical confinement, for instance by varying the stiffness of a suspension of polymer loops of a given volume fraction. In such a geometry, the transition could be characterised, for instance, by measuring the bulk rheology response of the system, as the elastic modulus should be non-zero in the gel phase.

Besides being of fundamental interest as an example of a topological phase transition in a soft condensed matter system, our results can be exploited to design DNA plasmids with desired topology. In this respect, it would be of interest to classify the typical links arising at different values of density and polymer stiffness. In analogy with the rich and tunable link and knot spectrum found with self-assembling rigid helices [86], we expect that by positioning our system close to the gelation transition we should be able to tilt the balance in the favour of either complex links or simple links of well-defined type. While we have not characterised the knot type of the recombinant rings, these are also likely to be non-trivial and parameter-dependent, adding further potential tunability to the system.

In the context of recombinant DNA *in vivo*, our results highlight the need for regulation of the biochemical action of recombination enzymes, which if left unchecked are likely to create a topological gel given the high density found under physiological conditions in the nuclei of living cells. Gelation is likely detrimental for the cell, just as DNA linking and knotting are, as it would lead to entanglements hindering, for instance, chromosome segregation during cell division. We hope that these ideas can be pursued in the future

by studying topological gelation with biomimetic experiments.

Chapter 4

Simplifying topological entanglements by entropic competition of slip-links

Knots and topological entanglements are often found in physical and biological systems [14, 87, 88]. Their uncontrolled formation and proliferation reduces the space of accessible configurations of generic polymer chains, in turn potentially affecting their mechanical [89] or biological [12, 90] functions. Entanglements are so inevitable and detrimental in the genome of living organisms that a specific class of highly conserved proteins – known as topoisomerases – has evolved to resolve them [91].

While it is typical to treat abundant topological constraints at the mean field level – for instance in the tube model of polymer melts [92] – exact and scaling results can be obtained via theories that replace entanglements with slip-links (SLs) which enforce contacts between polymer segments while allowing them to slide past each other [93, 94, 95, 96]. Because of this physically appealing analogy, systems of polymers with slip-links have been theoretically and numerically explored in the field of statistical and polymer physics, for instance to estimate the size of knots [24, 25] and the effective tube size in polymer melts [97].

Beyond the theoretical appeal of these systems, actual examples in which SLs affect the conformation of polymers can be found in nature and industry. For instance, the structural maintenance of chromosomes (SMC) proteins, such as cohesin and condensin, by bringing together two DNA segments, act as bridges [98] or SLs on the genome and, depending on the condition and organism, they can either actively move [99, 100, 101, 102] or passively diffuse

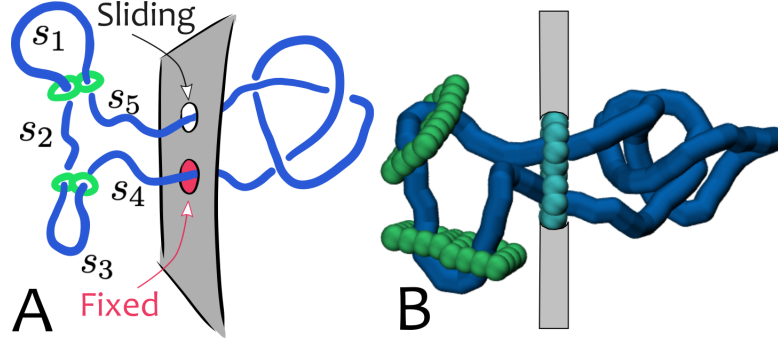


Figure 4.1: **A** Our set up is inspired by that of entropic competition [24]: a ring polymer with total length L is tied in a knot and contains diffusing SLs. The ring is forced to pass through two holes on a wall that separates the two sides. Only the contour passing through the top hole is allowed to slide, while the other is fixed. In the figure, the left-handed loop has $N_{SL} = 2$ SLs and is partitioned into $2N_{SL} + 1 = 5$ segments of which s_1 and s_3 are called peripheral; $s_2 + s_4 + s_5$ makes up the inner loop. The right hand side contains a 3_1 knot. **B** Snapshot from molecular dynamics simulations of the system in **A**. Figure adapted from Ref. [2].

(sliding-links) [103, 104, 105, 98], in turn extruding loops. These proteins are key for the organisation of the genome *in vivo* [106, 107, 108], for the structure of sister chromatids [109, 108] and to regulate chromatid decatenation and cell division [110]. Additionally, it was recently suggested that they cooperate with topoisomerases to maintain the genome topologically simple and entanglement-free [111]. While the structural role of SMC proteins is now widely explored, the mechanisms through which they regulate genome topology are far less understood or investigated. At the same time, SLs can be realised in synthetic and supramolecular chemistry using cyclodextrins [112] and are employed for instance to make molecular machines [113] and slide-ring gels [114]. In these cases, the benefit of using SLs is that they effectively act as mobile cross-links, thus imposing strong, yet plastic, topological constraints on the polymeric substrates. Thanks to this peculiar feature, gels of polyrotaxanes display toughness and self-healing abilities far superior than those of traditionally cross-linked materials [115]. Like in biological systems, also the SL-mediated topological regulation of synthetic slide-ring gels is only starting to be investigated [116].

In this work we study the interplay between SLs and topological entanglements formed by knots tied on a ring polymer (see Fig. 4.1). While our set up is inspired by the existing framework of entropic competition [117,

118, 24, 119, 25], here we consider SLs as real components of the system, modelling the presence of (diffusing) SMC proteins or cyclodextrins. In particular, we compute the loop size distribution and length of knotted segments for a variety of SL network topologies, and show that diffusing SLs are able to localise knotted segments purely by entropy, *without any external energy input*. Importantly, and in marked contrast with previous works on entropic competition, we show that the efficiency of entanglement localisation depends on the particular topology of the SL network and that including the action of topoisomerases (modelled as transient strand-crossings) leads to extremely fast and efficient simplification of complex knots. Our results suggest an entropy-driven mechanism through which generic SL-like molecules can regulate the topology of DNA or synthetic polymers.

4.1 Model

We prepare the initial configuration of our system by joining two bead-spring polymer segments (see chapter 2) on either sides of a wall; typically, we consider configurations in which the two polymer segments are either knotted or contain SLs. Since in our model the polymer cannot cross the wall, itself or the SLs, the topology on each side is preserved throughout the simulations (see below).

More specifically, we perform molecular dynamics simulations of bead-and-spring polymers made up of beads of size σ . The simulation box is separated into two equally large portions by an impenetrable wall and each bead interacting with the wall is subject to a repulsive force perpendicular to the wall described by the potential:

$$U_{\text{wall}}(r) = \begin{cases} \epsilon \left[\frac{2}{15} \left(\frac{\sigma}{r} \right)^9 - \left(\frac{\sigma}{r} \right)^3 - \frac{2}{3} \sqrt{\frac{5}{2}} \right] & r \leq r_c \\ 0 & r > r_c \end{cases}, \quad (4.1)$$

where r is the separation between the bead and the wall and $r_c = \left(\frac{2}{5} \right)^{\frac{1}{6}} \sigma$. Note that this expression is obtained by integrating the Lennard-Jones potential (see Ch. 2, Eq. (2.30)) over half space, namely by assuming that the interaction between each block of the wall and the bead is described by the Lennard-Jones (WCA, with the cutoff) potential. Beads of the polymer are strung together by FENE bonds (Eq. (2.29)) and interactions between beads are governed by the WCA potential (Eq. 2.30). To account for the effect of stiffness (arising, for instance, if the polymers represent chromatin fibres), we introduced a Kratky-Porod potential acting on triplets of neighbouring

beads (Eq. (2.31)). The initial configurations are designed such that a polymer passes through two holes in a wall, the size of which is small enough to let just one monomer through at a time. In this way, the polymer is divided into two segments, physically separated by the wall. The two segments do not interact with each other but for exchanging monomers through the holes in the wall, so that the total configurational entropy of the system is given by the sum of the entropy of the two sides. Additionally, we fix the position of one of the two beads closest to the wall so that the segments can exchange monomers only through the other hole (see Fig. 4.1). This set up is reminiscent of the one used to study entropic competition in Ref. [119].

Slip-links (SLs) are modelled as physical square hand-cuffs and are allowed to slide diffusively along the polymer. Specifically, we consider SLs with side 4σ , held together by four FENE bounds and opened in a “planar” arrangement with a sufficiently stiff (i.e. large l_p in Eq. (2.31)) Kratky-Porod potential.

4.1.1 Sampling of the system

To make sure that we are sampling our system ergodically and that are avoiding kinetic traps which may exist in MD simulations, we take snapshots of our system every 10^6 LAMMPS time steps ($10^4\tau_B$), comparable to the de-correlation time of a polymer of length up to at least 200 monomers. Within this time also the SLs loaded onto the substrate have travelled a significant amount of contour length [120] and so we assume that each sampled configuration is independent of others. Some of our results obtained with polymers of length 500 may thus display longer de-correlation times (see Figs. 4.2). Significantly improving the statistics would, however, require demanding simulations which would take months with tens of CPUs running in parallel. In Fig. 4.2 we plot the radius of gyration $R_g(t)$ of a polymer and the autocorrelation function $G_O(t)$ of a scalar observable O of the polymer (the length of one side, specifically) as a function of time. These are estimated as

$$R_g^2(t) = \frac{1}{L} \sum_{i=1}^L (\mathbf{r}_i(t) - \langle \mathbf{r}(t) \rangle)^2, \quad (3.18)$$

where $\mathbf{r}_i$ is the position of the i -th bead, L is the length of the polymer and $\langle \mathbf{r}(t) \rangle = \frac{1}{L} \sum_{i=1}^L \mathbf{r}_i(t)$, and

$$G_O(t) = \frac{1}{\sigma_O^2(T-t)} \sum_{i=1}^{T-t} (O(i) - \langle O \rangle) (O(i+t) - \langle O \rangle), \quad (4.2)$$

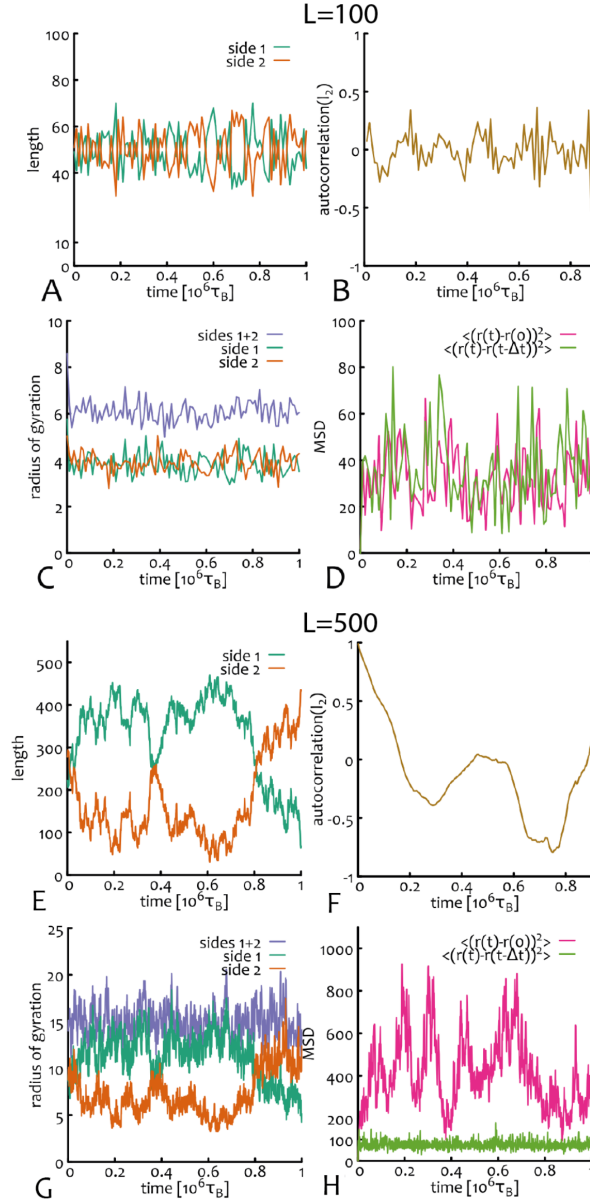


Figure 4.2: Conformational de-correlation time of a 100 (top) and 500 (bottom) beads long polymer in the 2SLs vs 2SLs configuration. **A,E** Length of the two competing sides of the polymer as a function of time. **B,F** Normalised autocorrelation function of the length of one of the two competing sides (labelled as side 2). **C,G** Radius of gyration in units of σ as a function of time. **D,H** Mean squared displacement in units of σ^2 of a bead of the polymer as a function of time. $\Delta t = 10^4 \tau_B$. Figure adapted from Ref. [2].

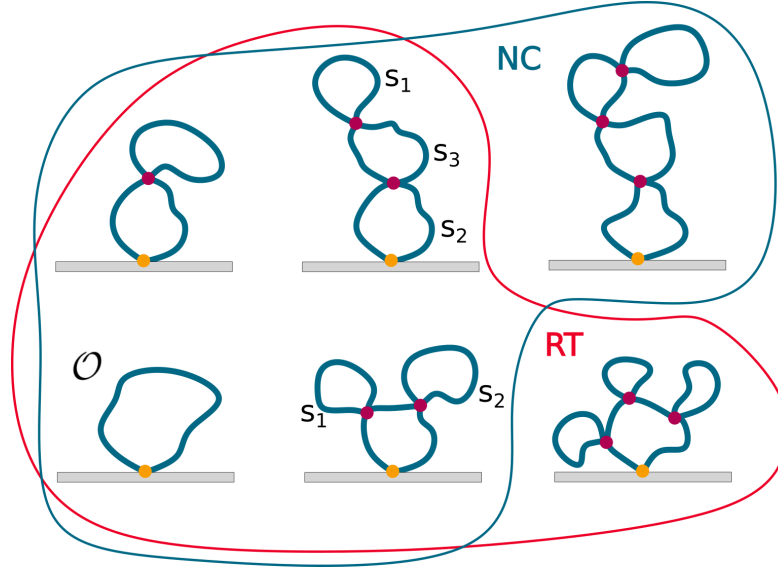


Figure 4.3: Examples of members of two families of topologies of polymer networks: the “surface round table” (red, RT) and “surface necklace” topologies (blue, NC). Orange vertices are grafted to a wall. $\mathcal{O}$ is a simple surface loop and, together with a surface loop with only one slip-links (top left), belongs to both the NC and RT families. The two topologies in the central column can also be seen as both grafted RT and grafted NC networks, as they would be identical without the wall.

where T is the number of sampled configurations and $O(i)$ is the value of O for the i -th configuration, $\langle O \rangle = \frac{1}{T} \sum_{i=1}^T O(i)$ and $\sigma_O^2 = \frac{1}{T} \sum_{i=1}^T (O(i) - \langle O \rangle)^2$. By comparing Fig. 4.2B and F one can appreciate that the 500 beads long chain displays a slower decay of the autocorrelation of the length of side 2 ($G_{l_2}(t)$). This autocorrelation appears to decay around $10^5 \tau_B$ while we sample the system every $10^4 \tau_B$ timesteps.

4.2 Competition of SLs

We first study the competition between symmetric SL networks: we load up to 3 SLs on each side of the wall in either a “round table” (RT) configuration, realised by loading the slip-links in series [24] (Fig. 4.3 and 4.4A,C,D), or a “necklace” (NC) configuration, characterised by nested loops [24] (Fig. 4.3 and 4.4B). How the total number of monomers of the system is partitioned between the two sides is predicted by scaling analysis as detailed below.

4.2.1 Equations for SLs vs SLs using Duplantier scaling for polymer networks

We start by considering monodisperse networks, where branches are equal in length. The number of configurations of an asymptotically long monodisperse polymer network of total length L and topology $\mathcal{G}$ in the dilute regime scales as [36]

$$\mathcal{Z}_{\mathcal{G}}(L) \propto \mu^L L^{\gamma_{\mathcal{G}}-1}, \quad (4.3)$$

where μ is the non-universal connectivity constant and

$$\gamma_{\mathcal{G}} - 1 = -\nu d\mathcal{L} + \sum_{\mathcal{M} \geq 1} n_{\mathcal{M}} \sigma_{\mathcal{M}} \quad (4.4)$$

is a topology dependent exponent: d is the dimension, $\mathcal{L}$ is the number of loops, $n_{\mathcal{M}}$ the number of vertices of multiplicity – i.e., the number of branches connected to the vertex – $\mathcal{M}$ and $\sigma_{\mathcal{M}}$ the star polymers exponents. ν is the universal metric exponent relating the size of a polymer to its length.

If the network is grafted to a surface, Eq. (4.3) changes to [36]

$$\mathcal{Z}_{\mathcal{G}}^S(L) \propto \mu^L L^{\gamma_{\mathcal{G}}^S-1}, \quad (4.5)$$

where

$$\gamma_{\mathcal{G}}^S - 1 = -\nu(d\mathcal{L} + \mathcal{V}_S - 1 + \mathcal{L}_S) + \sum_{\mathcal{M} \geq 1} (n_{\mathcal{M}} \sigma_{\mathcal{M}} + n_{\mathcal{M}}^S \sigma_{\mathcal{M}}^S). \quad (4.6)$$

$n_{\mathcal{M}}^S$ is the number of surface vertices (those grafted to the surface) of multiplicity $\mathcal{M}$, $\mathcal{V}_S = \sum_{\mathcal{M} \geq 1} n_{\mathcal{M}}^S$ is the total number of surface vertices, $\mathcal{L}_S = \sum_{\mathcal{M} \geq 1} \mathcal{M} n_{\mathcal{M}}^S$ is the number of branches connected to the surface and $\sigma_{\mathcal{M}}^S$ are the surface star polymers exponents.

Surface loop and surface round table exponents

Let us call surface loop ($\mathcal{O}$) a simple loop with a single vertex grafted to a surface (Fig. 4.3). A surface loop and the “surface” round table (RT) and necklace (NC) topologies we considered in our simulations (see Fig 4.3) have only one surface vertex of multiplicity 2: $\mathcal{V}_S = 1$ and $\mathcal{L}_S = 2$. Furthermore, every other vertex of both the RT and NC topologies has multiplicity 4, and the number of loops of both families of topologies is $N_{SL} + 1$, where N_{SL} is the number of vertices which are not grafted to the surface. According to Duplantier [36], then,

$$\sigma_2^S = 2\nu - 1, \quad (4.7)$$

hence, by using Eq. (4.6),

$$\gamma_{\mathcal{O}}^S - 1 = -d\nu - 1, \quad (4.8)$$

and

$$\gamma_{NC, N_{SL}}^S - 1 = \gamma_{RT, N_{SL}}^S - 1 = \gamma_{RT, N_{SL}} - 2 = -d\nu(N_{SL} + 1) + \sigma_4 N_{SL} - 1. \quad (4.9)$$

Surface round table and necklace scaling with the lengths of the loops

Since its vertices can slide along the contour and since it is not monodisperse, Eqs. (4.5) and (4.6) do not directly quantify the number of configurations of a slip-linked polymer network (i.e. polymer network with sliding vertices, namely the slip-links). However, they can be extended to the RT networks with N_{SL} slip-links if, analogously to what predicted for the RT networks without surface [24], we assume their number of configurations scales as

$$\begin{aligned} \mathcal{Z}_{RT, N_{SL}}(L) &\propto \mu^L A(s_1, \dots, s_{N_{SL}}, L) \left(L - \sum_{i=1}^{N_{SL}} s_i \right)^{\gamma_{RT, N_{SL}}^S - 1} \\ &\times \mathcal{X} \left(\frac{s_1}{L - \sum_{i=1}^{N_{SL}} s_i}, \dots, \frac{s_{N_{SL}}}{L - \sum_{i=1}^{N_{SL}} s_i} \right), \end{aligned} \quad (4.10)$$

where s_i , $i = 1, \dots, N_{SL}$ indicate the lengths of the loops subtended by the slip-links (“slip-linked” loops, e.g. s_1 and s_2 in Fig. 4.3), $A(s_1, \dots, s_{N_{SL}}, L)$ accounts for the sliding entropy of the slip-links and $\mathcal{X}$ is a function to be determined.

Round table We now suppose, as we see *a posteriori* in our simulations, the peripheral slip-linked loops of a RT network are much smaller than the internal loop. They contribute equally to Eq. (4.10):

$$\begin{aligned} \mathcal{Z}_{RT, N_{SL}}(L) &\propto \mu^L A(s_1, \dots, s_{N_{SL}}, L) \left(L - \sum_{i=1}^{N_{SL}} s_i \right)^{\gamma_{RT, N_{SL}}^S - 1} \\ &\times \prod_{i=1}^{N_{SL}} \left(\frac{s_i}{L - \sum_{j=1}^{N_{SL}} s_j} \right)^\alpha. \end{aligned} \quad (4.11)$$

Furthermore, on the basis of this assumption, we require the scaling contribution of the largest loop to be that of a surface loop¹. This sets α to

$$\alpha = \frac{\gamma_{RT, N_{SL}}^S - 1}{N_{SL}} - \frac{-d\nu - 1}{N_{SL}} = -d\nu + \sigma_4, \quad (4.12)$$

which is the same loop closure factor for the RT topologies without surface. Finally,

$$\mathcal{Z}_{RT, N_{SL}}(L) \propto \mu^L A(s_1, \dots, s_{N_{SL}}, L) \left(L - \sum_{i=1}^{N_{SL}} s_i \right)^{-d\nu-1} \prod_{i=1}^{N_{SL}} s_i^{-d\nu+\sigma_4}. \quad (4.13)$$

Surface necklace with 2 SLs

In the case of the necklace topology with 2 slip-links, we write

$$\begin{aligned} \mathcal{Z}_{NC,2}(L) &\propto \mu^L B(s_1, s_2, L) (L - s_1 - s_2)^{\gamma_{NC,2}^S - 1} \\ &\quad \times \left(\frac{s_1}{L - s_1 - s_2} \right)^{\alpha_1} \left(\frac{s_2}{L - s_1 - s_2} \right)^{\alpha_2}, \end{aligned} \quad (4.14)$$

where we labelled as 1 the slip-linked loop grafted to the wall and where $B(s_1, s_2, L)$ accounts for the sliding entropy. This time, we can not state $\alpha_1 = \alpha_2$ *a priori*. However, if we require the scaling contribution of the internal loop (s_3 in Fig. 4.3), which we assume to be the largest, to be that of a surface loop, and the contribution of the external slip-linked loop (s_1 in Fig. 4.3) to be the usual closure factor (Eq. (4.12)), the final result is

$$\mathcal{Z}_{NC,2}(L) \propto \mu^L B(s_1, s_2, L) (L - s_1 - s_2)^{-d\nu-1} s_1^{-d\nu+\sigma_4} s_2^{-d\nu+\sigma_4}. \quad (4.15)$$

Notice the consistency of Eqs. 4.13 and 4.15 with the hypothesis of the peripheral loops being small depends on the form of the sliding entropy terms A and B . For the surface necklace this might be a point of discussion; nonetheless, Eq. (4.15) represents our simulated data reasonably well (see Fig. 4.4B).

4.2.2 Scaling behaviours for competing SLs

The number of configuration of each side of a system of competing SLs scales as the integral over all lengths compatible with the given topology of either

¹Note that the sliding entropy of the slip-links is accounted for by A in Eq. (4.10). See [24] for the case without surface.

Eq. (4.13) or Eq. (4.15) divided by A or B respectively. Their product yields the scaling of the total number of configurations.

For instance, the number of configurations of two slip-links competing with each other (Fig. 4.4A) scales as:

$$\begin{aligned} \mu^L \int_{s_1^m}^{l-s_2^m-s_3^m} ds_1 (l-s_2^m-s_3^m) \frac{s_1^{-3\nu+\sigma_4}}{(l-s_1)^{3\nu+1}} \\ \times \int_{s_1^m}^{L-l-s_2^m-s_3^m} ds_1 (L-l-s_2^m-s_3^m) \frac{s_1^{-3\nu+\sigma_4}}{(L-l-s_1)^{3\nu+1}}, \end{aligned} \quad (4.16)$$

where the coefficients s_i^m account for the physical size of the slip-links: they are the minimum size of the segments of the slip-linked polymer; L and l are, respectively, the total length and the length of one side of the polymer. Note the integrals account for the sliding entropy of the slip-links.

We employ Eqs. (4.13) and (4.15) to derive semi-analytically the distribution of lengths of one side of systems with up to 3 SLs on each side. We then verify these predictions with molecular dynamics (MD) simulations and report the results in Fig. 4.4. Figure 4.4A shows that the distribution undergoes a unimodal-bimodal transition for increasing ring sizes (we consider 35 and 100 beads-long polymers). This transition, which, as shown in Fig. 4.4A, can be viewed as a pitchfork bifurcation, is analogous to the symmetry-breaking behaviour found with phantom chains in [117, 118].

We now turn to the case of 2 and 3 SLs on each side (Fig. 4.4B,C,D). We find that the spontaneous symmetry breaking is much more noticeable in the NC topology than in the RT one (Figs. 4.4B,C). Finally, the RT topology displays a stable symmetric state for all polymer lengths when $N_{SL} = 3$ SL are loaded on each side (Fig. 4.4D) with a remarkable absence of symmetry breaking. The histograms computed numerically in Fig. 4.4B,C,D nicely match the predicted distributions in most cases.

In light of these results we argue that the network topology arising from the loading of multiple SLs – either in series (RT) or in parallel (NC) – profoundly affects the spatial organisation of the underlying DNA or synthetic polymer. Further, our semi-analytical results are in line with previous works which showed that nested SLs (NC topology) promote growth, or extrusion, of the outer loop via a ratcheting process in which the diffusive dynamics of the outer loop is rectified by the entropic pressure of the inner ones [103].

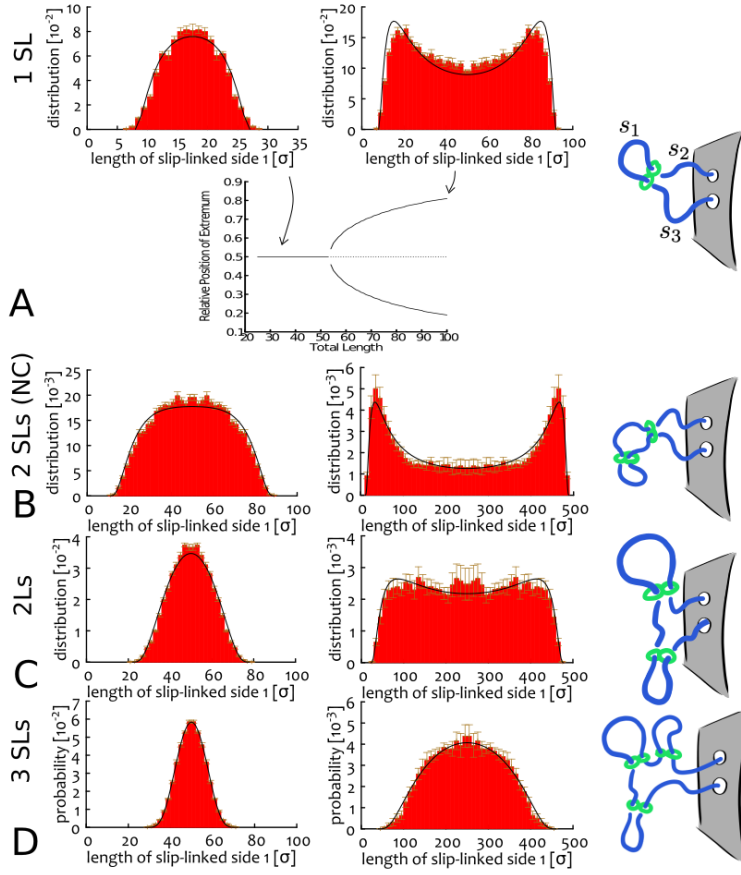


Figure 4.4: From top to bottom: **A** a single SL. The black solid lines in the panel to the bottom of row **A** are the maxima of the length distribution (panels to the top) as a function of the total length of the system. The dashed line represents the minima. **B** 2 SLs in a necklace configuration; **C** 2 SLs in a round-table configuration; **D** 3 SLs in a round-table configuration. Note that we simulate two SL networks with identical topology, one on each side of the wall. Each histogram collects information from 10^4 configurations sampled every 10^6 LAMMPS steps from 100 independent replicas. The black curves are obtained as described in the previous section. From left to right: increasing total length of the ring polymer (35, 100 beads in **A**) and (100, 500, beads) in **B-D**. Figure taken from Ref. [2].

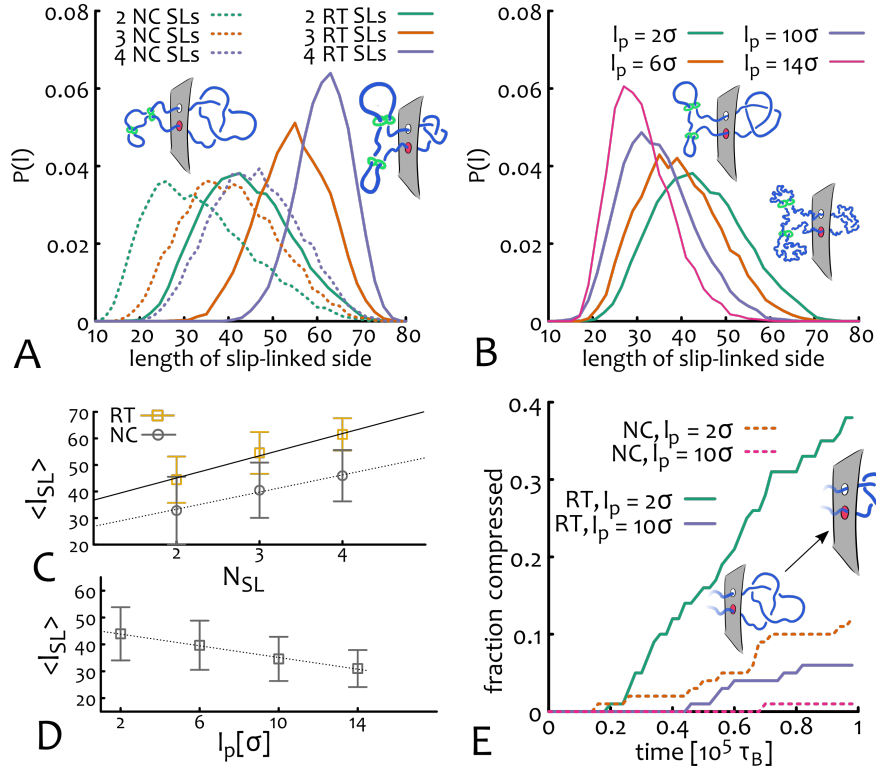


Figure 4.5: **A** Plot of the distribution of lengths of the side with the SLs, for different SL number and topology. The plot shows that the RT configurations (solid lines) favour longer SL sides and hence tighter knots than NC configurations (dashed lines). **B** Distribution of lengths of the side with the SLs, for different substrate persistence lengths: the curves refer to the case of a RT configuration with 2 SLs. The plot shows that stiffer substrates favour longer knotted sides. **C** Mean length ($\langle l_{SL} \rangle$) and standard error of the SL side length as a function of SL number and topology. The solid and dashed lines are best linear fits with slopes 8.4(8) for RT and 6.5(6) for NC. **D** Mean length ($\langle l_{SL} \rangle$) and standard error of the SL side length as a function of substrate persistence length l_p . The solid line is the best linear fit with slope $-1.09(5)l_p$. **E** Fraction of knots that have been localised – i.e., confined within 20% of the overall ring contour length – during the course of a simulation starting from a situation in which 80% of the contour length is on the knotted side. Higher flexibility and RT configurations both favour knot localisation. NC=necklace; RT=round-table; SLs=slip-links. Figure adapted from Ref. [2].

4.3 SLs versus knots

The results of the previous section suggest that entropic competition of SLs can regulate the distribution of loop lengths of an underlying polymer. We

now ask whether this entropic competition may also provide a mechanism for the localisation and topological simplification of entanglements such as knots that may occur on biological or synthetic polymers. To address this question, we now consider the case where the competing topologies are a SL network and a trefoil (3_1) knot (as in Fig. 4.1), and aim to identify the optimal conditions to localise and simplify the knotted loop. It should be noted that our calculations are performed in thermal equilibrium, so we seek entropic and topological mechanisms that bias the free energy towards a lower knotting probability. This is different from previous works which studied the effects of energy-consuming mechanisms driving unknotting [111].

Since the semi-analytical approach we used in the previous sections cannot be easily extended to this setup, we directly perform MD simulations of a 100 beads-long ring and address how different factors such as SL number, topology and polymer persistence length affect the entropic competition and, in turn, regulate the probability and efficiency of knot localisation.

First, we find a large difference between RT versus NC configurations (Fig. 4.5A): the larger sliding entropy and low looping cost of the RT set-up provides an entropic pressure which outcompetes the ratcheting effect of the NC topology. This result is unexpected, since it is known that NC configurations promote asymmetric polymer conformations and the formation of large loops with respect to RT networks (see Fig. 4.4B,C and [103]). Nevertheless, the RT arrangement appears to be more suited at overcoming the knot entropy. Quantitatively, the entropic pressure of the RT configurations is notable: already with 3 SLs the knotted side is localised to a shorter contour length than the SL side (Fig. 4.5A).

Additionally, our simulations reveal that the bending rigidity of the polymer substrate also plays an important role (Fig. 4.5B). Indeed, there are two potential and contrasting enthalpic effects that the bending rigidity may have on the entropic competition: on the one hand, the larger the stiffness, the longer the contour length required to form a knot without tight bends (this effect is particularly relevant for short loops); on the other hand, bending rigidity enhances the formation of large loops via diffusive loop extrusion [120]. Our simulations show that – for 2 SLs in a RT arrangement and a 100-bead polymer ring – the former effect dominates: the larger the persistence length, the larger the knotted side, rendering the localisation of entanglements less efficient (Fig. 4.5B). This result shows that the entropic pressure provided by the SLs can only provide localisation if the polymer substrate is sufficiently flexible.

More quantitatively, in Fig. 4.5C we show that the mean length of the SL-

side, $\langle l_{SL} \rangle$, grows linearly as a function of the number of SLs (in the range of N_{SL} explored here) and that the growth rate is faster for RT configurations. At the same time, $\langle l_{SL} \rangle$ appears to also decrease linearly with polymer stiffness as $\sim l_0 - l_p$ with l_0 a constant (Fig. 4.5D); this is expected, since the minimum number of monomers required to tie a knot with a given topology increases linearly with l_p (e.g., in a freely jointed chain). In a practical genomic or synthetic context, the results reported in Fig. 4.5A-D overall suggest that local SL density and substrate flexibility may provide handles to tune the typical size of knots. *In vivo* both these parameters depend on local DNA sequence and transcriptional activity, and we thus speculate that knot and entanglement localisation may be achieved to a different extent in different genomic regions; in particular, transcriptionally active genes (which are also thought to be more flexible) may harbour smaller knots and localised entanglements. At the same time, different synthetic polymers having different stiffness and different arrangements of cyclodextrins may be designed, so that the resulting cross-linked gels have tunable properties [115].

4.4 Kinetics of topological simplification

Another important aspect of the problem is *how fast* knots can be localised, and eventually removed.

To measure the kinetics of knot localisation, we perform MD simulations in which the system is initialised far from equilibrium with a large knot – occupying about 80% of the total contour length – and measure how long it takes for the system to revert the situation and compress the knotted size to 20% of the total contour length. Figure 4.5E shows that RT configurations and flexible substrates are faster at localising the knot, on top of being more efficient in steady state, as discussed before.

Finally, to address the question of whether diffusing SLs can drive the topological simplification of knots and entanglements we now include topoisomerase-mediated strand-crossing reactions in our model as follows. We study a system in which a short segment of the knotted side is allowed to undergo strand-crossings. We then initialise the system with a given knot (either a 7-crossings torus knot 7_1 , a 10-crossings torus knots 10_{124} , or a twist knot 7_2) taking up 75% of the total contour length on the right-hand side and load SLs in RT arrangements on the other side. We chose to set only a short segment (10 beads) close to the wall as crossable to model the presence of a topoisomerase closely upstream of a SL-like SMC protein, as suggested by recent experimental evidence [121, 122] (note that in this context we view the

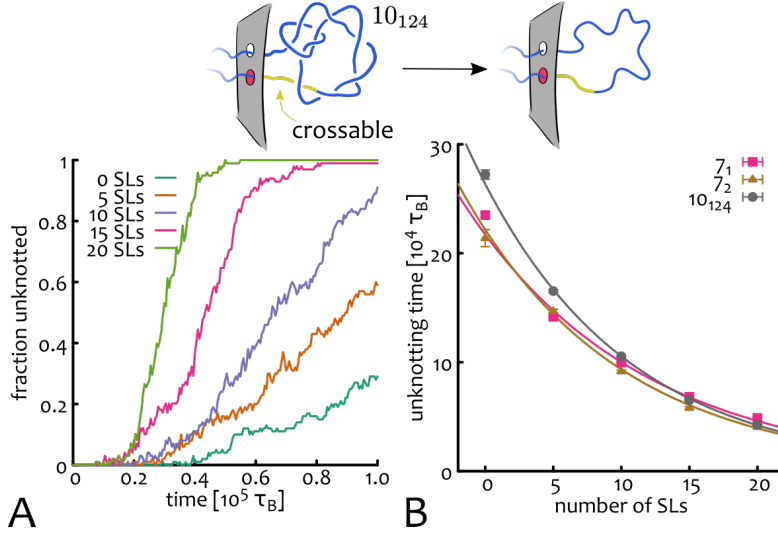


Figure 4.6: **A** Fraction of 10_{124} knots that have simplified to unknots during the course of simulations of a polymer with $N = 400$ beads, starting from a situation in which 80% of the contour length is on the knotted side. In this simulation, the 10 closest beads to the bead fixed at the wall on the knotted side are used to allow strand-crossings to model topoisomerase action to a SL. Loading more SLs increases the pressure on the knot and hence the rate of unknotting. **B** For all the knot topologies considered ($7_1, 7_2, 10_{124}$), the simplification time (the time it takes to simplify a knot to the unknot with unitary probability) shows an exponential decrease as a function of the number of SLs. For the 10_{124} , rates are computed by using the curves shown in **A**. SL=slip-link. Figure adapted from Ref. [2].

wall in Fig. 4.1 as a SL itself). This is achieved by setting to 0 the amplitude of the (WCA) potential regulating the interaction between the beads of this segment and each bead composing the rest of the polymer, excluding the two immediately next to the segment boundaries. The crossable segment is kept on the knot side, by placing it next to the fixed bead (wall, see Model section).

With this set-up, we observe that the more SLs are loaded onto the substrate, the faster the rate of the topological simplification (see Fig. 4.6A). The quantitative speed-up is striking, as, regardless of the topology of the knot, the rate grows exponentially with N_{SL} (Fig. 4.6B). An explanation of this is that the mean length of the SL side scales linearly with the number of SLs (Fig. 4.5C), so that the mean knotted length is $\langle l_K \rangle = L - \langle l_{SL} \rangle \simeq L - kN_{SL}$. Since the probability of observing a trivial knot on a polymer x segments

long scales as $P_0(L) \sim e^{-x/x_0}$ [9, 123], with x_0 a model-specific parameter, we expect $P_0(\langle l_K \rangle) \sim e^{-\langle l_K \rangle/L_0} \sim e^{N_{SL}/L_0}$. In other words, as the knot size becomes smaller, its entropic cost increases sharply and it becomes exponentially harder to prevent its simplification to the unknot. [Note that this argument implicitly assumes that simplification is slow with respect to localisation.]

4.5 Conclusions

In summary, in the context of a search for possible mechanisms of topological regulation and simplification of entanglements in DNA and other polymeric systems, here we have investigated the interplay between the entropy of SL networks and of knots.

In the case where no knots are tied on the ring, (SL-only case, Fig. 4.4), we provide semi-analytical predictions for the distribution of polymer lengths which compare well to MD simulations. We then consider a set-up in which an SL network competes entropically with a trefoil knot and we find that the network can localise the knot efficiently, by just relying on its higher entropic pressure and in the absence of any motor activity associated with the SLs (such an activity is known to be absent in yeast cohesin [98] and cyclodextrins [115]). We dissect the effects of the number of SLs, the topology of the SL network, and the polymer flexibility on the knot localisation discovering that round-table (or “in series”) SL arrangements are best suited at confining a knot. Additionally, when the substrate is stiff, the enthalpic contribution from the bending energy dominates and the knot swells, hampering localisation. Finally, we have also included topoisomerase in the model, which is relevant to study topological simplification of the genomes of living cells [111] or self-healing gels of polyrotaxanes with reversible bonds [124]. Strikingly, we find that entropic pressure alone is sufficient to simplify very complicated knots reliably and extremely fast, with the simplification rate increasing exponentially with number of SLs.

We hope our findings may help unveil the mechanisms of topological simplifications of DNA by means of SL-like proteins such as SMC. Our results may also inform strategies for regulating the topology of synthetic polymer networks, for instance by using SLs to control the size of knots inside them, which can in turn affect their material properties.

Chapter 5

Asymptotics of linked polygons

Entanglement *complexity* affects the physical and statistical properties of polymers, as seen for polymer knots translocating through small pores [20]. It is therefore important to predict how likely it is for a polymer to be knotted or linked to another polymer with a given topology. While the scaling behaviour, entropy and knot localisation of ring polymers with fixed topology have been studied extensively [23, 24, 25], how mutual entanglement affects the entropy of a system composed of multiple ring polymers is a largely unexplored, and mostly limited to the case in which no topological restrictions are included [125, 32, 126, 127, 128]. This is although *mutual* entanglement strongly impacts the physical properties of polymer solutions and plays an important role in biological processes such as replication [10]. To address this problem, we use Monte Carlo methods and scaling analysis to study the critical behaviour of the entropy of up to 4 rings linked together with a given topology.

We model polymer rings as lattice polygons, i.e. embeddings of a 1-sphere S_1 into the simple cubic lattice Z^3 (polymer loops on Z^3). Let us call $p_n^{(k)}(n_1, \dots, n_k | \tau)$ the number of embeddings of k polygons with $n_1, \dots, n_k$ edges forming a link with topology τ and total length $n = \sum_{i=1}^k n_i$. We are interested in counting $p_n^{(k)}(\tau) = \sum_{n_1, \dots, n_k} p_n^{(k)}(n_1, \dots, n_k | \tau)$, the number of embeddings with a total of n edges distributed in all possible ways among the k polygons. We first prove some rigorous results about the exponential growth rate of $p_n^{(k)}(\tau)$, we then make the reasonable assumption $p_n^{(k)}(\tau) \sim \mu(\tau) n^{\Gamma(\tau)-1}$ in the limit of asymptotically large n , and numerically investigate how the critical exponent $\Gamma(\tau)$ depends on the topological complexity. We provide evidence that one ring grows at the expense of the others, while the remaining components act as one or more roots sliding along

the dominant ring, each contributing +1 to the critical exponent.

5.1 Rigorous results about the exponential growth rate

Let us first recall some useful results. Let us call p_n the number of polygons with n edges. It is known that the limit

$$\lim_{n \rightarrow \infty} n^{-1} \log p_n = \sup_{n \geq 4} n^{-1} \log p_n \equiv \log \mu \quad (5.1)$$

exists [129]; μ is called growth or connectivity constant of the lattice. If we instead consider only the unknotted polygons with n edges, and if p_n^0 is their number, it can be shown that [9, 8]

$$\lim_{n \rightarrow \infty} n^{-1} \log p_n^0 = \sup_{n \geq 4} n^{-1} \log p_n^0 \equiv \log \mu_0 < \log \mu. \quad (5.2)$$

Eqs. (5.1) and (5.2) indicate that both the number of polygons and the number of unknotted polygons grow exponentially with the number of edges, and that unknotted polygons are exponentially rare. Few results are known about linked polygons. One is that the exponential growth rate of a set of k *unknotted* polygons with the same number of edges, n/k , linked with arbitrary topology τ^0 , does not depend on the link type nor on the number of components [130]:

$$\lim_{n \rightarrow \infty} n^{-1} \log p_n^{(k)}(n/k, n/k, \dots, n/k | \tau^0) = \log \mu_0. \quad (5.3)$$

Note that n in Eq. (5.3) is the total number of edges. A similar but weaker result holds true for *knotted* polygons linked with topology τ^1 [130]:

$$\log \mu_0 \leq \lim_{n \rightarrow \infty} n^{-1} \log p_n^{(k)}(n/k, n/k, \dots, n/k | \tau) \leq \log \mu. \quad (5.4)$$

Above, we used τ^0 to differentiate links with all components unknotted from the rest. For convenience, we now denote τ any given link topology, and, only when needed, specify if its components are unknotted or not. We now extend these two results to polydisperse linked polygons.

¹Note that the knot type of the components is included in τ , and the unknot is a knot type.

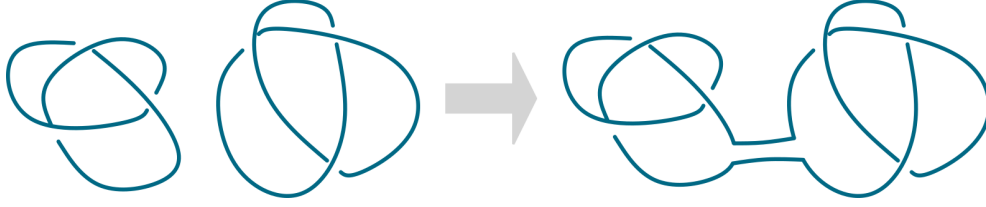


Figure 5.1: Sketch of the connected sum of two (trefoil) knots. Two small arcs, one from each knot, are removed; the knots are then connected by inserting a new pair of arcs.

Theorem 5 *If τ is an unsplittable link with k components with all components unknotted then*

$$\lim_{n \rightarrow \infty} n^{-1} \log p_n^{(k)}(\tau) = \log \mu_0.$$

Proof: A lower bound follows from Eq. 5.3, since $p_n^{(k)}(n/k, n/k, \dots, n/k | \tau) \leq p_n^{(k)}(\tau)$. The upper bound is obtained as follows. Consider k polygons with $n_1, n_2, \dots, n_k$, $\sum_{i=1}^k n_i = n$, edges respectively. Embed the first polygon. Embed each of the other $k - 1$ polygons, independently, in such a way that their barycentre is contained inside a cubic box of side $(kn/2)$ centred at the first polygon. This can be done in at most

$$p_{n_1}^0 p_{n_2}^0 \dots p_{n_k}^0 (kn/2)^{3(k-1)} \leq \mu_0^{n_1+n_2+\dots+n_k} (kn/2)^{3(k-1)} = \mu_0^n (kn/2)^{3(k-1)}$$

ways, where we have used the fact that $p_n^0 \leq \mu_0^n$ (Eq. (5.2)). All the links with topology τ and components of 1D size (i.e. number of edges) $n_1, \dots, n_k$ are included among those obtained by this construction. Therefore, summing over $n_1, n_2, \dots, n_k$,

$$p_n^{(k)}(\tau) \leq \mu_0^n (kn/2)^{3(k-1)} \binom{n-1}{k-1}.$$

Taking logarithms, dividing by n and letting $n \rightarrow \infty$ gives

$$\limsup_{n \rightarrow \infty} n^{-1} \log p_n^{(k)}(\tau) \leq \log \mu_0$$

which completes the proof. $\square$

If at least one component is knotted we have the following weaker result:

Theorem 6 *If τ is an unsplittable link with k components with at least one component knotted with fixed knot type then*

$$\log \mu_0 \leq \liminf_{n \rightarrow \infty} n^{-1} \log p_n^{(k)}(\tau) \leq \limsup_{n \rightarrow \infty} n^{-1} \log p_n^{(k)}(\tau) < \log \mu.$$

Proof: The lower bound $\log \mu_0 \leq \liminf_{n \rightarrow \infty} n^{-1} \log p_n^{(k)}(\tau)$ follows from inequality (5.4). The upper bound is obtained with the same construction used to prove Theorem 5, except we consider k polygons with knot types K_1, K_2 , etc., instead of k unknotted polygons. If the number of edges of the k polygons is $n_1, \dots, n_k$, subject to $\sum_{i=1}^k n_k = n$,

$$p_n^{(k)}(n_1, n_2, \dots, n_k | \tau) \leq (kn/2)^{3(k-1)} p_{n_1}(K_1) p_{n_2}(K_2) \dots p_{n_k}(K_k)$$

and summing over $n_1, n_2 \dots$ then gives

$$p_n^{(k)}(\tau) \leq (kn/2)^{3(k-1)} 2^{k-1} p_n(K_1 \# K_2 \# \dots \# K_k) \binom{n-1}{k-1},$$

where $K_1 \# K_2 \# \dots \# K_k$ is the connected sum of $K_1, K_2, \dots, K_k$ ² (Fig. 5.1), and hence

$$\limsup_{n \rightarrow \infty} n^{-1} \log p_n^{(k)}(\tau) < \log \mu$$

which completes the proof. $\square$

5.2 Growth rate

We recall that, including the first order correction to the scaling, it is generally assumed, for any knot topology K ,

$$p_n(K) = A(K) n^{\alpha(K)-3} \mu(K)^n \left(1 + \frac{B(K)}{n^\Delta} + \dots \right). \quad (5.5)$$

Analogously, it is reasonable to assume, for any given topology τ of links with k components,

$$p_n^k(\tau) = A(\tau) n^{\Gamma(\tau)-1} \mu(\tau)^n \left(1 + \frac{B(\tau)}{n^\Delta} + \dots \right), \quad (5.6)$$

²The inequality is obtained by catenating (Fig. 5.1) each given set of configurations of the knots $K_1, K_2, \dots, K_k$; to do so, it might be necessary to rotate one or more of them, hence the factor 2^{k-1} .

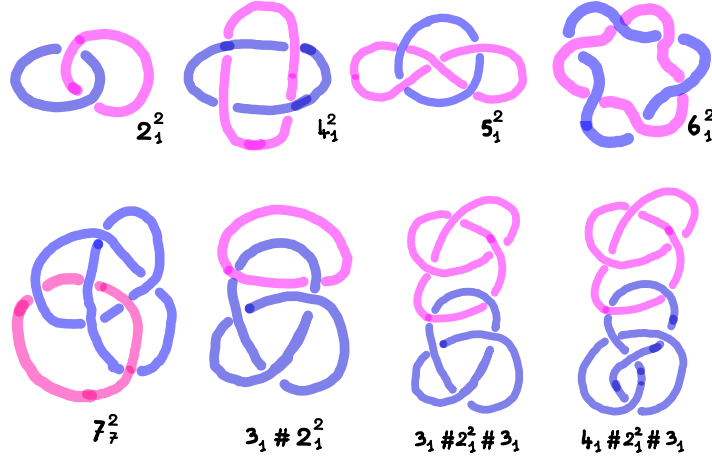


Figure 5.2: The two-component link types considered in this work. The links in the top row have both components unknotted, whereas the links in the bottom row have at least one component that is a prime knot. Figure taken from Ref. [3]. To identify a specific prime link type we use the Alexander-Briggs notation (see footnote 3). $\tau_1 \# \tau_2$ is the connected sum of τ_1 and τ_2 .

$\mu(\tau)$ is the growth constant, and $\Gamma(\tau)$ is the entropic exponent.

We aim to numerically validate theorems 5 and 6, and to estimate $\Gamma(\tau)$. To do so, we first build a Markov chain Monte Carlo (MCMC) algorithm to sample sets of linked polygons with fixed topology according to a known probability distribution (see theory of MCMC methods in chapter 2); the values of $\mu(\tau)$ and $\Gamma(\tau)$ are then extrapolated by comparing the numerical ensemble averages of some observables of interest with their theoretical counterpart obtained from Eq. (5.6).

The algorithm we use to sample sets of linked polygons with fixed link type is based on a variation of the BFACF algorithm [46, 47, 48], a grand canonical MCMC method commonly used to sample self-avoiding lattice polygons. Details of the algorithm and a sketch of the proof that its ergodic classes are the link types are included in chapter 2.

At long times, the BFACF algorithm samples the configuration space distribution (see chapter 2 for more details)

$$\pi_{\tau_0}(\omega) = \frac{1}{\Phi} n^q F^n \chi(\tau(\omega), \tau_0), \quad (5.7)$$

where τ_0 is the link type of the starting configuration, Φ is a normalisation constant, $\chi(a, b)$ is 1 when a and b are the same link type and zero otherwise,

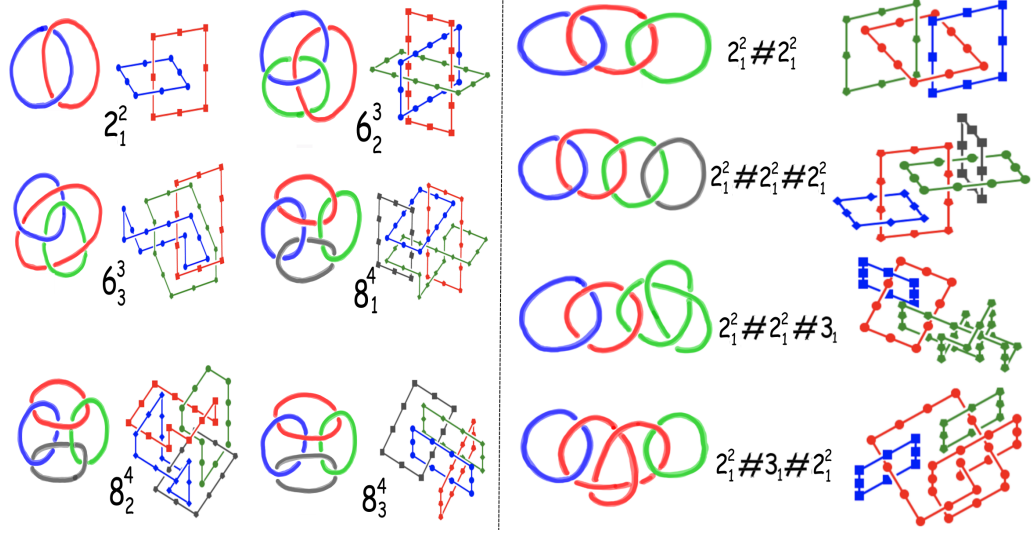


Figure 5.3: The links with more than two components considered here. The links in the left panel are prime, those in the right panel are composite. Figure taken from Ref. [4].

and n is the number of edges of ω . q and F are two parameters; the former affects the sampling efficiency, the latter controls the mean number of edges of the sampled configurations. F , the step fugacity, is of crucial importance: to test the scaling behaviour (5.6) we need to sample from a large range of lengths and approach the limit $n \rightarrow \infty$; this is achieved by running the BFACF algorithm at different values of F , as detailed below. From Eqs. (5.6) and (5.7), the mean number of edges in the k polygons comprising the link of type τ , sampled at q and F fixed, is given by

$$\langle n(\tau) \rangle \sim \frac{[\Gamma(\tau) + q]\mu(\tau)F}{1 - F\mu(\tau)} \left(1 - \frac{B(\tau)\Delta[1 - F\mu(\tau)]^\Delta}{\Gamma(\tau) + q} \right). \quad (5.8)$$

$\mu(\tau)$ and $\Gamma(\tau)$ can be computed by estimating $\langle n \rangle$ by Monte Carlo sampling with the algorithm introduced above, and considering its behaviour as $(1 - \mu(\tau)F) \rightarrow 0$. $F_c = 1/\mu(\tau)$ is the critical value of F for which the mean number of edges diverges and is inversely proportional to the growth rate $\mu(\tau)$. The Monte Carlo sampling was carried out by employing the multiple chain method described in chapter 2 and $q = 3$ [23]. We consider the following link types³ τ :

³We use the Alexander-Briggs notation for prime links: $\tau = X_z^y$, where X is the number of crossings, y the number of components and z an index labelling different link types with the same number of crossings and components.

1. The Hopf (2_1^2), Solomon (4_1^2), Whitehead (5_1^2) links, and 6_1^2 (see Fig. 5.2 top row). They are all characterized by having both components unknotted.
2. Two-component catenanes, based on 2_1^2 , in which one or both components are either a trefoil (3_1) or a figure-eight (4_1) knot (see Figure 5.2, bottom row). To highlight the presence of knots we use the notation $k_1 \# 2_1^2 \# k_2$ where k_1 and k_2 are the knot types of the rings 1 and 2 respectively. For comparison we consider the prime link 7_7^2 where one ring is a trefoil.
3. The three and four components prime links 6_2^3 (Borromean rings) and 6_3^3 , 8_1^4 , 8_2^4 and 8_3^4 (left of Fig. 5.3)
4. Three and four components catenanes based on 2_1^2 , in which either all components are unknot, $2_1^2 \# 2_1^2$ and $2_1^2 \# 2_1^2 \# 2_1^2$ (top-right of Fig. 5.3), or one is a trefoil and the others unknots, $2_1^2 \# 2_1^2 \# 3_1$ and $2_1^2 \# 3_1 \# 2_1^2$ (bottom-right of Fig. 5.3).

To leading order, we can use Eq. (5.8) to approximate $1/\langle n \rangle$ as

$$\langle n \rangle^{-1} \sim \frac{1 - F\mu(\tau)}{(\Gamma(\tau) + q)\mu(\tau)F} = \frac{1}{(\Gamma(\tau) + q)\mu(\tau)F} - \frac{1}{\Gamma(\tau) + q}. \quad (5.9)$$

Eq. (5.9) suggests that we can estimate $\mu(\tau)$ by extrapolating to that value of F for which $\langle n \rangle^{-1}$ is zero.

In Figures 5.4 (for links with two components) and 5.5 (for links with three and four components) we show the behaviour of $\langle n \rangle^{-1}$ as a function of F^{-1} for all the links considered. Although corrections to the scaling become more important as the link complexity increases, as $F \rightarrow F_c = 1/\mu(\tau)$ a linear behaviour is observed, in agreement with Eq. (5.9). By performing a linear extrapolation for the most asymptotic data we obtain the following estimates of $\mu(\tau)$.

$$\begin{aligned}
 \mu(2_1^2) &= 4.683 \pm 0.016 \\
 \mu(4_1^2) &= 4.682 \pm 0.029 \\
 \mu(5_1^2) &= 4.68 \pm 0.17 \\
 \mu(6_1^2) &= 4.68 \pm 0.12 \\
 \mu(7_7^2) &= 4.681 \pm 0.098 \\
 \mu(3_1 \# 2_1^2) &= 4.683 \pm 0.042 \\
 \mu(3_1 \# 2_1^2 \# 3_1) &= 4.68 \pm 0.11 \\
 \mu(4_1 \# 2_1^2 \# 3_1) &= 4.68 \pm 0.14
 \end{aligned} \quad (5.10)$$

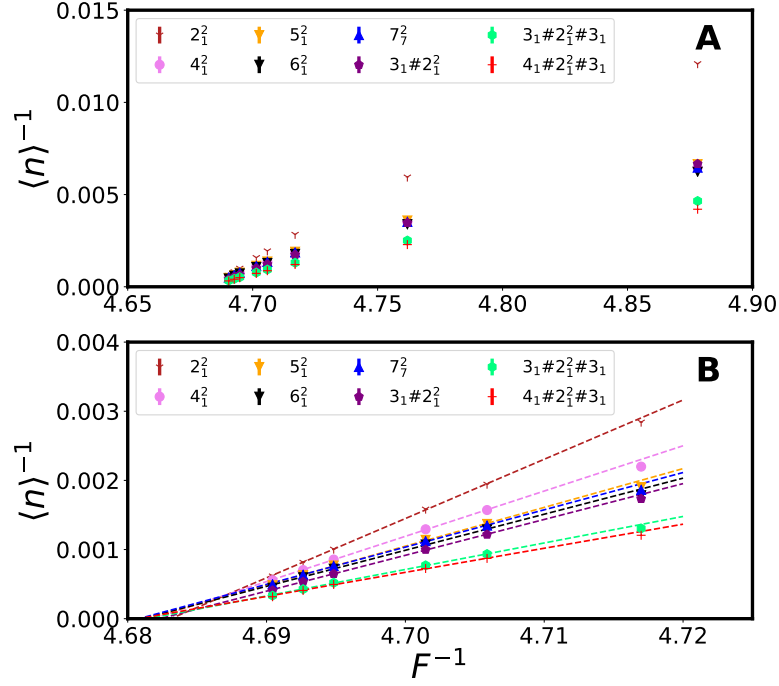


Figure 5.4: **A** Plot of $\langle n \rangle^{-1}$ as a function of $1/F$ for all the link types with two components considered (see label). **B** Zoom of panel A showing the linear behaviour as F approaches its critical value F_c . Figure adapted from Ref. [3].

$$\begin{aligned}
 \mu(6_2^3) &= 4.68 \pm 0.16 \\
 \mu(6_3^3) &= 4.682 \pm 0.08 \\
 \mu(8_1^4) &= 4.68 \pm 0.14 \\
 \mu(8_2^4) &= 4.68 \pm 0.15 \\
 \mu(8_3^4) &= 4.68 \pm 0.12 \\
 \mu(2_1^2 \# 2_1^2) &= 4.683 \pm 0.06 \\
 \mu(2_1^2 \# 2_1^2 \# 2_1^2) &= 4.683 \pm 0.06 \\
 \mu(3_1 \# 2_1^2 \# 2_1^2) &= 4.683 \pm 0.07 \\
 \mu(2_1^2 \# 3_1 \# 2_1^2) &= 4.683 \pm 0.04
 \end{aligned} \tag{5.11}$$

When all the components are unknotted we know from Theorem 5 that $\mu(\tau)$ is independent of τ and the numerical results are consistent with this. For the cases with knotted components the numerical estimates are also consistent with $\mu(\tau)$ being independent of τ .

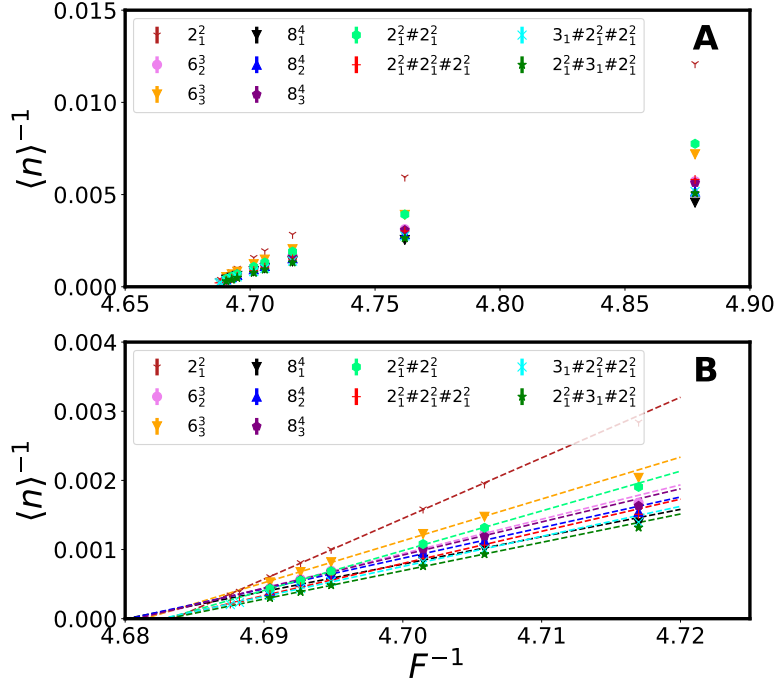


Figure 5.5: **A** Plot of $\langle n \rangle^{-1}$ as a function of $1/F$ for link types with more than two components and **B** zoom in on the region close to the critical value F_c . Figure taken from Ref. [4].

5.2.1 Critical exponents

To estimate $\Gamma(2_1^2)$ with a sufficiently high degree of confidence we triple the data available for 2_1^2 by performing two more simulations with different sets of F values,

$$F = \{0.206, 0.208, 0.2105, 0.2108, 0.2123, 0.2126, 0.2128, 0.21305\}$$

and

$$F = \{0.204, 0.207, 0.2085, 0.209, 0.2095, 0.211, 0.2114, 0.2117\}.$$

A direct estimate of $\Gamma(2_1^2)$ can then be obtained by inserting the estimate $\mu_0 = 4.683$ in the scaling law (5.9), compute for each value of F and the corresponding estimate of $\langle n \rangle$ an effective value $\Gamma(\tau|F)$, and finally extrapolate $\Gamma(\tau|F)$ as $F \rightarrow F_c$. The values $\Gamma(\tau|F)$ are reported in Figure 5.6 and a simple linear extrapolation of the most asymptotic data (see Figure 5.6B) gives:

$$\Gamma(2_1^2) = -0.766 \pm 0.004 \quad (5.12)$$

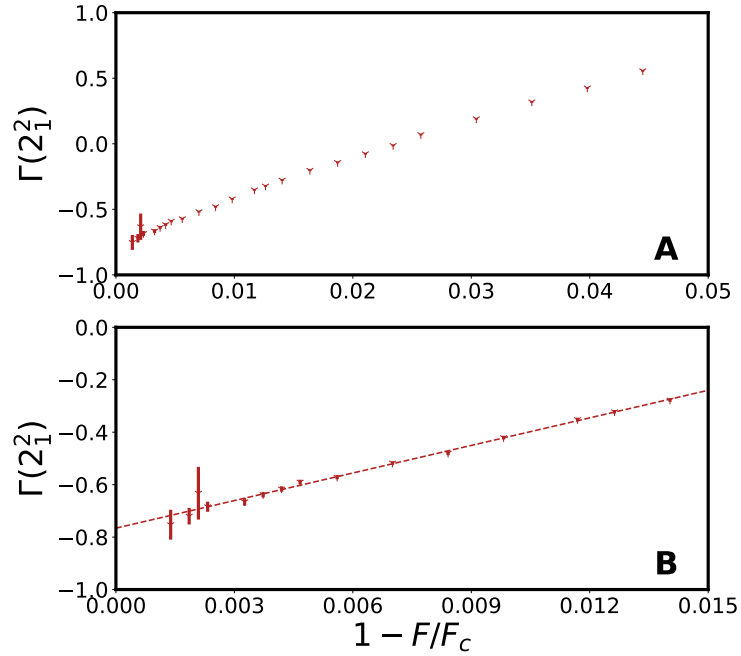


Figure 5.6: **A** Plot of the “effective” value of $\Gamma(2_1^2)(F)$, as obtained by equation (5.9) **B** Zoom of panel A showing the linear behaviour as F approaches its critical value F_c . Figure adapted from Ref. [3].

By recalling the best estimate of the entropic exponent for an unknotted ring, $\alpha_\emptyset = 0.237$ [23, 34] and the relation $\Gamma_\emptyset = \alpha_\emptyset - 2 = -1.763$ the estimate (5.12) suggests that $\Gamma(2_1^2) = \Gamma_\emptyset + 1^4$. In analogy with the picture proposed in Ref. [23] for prime knots this result suggests that, in the large n limit, the dominant configuration of the Hopf link corresponds to an unknotted rooted ring. This would also imply that the 1D size of the mutual entanglement should grow as $o(n)$ (see below for a more detailed discussion of this property).

If we accept that all the links have the same exponential growth rate then we can take ratios of equation (5.8) for different link types and investigate the dependence of the entropic exponent $\Gamma(\tau)$ on τ . By assuming $\Delta(\tau_1) = \Delta(\tau_2) = 1/2$ [23, 40], at leading order we obtain

$$\frac{\langle n(\tau_1) \rangle}{\langle n(\tau_2) \rangle} \simeq \frac{\Gamma(\tau_1) + q}{\Gamma(\tau_2) + q} \left(1 + c(1 - F\mu_0)^{1/2} \right). \quad (5.13)$$

By plotting $\langle n(\tau_1) \rangle / \langle n(\tau_2) \rangle$ against $\sqrt{1 - F\mu_0}$ we have curves that behave linearly as F goes to $F_c = 1/\mu_0$. See Figure 5.7

⁴This agrees with an independent estimate by Rechnitzer [131].

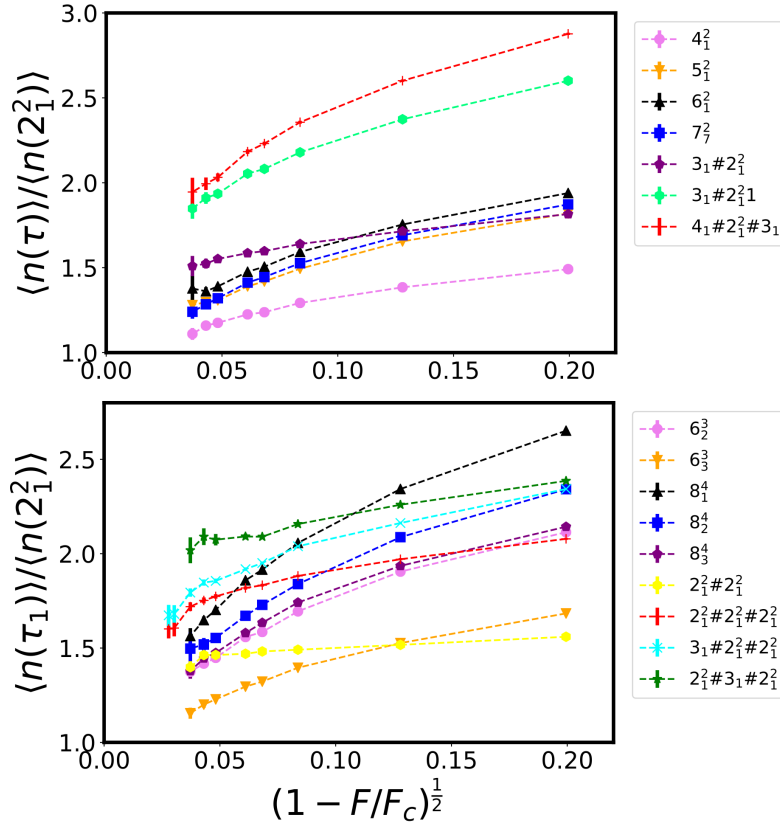


Figure 5.7: $\langle n(\tau) \rangle / \langle n(2_1^2) \rangle$ as a function of $\sqrt{1 - F/F_c}$. Extrapolating the curves as $F \rightarrow F_c$ gives estimates of the critical exponents. Figure adapted from Refs. [3] and [4].

By using linear extrapolations we estimate the following ratios, where $\Gamma_q(\tau) = \Gamma(\tau) + q$ and $q = 3$.

$$\begin{aligned}
 \Gamma_q(4_1^2) / \Gamma_q(2_1^2) &= 1.02 \pm 0.06 \\
 \Gamma_q(5_1^2) / \Gamma_q(2_1^2) &= 1.07 \pm 0.08 \\
 \Gamma_q(6_1^2) / \Gamma_q(2_1^2) &= 1.12 \pm 0.12 \\
 \Gamma_q(7_2^2) / \Gamma_q(2_1^2) &= 1.02 \pm 0.08 \\
 \Gamma_q(3_1 \# 2_1^2) / \Gamma_q(2_1^2) &= 1.43 \pm 0.09 \\
 \Gamma_q(3_1 \# 2_1^2 \# 3_1) / \Gamma_q(2_1^2) &= 1.6 \pm 0.1 \\
 \Gamma_q(4_1 \# 2_1^2 \# 3_1) / \Gamma_q(2_1^2) &= 1.6 \pm 0.2
 \end{aligned} \tag{5.14}$$

for link types with two components, and

$$\begin{aligned}
\Gamma_q(6_2^3)/\Gamma_q(2_1^2) &= 1.03 \pm 0.06 \\
\Gamma_q(6_3^3)/\Gamma_q(2_1^2) &= 0.99 \pm 0.05 \\
\Gamma_q(8_1^4)/\Gamma_q(2_1^2) &= 1.06 \pm 0.07 \\
\Gamma_q(8_2^4)/\Gamma_q(2_1^2) &= 1.07 \pm 0.1 \\
\Gamma_q(8_3^4)/\Gamma_q(2_1^2) &= 0.97 \pm 0.07 \\
\Gamma_q(2_1^2 \# 2_1^2)/\Gamma_q(2_1^2) &= 1.42 \pm 0.06 \\
\Gamma_q(2_1^2 \# 2_1^2 \# 2_1^2)/\Gamma_q(2_1^2) &= 1.4 \pm 0.1 \\
\Gamma_q(3_1 \# 2_1^2 \# 2_1^2)/\Gamma_q(2_1^2) &= 1.4 \pm 0.1 \\
\Gamma_q(2_1^2 \# 3_1 \# 2_1^2)/\Gamma_q(2_1^2) &= 1.9 \pm 0.1
\end{aligned} \tag{5.15}$$

for link types with three and four components.

For each of the prime links $7_7^2, 4_1^2, 5_1^2, 6_2^2, 6_3^2, 8_1^4, 8_2^4$ and 8_3^4 , these ratios are very close to unity and it is likely that the exponent is independent of prime link type. We know from 5.12 that $\Gamma(2_1^2)$ is about -0.76 , so the ratios $\Gamma_q(\tau)/\Gamma_q(2_1^2)$ where τ is $3_1 \# 2_1^2, 2_1^2 \# 2_1^2, 2_1^2 \# 2_1^2 \# 2_1^2$ and $3_1 \# 2_1^2 \# 2_1^2$ are all consistent with $\Gamma(\tau) = \Gamma(2_1^2) + 1$, i.e. $\Gamma_q(\tau)/\Gamma_q(2_1^2)$ around 1.45. In a similar way, the exponent differences $\Gamma(3_1 \# 2_1^2 \# 3_1) - \Gamma(2_1^2)$ and $\Gamma(4_1 \# 2_1^2 \# 3_1) - \Gamma(2_1^2)$ are between 0.9 and 1.8. This range is large but includes 1. For $\tau = 2_1^2 \# 3_1 \# 2_1^2$ we have $\Gamma_q(\tau)/\Gamma_q(2_1^2) = 1.9 \pm 0.1$, which is consistent with $\Gamma(2_1^2 \# 3_1 \# 2_1^2) = \Gamma(2_1^2) + 2$.

In the same way that the Hopf link behaves like a single polygon with a root, the links $3_1 \# 2_1^2, 3_1 \# 2_1^2 \# 3_1, 4_1 \# 2_1^2 \# 3_1, 2_1^2 \# 2_1^2, 2_1^2 \# 2_1^2 \# 2_1^2$ and $3_1 \# 2_1^2 \# 2_1^2$ all behave like a single polygon with two roots and $2_1^2 \# 3_1 \# 2_1^2$ behaves like a single polygon with three roots. This is discussed further in the next section.

5.2.2 Entropic competition

In the model that we are considering here the total number of edges in the link is fixed but the number of edges in the different components can vary. In this section we investigate how the numbers of edges in the various components change as n increases. Suppose, for the moment, that n_1 is the number of edges in the largest component of the link. In Figure 5.9 we plot $\langle n_1/n \rangle$ as a function of $\sqrt{1 - F/F_c}$ for various link types. As F approaches $F_c = 1/\mu_0$, i.e. as $\langle n \rangle$ increases, it is clear that one ring grows at the expense of the others. In every case, at the largest values of F , more than half the edges are in the largest ring. See Fig. 5.8 for a few snapshots of sampled configurations.

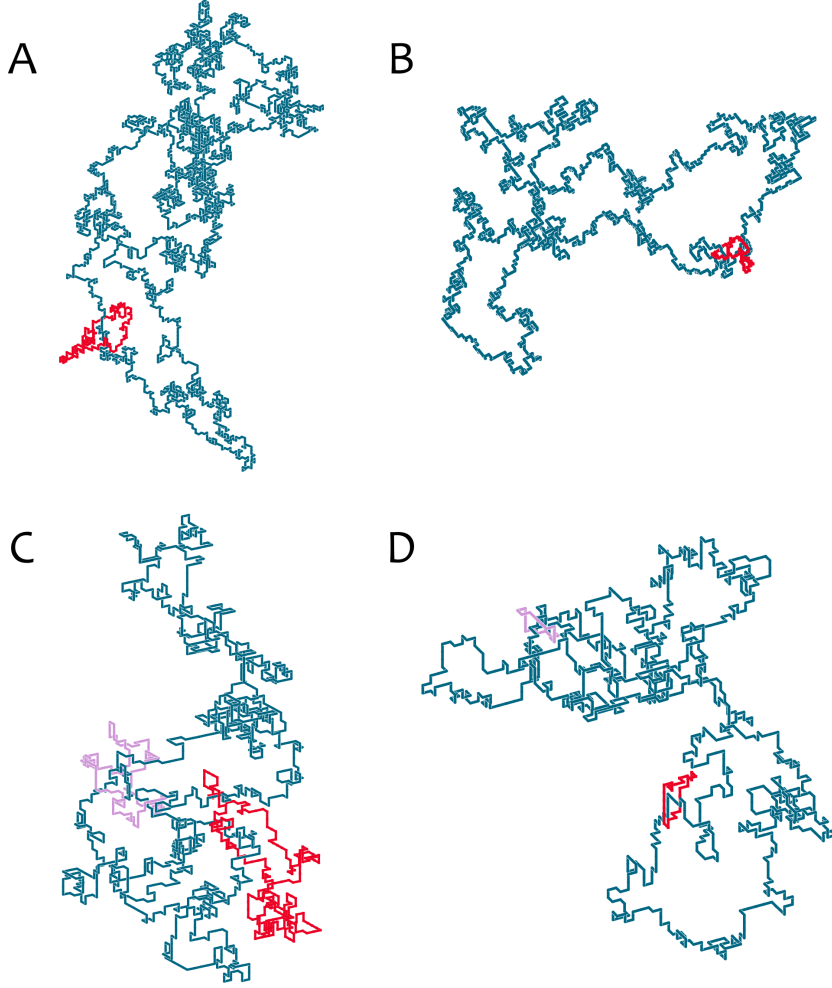


Figure 5.8: Examples of links on the cubic lattice sampled with the BFACF algorithm with $F = 0.2132$ (and $q = 3$). **A**, **B** and **C** are, respectively examples of the prime links 2_1^2 (Hopf), 5_1^2 (Whitehead) and 6_2^3 (Borromean). **D** is $2_1^2 \# 3_1 \# 2_1^2$. The blue component of $2_1^2 \# 3_1 \# 2_1^2$ is a 3_1 . It is evident that one of the components is significantly larger than the others.

We first consider the prime links 2_1^2 , 4_1^2 , 5_1^2 , 6_1^2 , 6_2^3 , 6_3^3 , 8_1^4 , 8_2^4 and 8_3^4 . In each of these cases, the rings are indistinguishable before numbering. When we examine the numbers of edges in each component as a function of n , we find that one component grows at the expense of the others, so we have spontaneous symmetry breaking. See Figure 5.9. The rings are indistinguishable so any of the rings can be the one that grows. See Figures 5.10 and 5.11. It is important to notice that the averages in Figures 5.10 and 5.11 are over a set of realizations. In some realizations one ring grows, while in another set

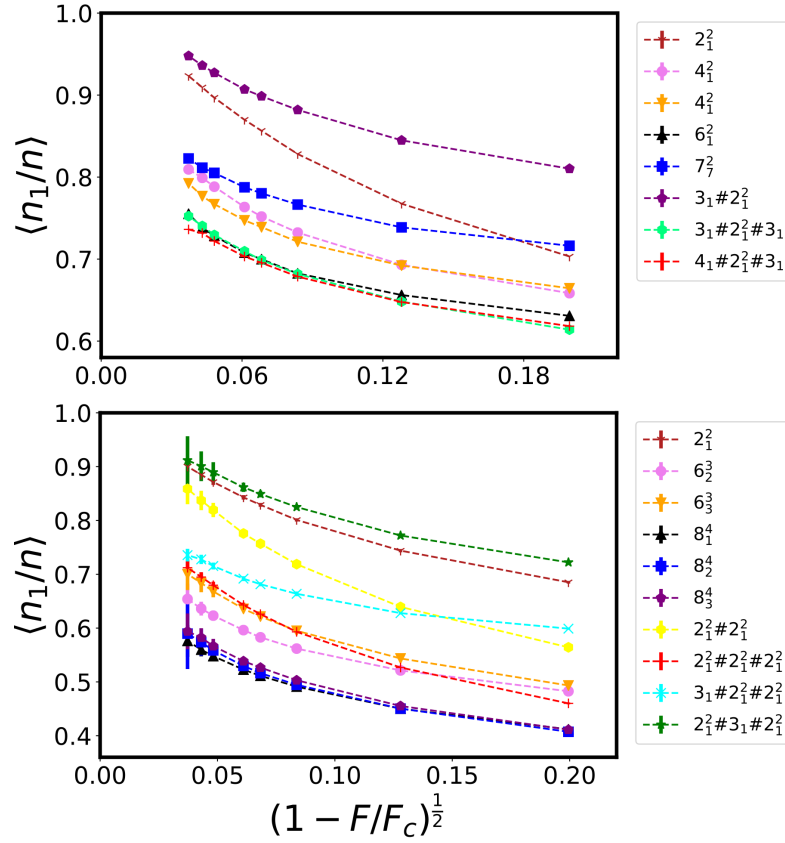


Figure 5.9: The growth of the largest ring, for various link types, as F changes. Even for the links with four components, at the largest values of F more than half the edges are in the largest ring. Bottom panel taken from Ref. [4].

a different ring grows. We note that the number of edges in the largest component increases linearly with increasing n but each of the rings is equally likely to grow.

We looked at composite links with all components unknotted, particularly the three-component link $2_1^2\#2_1^2$ and the four-component link $2_1^2\#2_1^2\#2_1^2$. In the three-component case we have one component linked to the other two components and this central component grows at the expense of the other two components. See Figures 5.13 and 5.15. For the four-component link $2_1^2\#2_1^2\#2_1^2$ one of the two central components grows at the expense of the other rings. See Figures 5.9 and 5.13. There is no reason to prefer one of the two central components over the other and we have spontaneous symmetry breaking. This is seen in Figure 5.15 where two peaks develop.

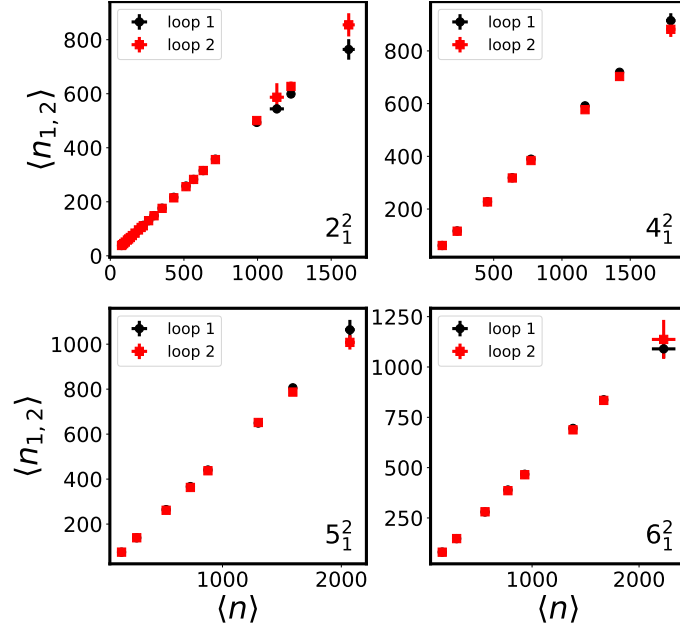


Figure 5.10: The ring 1D sizes for prime links with two unknotted components as a function of the total length of the links.

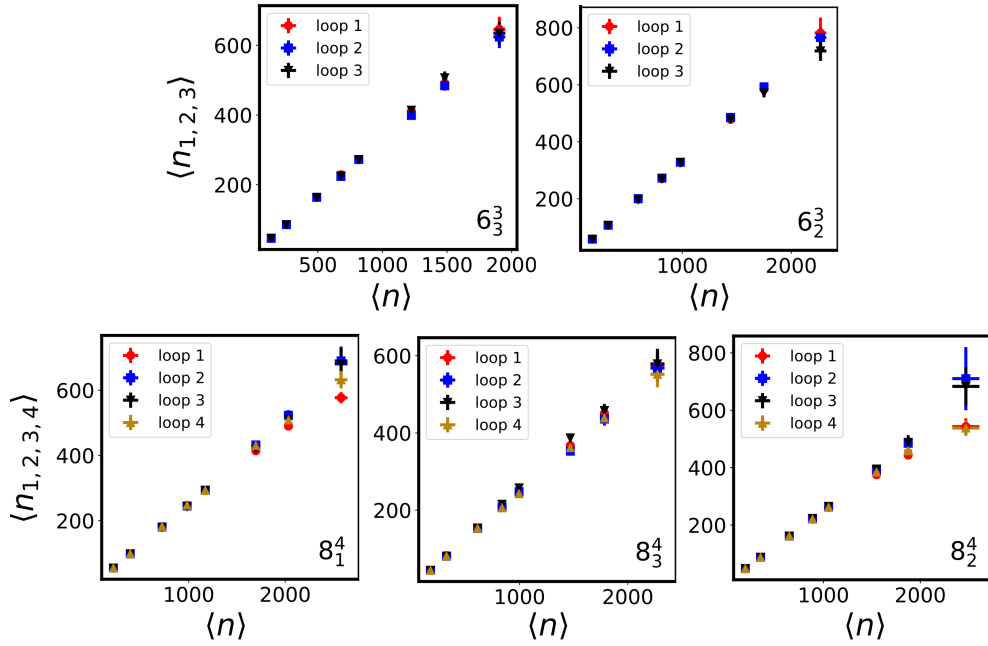


Figure 5.11: The ring 1D sizes for prime links with three and four components as a function of the total length of the links. Figure adapted from Ref. [4].

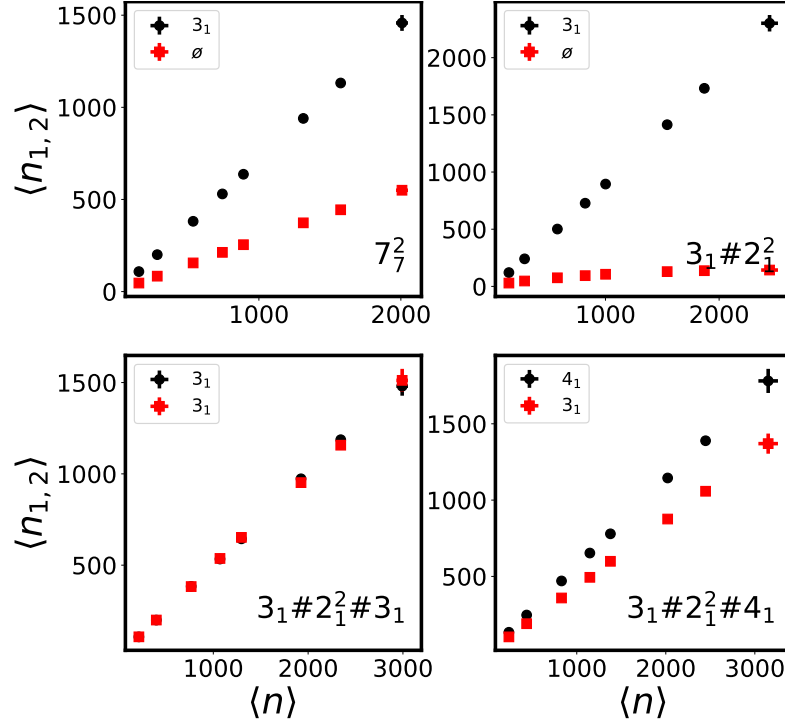


Figure 5.12: The ring 1D sizes for links with two components, at least one of which is a 3_1 or 4_1 knot as a function of the total length of the links. 7_7^2 is a prime link, $3_1 \# 2_1^2$, $3_1 \# 2_1^2 \# 3_1$ and $4_1 \# 2_1^2 \# 3_1$ are composite links.

We have also looked at examples of composite links with one or two components knotted, namely $3_1 \# 2_1^2$, $3_1 \# 2_1^2 \# 3_1$, $4_1 \# 2_1^2 \# 3_1$, $3_1 \# 2_1^2 \# 2_1^2$ and $2_1^2 \# 3_1 \# 2_1^2$. $3_1 \# 2_1^2$ is a two-component link with one component a trefoil. Only the trefoil component grows linearly with the number of edges. The prime link 7_7^2 shows similar behaviour with the trefoil growing at the expense of the unknot. This is seen in Figures 5.12 and 5.14.

$3_1 \# 2_1^2 \# 2_1^2$ and $2_1^2 \# 3_1 \# 2_1^2$ are both three-component links with one component a trefoil. In the first case one of the two outer rings is a trefoil while in the second case the central ring is a trefoil. For $3_1 \# 2_1^2 \# 2_1^2$ either the trefoil or the central ring grows roughly linearly with n while the unknotted outer ring grows more slowly. See Figures 5.13 and 5.15. For $2_1^2 \# 3_1 \# 2_1^2$ the central (knotted) ring grows roughly linearly with n while the two outer (unknotted) rings grow sub-linearly. This can be seen clearly in Figures 5.13, 5.15 and 5.8D. The two trefoils of the 2-component link $3_1 \# 2_1^2 \# 3_1$ are indistinguishable; as for the prime links with all components unknotted, only one of the two rings grows at the expense of the other and both of the rings can

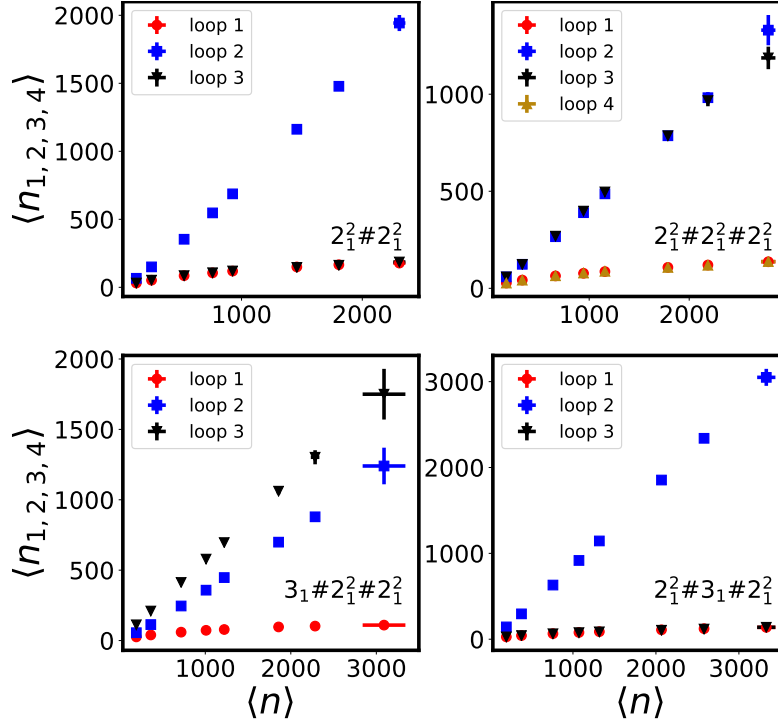


Figure 5.13: The ring 1D sizes for composite links with more than two components as a function of the total length of the links. The colouring scheme follows that used for the sketches in Fig. 5.15. For $2_1^2 \# 2_1^2$, the blue squares indicate the central component. For $2_1^2 \# 2_1^2 \# 2_1^2$, the mean length of the two central components are in blue and black. For $3_1 \# 2_1^2 \# 2_1^2$, black is used for the knotted component (3_1), blue for the central (unknotted) component. The mean length of the knotted (3_1) component of $2_1^2 \# 3_1 \# 2_1^2$ is in red. Figure taken from Ref. [4].

be the one that grows. See Figures 5.12 and 5.14. Finally, $4_1 \# 2_1^2 \# 3_1$ is a 2-component link with one components a trefoil and the other a figure eight. We discuss this case below.

Previous studies on the problem of a ring sliding through a rigid slip-link [119, 25, 132] or translocating thorough a pore [132] suggest that the partition of the total number of edges n between two components is the result of the entropic competition between the components and ultimately on whether they are knotted or not. We can understand most of our results in terms of entropy maximization. Note that here the partition is naturally induced by the link topology and not imposed by an external geometrical constraint.

For the prime links that we considered, with the exception of 7_7^2 , all the

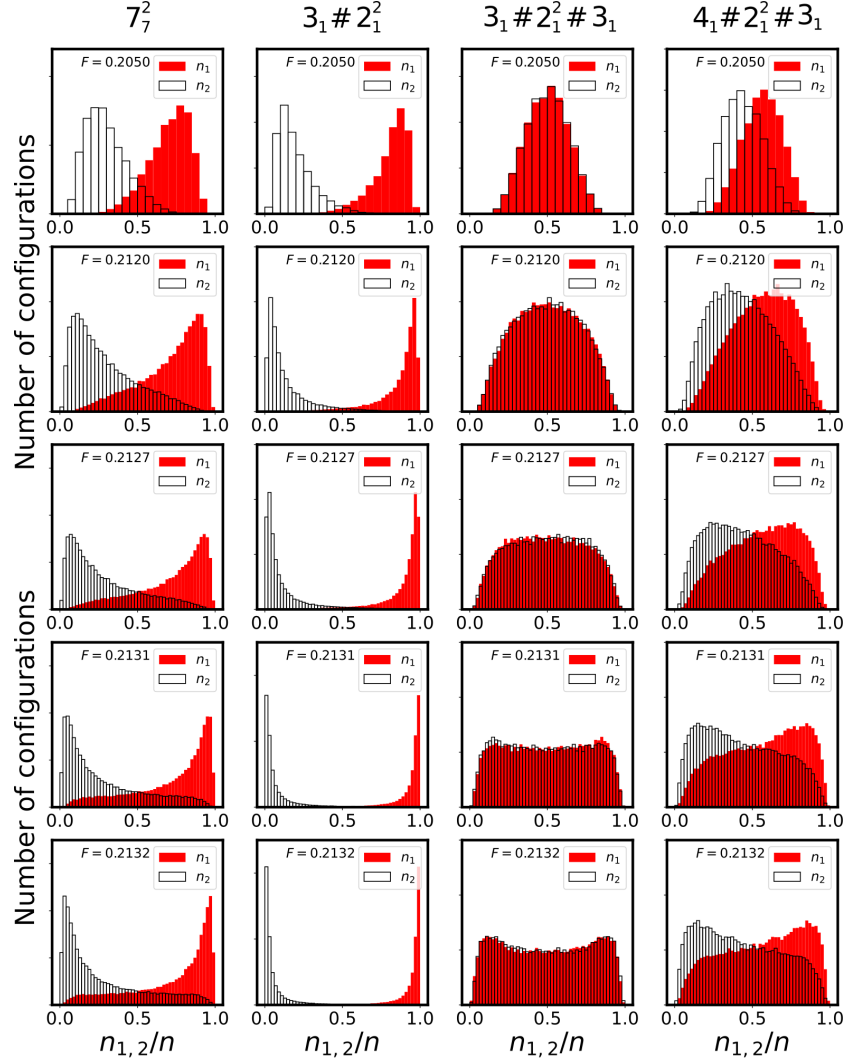


Figure 5.14: Fractions n_1/n and n_2/n for the links with two components at least one of which is knotted. Each column refers to a link type with different values of F . The five panels in each column correspond to five different values of F , $F = 0.2050$, $F = 0.2120$, $F = 0.2127$, $F = 0.2131$ and $F = 0.2132$. Figure taken from Ref. [3].

components are equivalent and the entropy is maximized by having any one component grow at the expense of the others. Since the components are equivalent, any one of them can grow. See Figures 5.10 and 5.11. In each case, the link behaves like a polygon with one root [23], consistent with the exponent estimates reported in section 5.2.1. For the three-component link $2_1^2 \# 2_1^2$, when the central ring grows it behaves like a polygon with two quasi-

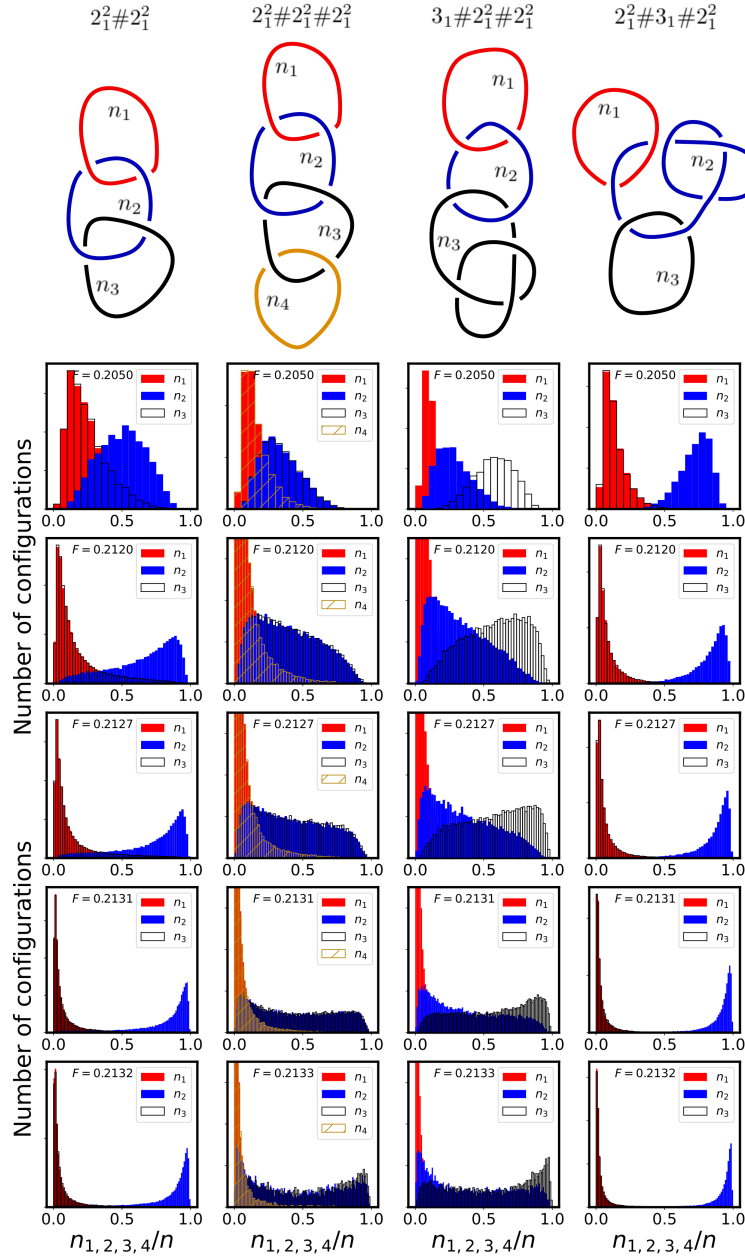


Figure 5.15: Fractions n_1/n , n_2/n , n_3/n (and n_4/n) for the composite links with three and four components considered here. Each column refers to a link type with different values of F . The five panels in each column correspond to five different values of F , $F = 0.2050$, $F = 0.2120$, $F = 0.2127$, $F = 0.2131$ and $F = 0.2132$ or 0.2133 . Figure taken from Ref. [4].

independent roots, again consistent with the exponent estimates in section 5.2.1. If either of the outer rings grew it would be behaving like a polygon with one root which has lower entropy. Similarly, one of the two central rings of the four-components link $2_1^2 \# 2_1^2 \# 2_1^2$ is expected to grow, as a polygon with two roots has larger entropy than one with a single root.

Since a knot has larger entropy than the unknot [23], we expect that the unknotted component of the link $3_1 \# 2_1^2$ is, most of the time, the smaller one. Asymptotically, we expect it to behave like a polygon with two quasi-independent roots, one being associated with the weakly localised physical knot of the trefoil [23]. The prime link 7_7^2 shows similar behaviour, with the trefoil growing at the expense of the unknot. However, the expected number of roots in this case is one, since 7_7^2 is a prime link (the knotted portion of the trefoil and the unknot are interlocked). For the knotted case $3_1 \# 2_1^2 \# 2_1^2$, if either the trefoil or the central ring grows, the system behaves as a polygon with two roots, so both these cases are more favourable than the other unknotted ring growing (we would have a polygon with one root). Again the system adapts to maximize its entropy. The growth of the ring containing the trefoil is somewhat favoured over the growth of the central ring. This is an interesting second order effect for which we do not have an explanation. It would be interesting to investigate this further. For the case $2_1^2 \# 3_1 \# 2_1^2$ the situation is simpler. When the central (knotted) ring grows, it behaves as a polygon with three roots and this maximizes the entropy of the system. This is all consistent with, and helps to explain, the exponent estimates in section 5.2.1.

If both components of a 2-component link have the same knot type, the symmetry of the link is recovered and we expect an asymptotic behaviour similar to the link 2_1^2 . On the other hand, we know that knots are weakly localized and this should complicate the approach of the system to the asymptotic limit. This is the case of the link $3_1 \# 2_1^2 \# 3_1$; as shown in Figure 5.14, third column, the bimodal shapes of both distributions of edges can just be glimpsed for the largest value of F . Once again, the entropy is maximised by having one component grow at the expense of the other; consequently, the link behaves like a polygon with two roots, one of which is associated with the weakly localised physical knot of the dominant ring. Corrections to scaling are even more severe for the link $4_1 \# 2_1^2 \# 3_1$: for intermediate values of n the system looks asymmetric with the distributions (Fig. 5.14, last column) of the first (4_1) and the second (3_1) ring developing a single peak respectively close to one and zero. We expect that the exponential growth rate and the critical exponent are the same for the trefoil and the figure eight, but they will have different amplitudes and different corrections to scaling. For the

most asymptotic value of F (rightmost panel) each distribution starts to develop a shoulder that might be a precursor of a bimodal shape. More work is required to understand why the figure-eight is favoured at intermediate values of n .

5.3 Discussion

We have investigated the asymptotic behaviour of multi-component links where the edges can be distributed among the components in all possible ways, consistent with a fixed link type. We proved two theorems about their exponential growth rate and used a Monte Carlo approach, based on the BFACF algorithm, to investigate the values of the critical exponent and how it depends on link type. The exponent values are consistent with one ring growing at the expense of the others while the remaining rings effectively act as one or more roots on the growing ring. Each root increases the exponent by 1. Which ring grows is determined by which maximizes the entropy of the system.

Chapter 6

Conclusions

Entanglement abundance and complexity strongly affect the biological and mechanical function of polymer systems. In the lifecycle of living organisms, genome regulation is essential to avoid nefarious consequences, and the fine tuning of the topological properties of synthetic polymer systems is essential to target certain desired rheological properties. Nevertheless, how DNA entanglement is regulated *in vivo* is still unclear, and the characteristics of many topologically constrained polymer systems, due to their difficult practical and theoretical treatment, are currently unknown. In this work we begin to address this gap of knowledge by modelling the proliferation and regulation of genome entanglement and by investigating the properties of linked ring polymers.

In chapter 3, by using theory and simulations, we examined a solution of ring polymers undergoing recombination inside a sphere. We showed that the system harbours a transition or crossover from a “short rings” regime, which steady states are a mixture of short and unlinked rings polymers, to a “long ring” regime, characterised by the self-assembly of a network of long linked loops. This transition can be triggered either by increasing the stiffness of the polymers or by decreasing the confinement radius, and, by analogy with the gelation transition for sticky colloids, was labelled *topological gelation*. We argued that, geometrically, topological gelation is explained by the competition between the bending energy and the configurational entropy of the recombinant rings, while the topological properties of the gel phase are a consequence of the typical volume occupied by the rings exceeding the overlapping point. Since the topological gelation of DNA would likely be highly detrimental for living cells, our results highlight the importance of regulating the action of recombination enzymes according to the stiffness and volume fraction of the

DNA substrate. How this can be practically achieved, however, requires further investigation; we argue, for instance, that the topological simplification mechanism proposed in chapter 4, enacted by SMC proteins, might be able, to some extent, to balance out the topological activity of recombination. On the other hand, as emphasized, for instance, by studying the escape dynamics from a permeabilised sphere of the recombinant rings, the topological gel might also be exploited to assemble DNA plasmids with desired topologies. Since it is known that topological gels possess toughness far superior to that of traditional chemical gels, it would be interesting to study the properties of DNA gels assembled by recombination. On this note, analysing the spectra of topologies which emerge close to the gelation point would be of great importance, since, as emphasised in chapter 5, ring polymers linked with different link types behave, generally speaking, differently.

In chapter 4, we suggested a pathway to efficiently resolve entanglement in DNA and other polymer systems. Specifically, by using scaling analysis and computer simulations, we inspected the interplay between slip-links and knots. First we showed that the topology of two competing slip-linked polymer networks determines how they exchange monomers, suggesting a potential strategy to fine tune the geometric details of the system. We then showed that a set of slip-links is able to quickly confine a knot within a small portion of a polymer, and that the efficiency of the localisation is hampered on stiffer substrates and regulated by the topology of the slip-linked network. We then used this result to prove that entropic competition with slip-links can swiftly resolve entanglement by transient strand-crossing. In the contest of DNA, this finding informs of a potential mechanism of topological simplification enacted by SMC proteins and topoisomerase. Since how SMC proteins extrude DNA loops has not been completely uncovered yet, checking how effective this mechanism would be in the wider genome picture (e.g. recombinant gel) might help understanding if and in which contest diffusive loop extrusion is sufficient and when an active (i.e. with external energy input) mechanism is instead required. More generally, the ability to partition a polymer in a controlled, but reversible way, by using desired network topologies, might be useful in designing devices such as molecular switches. How to practically implement this, however, requires further investigation.

Finally, in chapter 5, we studied the configurational entropy of ring polymers linked with fixed topology, where the monomers distribute among the components in all possible ways. First we proved some rigorous results about the exponential growth rate of their number of configurations. We then implemented a Monte Carlo method, based on the BFACF algorithm, to sample linked polygons on the cubic lattice, and proved that its ergodic classes are

the link types. We proceeded to estimate the growth rate and entropic exponents for various topologies and showed that one of the components of a link grows at the expense of the others, and as such, asymptotically, the link looks like an unknot crowned with one or more sliding roots, each contributing +1 to the entropic exponent. The number of roots is determined by the number of components and the presence of knots, and it obeys entropy maximisation. Although here we have taken a step forward in understanding the entropic properties of polymer links, the applicability of our results is mostly restricted to melts of living (or recombinant) polymers and, indirectly, to slip-linked polymer networks. Since in many circumstances real ring polymers cannot freely exchange monomers (e.g DNA catenanes formed during replication), studying the scaling behaviour of a system of polygons with fixed number of edges linked with a given fixed topology would be equally interesting. As can be deduced from chapter 5, for instance, this would shed light on how the size of the physically linked portion of two linked ring polymers grows with the number of edges of the components, information which we argue might be of substantial importance in the contest of entanglement regulation in polymer systems.

Bibliography

- [1] A. Bonato, D. Marenduzzo, D. Michieletto, and E. Orlandini. Topological gelation of reconnecting polymers. *Proceedings of the National Academy of Sciences*, 119(44):e2207728119, 2022.
- [2] A. Bonato, D. Marenduzzo, and D. Michieletto. Simplifying topological entanglements by entropic competition of slip-links. *Physical Review Research*, 3(4):043070, 2021.
- [3] A. Bonato, E. Orlandini, and S .G. Whittington. Asymptotics of linked polygons. *Journal of Physics A: Mathematical and Theoretical*, 53(38):385002, 2020.
- [4] A. Bonato, E. Orlandini, and S. G. Whittington. Asymptotics of multicomponent linked polygons. *Journal of Physics A: Mathematical and Theoretical*, 54(23):235002, 2021.
- [5] J. Arsuaga, M. Vázquez, S. Trigueros, D. W. Sumners, and J. Roca. Knotting probability of DNA molecules confined in restricted volumes: DNA knotting in phage capsids. *Proc. Natl. Acad. Sci. USA*, 99(8):5373–7, apr 2002.
- [6] H. L. Frisch and E. Wasserman. Chemical topology1. *Journal of the American Chemical Society*, 83(18):3789–3795, 1961.
- [7] M. Delbruck. Chemical topology1. *Journal of the American Chemical Society*, 83(18):3789–3795, 1961.
- [8] D. W. Sumners and S. G. Whittington. Knots in self-avoiding walks. *Journal of Physics A: Mathematical and General*, 21(7):1689–1694, apr 1988.
- [9] N. Pippenger. Knots in random walks. *Discrete Applied Mathematics*, 25(3):273–278, 1989.

- [10] D. E. Adams, E. M. Shekhtman, E. L. Zechiedrich, M. B. Schmid, and N. R. Cozzarelli. The role of topoisomerase iv in partitioning bacterial replicons and the structure of catenated intermediates in dna replication. *Cell*, 71(2):277–288, 1992.
- [11] S. DiNardo, K. Voelkel, and Sternglanz R. Dna topoisomerase ii mutant of *saccharomyces cerevisiae*: topoisomerase ii is required for segregation of daughter molecules at the termination of dna replication. *Proceedings of the National Academy of Sciences*, 81(9):2616–2620, 1984.
- [12] L. Postow, N. J. Crisona, B. J. Peter, C. D. Hardy, and N. R. Cozzarelli. Topological challenges to DNA replication: conformations at the fork. *Proc. Nat. Acad. Sci. USA*, 98(15):8219–26, 2001.
- [13] G. Jannink, B. Duplantier, and J. L. Sikorav. Forces on chromosomal DNA during anaphase. *Biophys. J.*, 71(July):451–465, 1996.
- [14] E. J. van Rijssel, J. B. Trimpos, and M. H. Booster. Mechanical performance of square knots and sliding knots in surgery: comparative study. *Am. J. Obstet. Gynecol.*, 162(January):93–97, 1990.
- [15] Thomas J. Nicholls, Cristina A. Nadalutti, Elisa Motori, Ewen W. Sommerville, Gráinne S. Gorman, Swaraj Basu, Emily Hoberg, Doug M. Turnbull, Patrick F. Chinnery, Nils-Göran Larsson, Erik Larsson, Maria Falkenberg, Robert W. Taylor, Jack D. Griffith, and Claes M. Gustafsson. Topoisomerase 3 α is required for decatenation and segregation of human mtdna. *Molecular Cell*, 69(1):9–23.e6, 2018.
- [16] Mila Ilić, Irene C Zaalberg, Jonne A Raaijmakers, and René H Medema. Life of double minutes: generation, maintenance, and elimination. *Chromosoma*, pages 1–19, 2022.
- [17] Stephen C.Y. Ip, Migena Bregu, François Xavier Barre, and David J. Sherratt. Decatenation of DNA circles by FtsK-dependent Xer site-specific recombination. *EMBO Journal*, 22(23):6399–6407, 2003.
- [18] D. Marenduzzo, C. Micheletti, E. Orlandini, and D. W. Sumners. Topological friction strongly affects viral dna ejection. *Proceedings of the National Academy of Sciences*, 110(50):20081–20086, 2013.
- [19] S. F. Edwards. The statistical mechanics of polymerized material. *Proceedings of the Physical Society*, 9(92), 1967.

- [20] V. Narsimhan, C. B. Renner, and P. S. Doyle. Translocation dynamics of knotted polymers under a constant or periodic external field. *Soft Matter*, 12:5041–5049, 2016.
- [21] S. A. Wasserman and N. R. Cozzarelli. Supercoiled DNA-directed knotting by T4 topoisomerase. *Journal of Biological Chemistry*, 266(30):20567–20573, 1991.
- [22] P. G. de Gennes. *Scaling concepts in polymer physics*. Cornell University Press, 1979.
- [23] E Orlandini, MC Tesi, EJ Janse van Rensburg, and Stuart G Whittington. Entropic exponents of lattice polygons with specified knot type. *Journal of Physics A: Mathematical and General*, 29(12):L299, 1996.
- [24] R. Metzler, A. Hanke, P. G. Dommersnes, Y. Kantor, and M. Kardar. Equilibrium shapes of flat knots. *Phys. Rev. Lett.*, 88:188101, Apr 2002.
- [25] B. Marcone, E. Orlandini, A. L. Stella, and F. Zonta. What is the length of a knot in a polymer? *Journal of Physics A: Mathematical and General*, 38(1):L15–L21, dec 2004.
- [26] F. Ferrari. A new strategy to microscopic modeling of topological entanglement in polymers based on field theory. *Nuclear Physics B*, 948:114778, 2019.
- [27] B. D. Hughes et al. *Random walks and random environments: random walks*, volume 1. Oxford University Press, 1995.
- [28] T. Hara and G. Slade. Self-avoiding walk in five or more dimensions i. the critical behaviour. *Commun. Math. Phys.*, 147:101–136, 1992.
- [29] P. J. Flory. *Principles of polymer chemistry*. Cornell university press, 1953.
- [30] B. Duplantier. Polymer chains in four dimensions. *Nuclear Physics B*, 275(2):319–355, 1986.
- [31] P. Grassberger, R. Hegger, and Schafer L. Self-avoiding walks in four dimensions: logarithmic corrections. *Journal of Physics A: Mathematical and General*, 27(22):7265–7282, nov 1994.
- [32] J Des Cloizeaux. Polymers in solutions: principles and applications of a direct renormalization method. *Journal de Physique*, 42(5):635–652, 1981.

- [33] S. F. Edwards. The statistical mechanics of polymers with excluded volume. *Proceedings of the Physical Society*, 85(4):613, 1965.
- [34] N. Clisby. Accurate estimate of the critical exponent ν for self-avoiding walks via a fast implementation of the pivot algorithm. *Phys. Rev. Lett.*, 104:055702, Feb 2010.
- [35] H. P. Hsu, W. Nadler, and P. Grassberger. Scaling of star polymers with 1- 80 arms. *Macromolecules*, 37(12):4658–4663, 2004.
- [36] B. Duplantier. Statistical mechanics of polymer networks of any topology. *Journal of Statistical Physics*, 54(3):581–680, 1989.
- [37] J. Batoulis and K. Kremer. Thermodynamic properties of star polymers: good solvents. *Macromolecules*, 22(11):4277–4285, 1989.
- [38] Renzo L Ricca and Bernardo Nipoti. Gauss’ linking number revisited. *Journal of Knot Theory and Its Ramifications*, 20(10):1325–1343, 2011.
- [39] C. E. Soteros, D. W. Sumners, and S. G. Whittington. Entanglement complexity of graphs in z_3 . In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 111, pages 75–91. Cambridge University Press, 1992.
- [40] B. Li, N. Madras, and A. D. Sokal. Critical exponents, hyperscaling, and universal amplitude ratios for two-and three-dimensional self-avoiding walks. *Journal of Statistical Physics*, 80(3):661–754, 1995.
- [41] K. L. Chung. *Markov Chains with Stationary Transition Probabilities*. Springer, 1967.
- [42] Alberto Berretti and Alan D Sokal. New monte carlo method for the self-avoiding walk. *Journal of Statistical Physics*, 40(3):483–531, 1985.
- [43] N. Madras and A. D. Sokal. Nonergodicity of local, length-conserving monte carlo algorithms for the self-avoiding walk. *Journal of statistical physics*, 47(3):573–595, 1987.
- [44] C. J. Geyer. *Computing Science and Statistics: Proc. 23rd Symp. on the Interface*. Interface Foundation of North America, 1991.
- [45] E. J. Jense van Rensburg and S. G. Whittington. The bfacf algorithm and knotted polygons. *Journal of Physics A: Mathematical and General*, 24(23):5553, 1991.

- [46] B. Berg and D. Foerster. Random paths and random surfaces on a digital computer. *Physics Letters B*, 106(4):323–326, 1981.
- [47] C. Aragao De Carvalho and S. Caracciolo. A new monte-carlo approach to the critical properties of self-avoiding random walks. *Journal de Physique*, 44(3):323–331, 1983.
- [48] C. Aragao De Carvalho, S. Caracciolo, and J. Fröhlich. Polymers and φ^4 theory in four dimensions. *Nuclear Physics B*, 215(2):209–248, 1983.
- [49] L. H. Kauffmans. Knot diagrammatics. *Handbook of knot theory*, pages 233–318, 2005.
- [50] D. Rolfsen. Knots and links. mathematical lecture series. 7. berkeley, ca.: Publish or perish, 1976.
- [51] N. Madras and G. Slade. *The self-avoiding walk*. Springer Science & Business Media, 2013.
- [52] S. D. Bond and B. J. Leimkuhler. Molecular dynamics and the accuracy of numerically computed averages. *Acta Numerica*, 16:1–65, 2007.
- [53] W. Hayes and K. R. Jackson. A survey of shadowing methods for numerical solutions of ordinary differential equations. *Applied Numerical Mathematics*, 53(2-4):299–321, 2005.
- [54] D. Donnelly and E. Rogers. Symplectic integrators: An introduction. *American Journal of Physics*, 73(10):938–945, 2005.
- [55] L. Verlet. Computer” experiments” on classical fluids. i. thermodynamical properties of lennard-jones molecules. *Physical review*, 159(1):98, 1967.
- [56] T. Schneider and E. Stoll. Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions. *Physical Review B*, 17(3):1302, 1978.
- [57] Steve Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of computational physics*, 117(1):1–19, 1995.
- [58] v W. Kuhn and F. Grun. Relations between elastic constants and the strain birefringence of high-elastic substances. *Kolloid Zh*, 101(3):248–271, 1942.

- [59] Harold R. Warner Jr. Kinetic theory and rheology of dilute suspensions of finitely extendible dumbbells. *Industrial & Engineering Chemistry Fundamentals*, 11(3):379–387, 1972.
- [60] T. A. Hunt and B. D. Todd. A comparison of model linear chain molecules with constrained and flexible bond lengths under planar Couette and extensional flows. *Molecular Simulation*, 35(14):1153–1167, 2009.
- [61] G. S. Grest and K. Kremer. Molecular dynamics simulation for polymers in the presence of a heat bath. *Phys. Rev. A*, 33:3628–3631, May 1986.
- [62] J. E. Jones. On the determination of molecular fields.—i. from the variation of the viscosity of a gas with temperature. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 106(738):441–462, 1924.
- [63] J. Z. Zhang, X. Y. Peng, S. Liu, B. P. Jiang, S. C. Ji, and X. C. Shen. The persistence length of semiflexible polymers in lattice monte carlo simulations. *Polymers*, 11(2):295, 2019.
- [64] O. Kratky and G. Porod. Röntgenuntersuchung gelöster fadenmoleküle. *Recueil des Travaux Chimiques des Pays-Bas*, 68(12):1106–1122, 1949.
- [65] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comp. Phys. Comm.*, 271:108171, 2022.
- [66] A. D. Bates and A. Maxwell. *DNA topology*. Oxford University Press, 2005.
- [67] D. J. Sherratt, L. K. Arciszewska, G. Blakely, S. Colloms, K. Grant, N. Leslie, and R. McCulloch. Site-specific recombination and circular chromosome segregation. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 347(1319):37–42, 1995.
- [68] P. D. Sadowski. Site-specific genetic recombination: hops, flips, and flops. *The FASEB Journal*, 7(9):760–767, 1993.

- [69] M. Vazquez, S. D. Colloms, and D. W. Sumners. Tangle analysis of Xer recombination reveals only three solutions, all consistent with a single three-dimensional topological pathway. *Journal of Molecular Biology*, 346(2):493–504, 2005.
- [70] I. Grainge, M. Bregu, M. Vazquez, V. Sivanathan, S. C. Y. Ip, and D. J. Sherratt. Unlinking chromosome catenanes in vivo by site-specific recombination. *EMBO Journal*, 26(19):4228–4238, 2007.
- [71] D. W. Sumners, C. Ernst, S. J. Spengler, and N. R. Cozzarelli. Analysis of the mechanism of DNA recombination using tangles, 1995.
- [72] R. G. Petschek, P. Pfeuty, and J. C. Wheeler. Equilibrium polymerization of chains and rings: A bicritical phenomenon. *Phys. Rev. A*, 34(3):2391, 1986.
- [73] L. Ringrose, V. Lounnas, L. Ehrlich, F. Buchholz, R. Wade, and A. F. Stewart. Comparative kinetic analysis of flp and cre recombinases: mathematical models for dna binding and recombination. *Journal of molecular biology*, 284(2):363–384, 1998.
- [74] M. E. Cates. Nonlinear viscoelasticity of wormlike micelles (and other reversibly breakable polymers). *Journal of Physical Chemistry*, 94(1):371–375, 1990.
- [75] R. J. Hommel and M. D. Graham. Constitutive modeling of dilute wormlike micelle solutions: Shear-induced structure and transient dynamics. *J. Non-Newton. Fluid Mech.*, 295:104606, 2021.
- [76] SF Edwards and KF Freed. The entropy of a confined polymer. I. *Journal of physics A: General physics*, 2:145–150, 1969.
- [77] D Marenduzzo. Statistics of confined polymers and the melting of a dna spool. *EPL (Europhysics Letters)*, 85(3):38005, 2009.
- [78] T. Cremer and C. Cremer. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat. Rev. Genet.*, 2(4):292–301, apr 2001.
- [79] P. R. Cook and D. Marenduzzo. Entropic organization of interphase chromosomes. *J. Cell. Biol.*, 186(6):825–34, sep 2009.
- [80] M. Falk, Y. Feodorova, N. Naumova, M. Imakaev, B. R. Lajoie, H. Leonhardt, B. Joffe, J. Dekker, G. Fudenberg, I. Solovei, and L. A.

- Mirny. Heterochromatin drives compartmentalization of inverted and conventional nuclei. *Nature*, 570(7761):395–399, 2019.
- [81] M. Chiang, D. Michieletto, C. A. Brackley, N. Rattana-virotkul, H. Mohammed, D. Marenduzzo, and T. Chandra. Polymer Modeling Predicts Chromosome Reorganization in Senescence. *Cell Reports*, 28(12):3212–3223.e6, 2019.
- [82] David H Van Winkle and CA Murray. Layering in colloidal fluids near a smooth repulsive wall. *J. Chem. Phys.*, 89(6):3885–3891, 1988.
- [83] G. B. Brandani, M. Schor, C. E. MacPhee, H. Grubmüller, U. Zachariae, and D. Marenduzzo. Quantifying Disorder through Conditional Entropy: An Application to Fluid Mixing. *PLoS ONE*, 8(6):e65617, 2013.
- [84] Kurt W. Kohn. Principles and practice of dna filter elution. *Pharmacology & Therapeutics*, 49(1):55–77, 1991.
- [85] Pavel Hozák, A Bassim Hassan, Dean A Jackson, and Peter R Cook. Visualization of replication factories attached to a nucleoskeleton. *Cell*, 73(2):361–373, 1993.
- [86] Guido Polles, Davide Marenduzzo, Enzo Orlandini, and Cristian Micheletti. Self-assembling knots of controlled topology by designing the geometry of patchy templates. *Nat. Commun.*, 6:6423, 2015.
- [87] J. Chen, C. A. Rauch, J. H. White, P. T. Englund, and N. R. Cozzarelli. The topology of the kinetoplast DNA network. *Cell*, 80(1):61–9, jan 1995.
- [88] S. Wasserman, J. Dungan, and N. Cozzarelli. Discovery of a predicted DNA knot substantiates a model for site-specific recombination. *Science (80-.)*, 229(4709):171–174, 1985.
- [89] V. P. Patil, J. D. Sandt, M. Kolle, and J. Dunkel. Topological mechanics of knots and tangles. *Science*, 367(6473):71–75, 2020.
- [90] J. M. Sogo, A. Stasiak, M. L. Martínez-Robles, D. B. Krimer, P. Hernández, and J. B. Schwartzman. Formation of knots in partially replicated DNA molecules. *J. Mol. Biol.*, 286(3):637–43, feb 1999.
- [91] J. C. Wang. DNA topoisomerases. *Annu. Rev. Biochem.*, 54:665–97, jan 1985.

- [92] M. Doi and S. F. Edwards. *The theory of polymer dynamics*. Oxford University Press, 1988.
- [93] R. C. Ball, M. Doi, S. F. Edwards, and M. Warner. Elasticity of entangled networks. *Polymer*, 22:1010, 1981.
- [94] S. F. Edwards and T. Vilgis. The effect of entanglements in rubber elasticity. *Polymer*, 27(4):483–492, 1986.
- [95] P. G. Higgs and R. C. Ball. Trapped entanglements in rubbers a unification of models. *Epl*, 8(4):357–361, 1989.
- [96] D. Michieletto. On the tree-like structure of rings in dense solutions. *Soft Matter*, 12:9485–9500, 2016.
- [97] A. E. Likhtman. The tube axis and entanglements in polymer melts. *Soft Matter*, 10(12):1895–1904, 2014.
- [98] J. K. Ryu, C. Bouchoux, H. W. Liu, E. Kim, M. Minamino, R. de Groot, A. J. Katan, A. Bonato, D. Marenduzzo, D. Michieletto, F. Uhlmann, and C. Dekker. Phase separation induced by cohesin SMC protein complexes. *Sci. Adv.*, 7(7):1–10, 2021.
- [99] G. Fudenberg, M. Imakaev, C. Lu, A. Goloborodko, N. Abdennur, and L. A. Mirny. Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep.*, 15:2038–2049, 2016.
- [100] X. Wang, H. B. Brandão, T. B. K. Le, M. T. Laub, and D. Z. Rudner. Travel From Origin To Terminus. *Science*, 527(February):524–527, 2017.
- [101] M. Ganji, I. A. Shaltiel, S. Bisht, E. Kim, A. Kalichava, C. H. Haering, and C. Dekker. Real-time imaging of DNA loop extrusion by condensin. *Science (80-.)*, 360(6384):102–105, 2018.
- [102] I. F. Davidson, B. Bauer, D. Goetz, W. Tang, G. Wutz, and J. M. Peters. DNA loop extrusion by human cohesin. *Science*, 366(December):1338–1345, 2019.
- [103] C. A. Brackley, J. Johnson, D. Michieletto, A. N. Morozov, M. Nicodemi, P. R. Cook, and D. Marenduzzo. Nonequilibrium Chromosome Looping via Molecular Slip Links. *Phys. Rev. Lett.*, 119(13):138101, 2017.

- [104] I. F. Davidson, D. Goetz, M. P. Zaczek, M. I. Molodtsov, P. J. Huis in 't Veld, F. Weissmann, G. Litos, D. A. Cisneros, M. Ocampo-Hafalla, R. Ladurner, F. Uhlmann, A. Vaziri, and J. M. Peters. Rapid movement and transcriptional re-localization of human cohesin on DNA. *EMBO J.*, 35(24):2671–2685, 2016.
- [105] J. Stigler, G. Çamdere, D. E. Koshland, and E. C. Greene. Single-Molecule Imaging Reveals a Collapsed Conformational State for DNA-Bound Cohesin. *Cell Rep.*, 15(5):988–998, may 2016.
- [106] K. Nasmyth. Disseminating the Genome: Joining, Resolving, and Separating Sister Chromatids During Mitosis and Meiosis. *Annu Rev Gen*, 35(1):673–745, 2001.
- [107] M. Hirano and T. Hirano. Hinge-mediated dimerization of SMC protein is essential for its dynamic interaction with DNA. *EMBO J.*, 21(21):5733–5744, 2002.
- [108] J. H. Gibcus, K. Samejima, A. Goloborodko, I. Samejima, N. Naumova, J. Nuebler, M. T. Kanemaki, L. Xie, J. R. Paulson, W. C. Earnshaw, L. A. Mirny, and J. Dekker. A pathway for mitotic chromosome formation. *Science (80-.)*, 359(6376), 2018.
- [109] A. Goloborodko, J. F. Marko, and L. A. Mirny. Chromosome Compaction by Active Loop Extrusion. *Biophys. J.*, 110(10):2162–2168, 2016.
- [110] E. Piskadlo and R. A. Oliveira. A topology-centric view on mitotic chromosome architecture. *Int. J. Mol. Sci.*, 18(12):1–21, 2017.
- [111] E. Orlandini, D. Marenduzzo, and D. Michieletto. Synergy of topoisomerase and structural-maintenance-of-chromosomes proteins creates a universal pathway to simplify genome topology. *Proc. Natl. Acad. Sci.*, 116(17):8149–8154, 2019.
- [112] Y. Inoue, M. Miyauchi, H. Nakajima, Y. Takashima, H. Yamaguchi, and A. Harada. Self-threading of a poly(ethylene glycol) chain in a cyclodextrin-ring: control of the exchange dynamics by chain length. *J. Am. Chem. Soc.*, 128(28):8994–5, jul 2006.
- [113] R. S. Forgan, J. P. Sauvage, and J. F. Stoddart. Chemical topology: Complex molecular knots, links, and entanglements. *Chem. Re*, 111(9):5434–5464, 2011.

- [114] K. Ito. Novel Cross-Linking Concept of Polymer Network: Synthesis, Structure, and Properties of Slide-Ring Gels with Freely Movable Junctions. *Polym. J.*, 39(6):489–499, 2007.
- [115] A. Bin Imran, K. Esaki, H. Gotoh, T. Seki, K. Ito, Y. Sakai, and Y. Takeoka. Extremely stretchable thermosensitive hydrogels by introducing slide-ring polyrotaxane cross-linkers and ionic groups into the polymer network. *Nature Communications*, 5:1–8, 2014.
- [116] Y. Yasuda, M. Toda, K. Mayumi, H. Yokoyama, H. Morita, and K. Ito. Sliding Dynamics of Ring on Polymer in Rotaxane: A Coarse-Grained Molecular Dynamics Simulation Study. *Macromolecules*, 52:3787–3793, 2019.
- [117] J. Rieger. Unusual bifurcation in a constrained random walk system. *J. Phys. A: Math. Gen.*, 21:L1085, 1988.
- [118] J. U. Sommer. Degenerate ground states of simple slip-link systems . *J. Chem. Phys.*, 97:5777, 1992.
- [119] R. Zhandi, Y. Kantor, and M. Kardar. Entropic competition between knots and slip-links. *ARI The Bulletin of the ITU*, 53:6, 2003.
- [120] A. Bonato, C. A. Brackley, J. Johnson, D. Michieletto, and D. Marenduzzo. Chromosome Compaction and Chromatin Stiffness Enhance Diffusive Loop Extrusion by Slip-Link Proteins. *Soft Matter*, 16:2406–2414, 2020.
- [121] L. Vian, A. Pekowska, S. S. P. Rao, D. Levens, E. Lieberman Aiden, R. Casellas, L. Vian, K. R. Kieffer-Kwon, S. Jung, L. Baranello, S. C. Huang, L. El Khattabi, M. Dose, N. Pruett, A. L. Sanborn, A. Canela, Y. Maman, A. Oksanen, W. Resch, X. Li, B. Lee, A. L. Kovalchuk, Z. Tang, S. Nelson, M. Di Pierro, R. R. Cheng, I. Machol, B. Glenn St Hilaire, N. C. Durand, M. S. Shamim, E. K. Stamenova, J. N. Onuchic, Y. Ruan, and A. Nussenzweig. The Energetics and Physiological Impact of Cohesin Extrusion. *Cell*, 173:1–14, 2018.
- [122] L. Uusküla-Reimand, H. Hou, P. Samavarchi-Tehrani, M. V. Rudan, M. Liang, A. Medina-Rivera, H. Mohammed, D. Schmidt, P. Schwalie, E. J. Young, J. Reimand, S. Hadjur, A. C. Gingras, and M. D. Wilson. Topoisomerase II beta interacts with cohesin and CTCF at topological domain borders. *Genome Biol*, 17(1):1–22, 2016.

- [123] C. Micheletti, D. Marenduzzo, E. Orlandini, and D. W. Summers. Knotting of random ring polymers in confined spaces. *Journal of Chemical Physics*, 124(6):064903, 2006.
- [124] M. Nakahata, S. Mori, Y. Takashima, H. Yamaguchi, and A. Harada. Self-Healing Materials Formed by Cross-Linked Polyrotaxanes with Reversible Bonds. *Chem*, 1(5):766–775, 2016.
- [125] A. V. Vologodskiĭ, A. V. Lukashin, and M. D. Frank-Kamenetskiĭ. Topological interaction between polymer chains. *Soviet Journal of Experimental and Theoretical Physics*, 40:932, 1975.
- [126] B. Duplantier. Linking numbers of closed manifolds at random in $\mathbb{R}^n$, inductances and contacts. *Communications in Mathematical Physics*, 85(2):221–264, 1982.
- [127] E. Orlandini, E. J. Janse Van Rensburg, M. C. Tesi, and S. G. Whittington. Random linking of lattice polygons. *Journal of Physics A: Mathematical and General*, 27(2):335, 1994.
- [128] M. Atapour, C. E. Soteros, C. Ernst, and S. G. Whittington. The linking probability for 2-component links which span a lattice tube. *Journal of Knot Theory and Its Ramifications*, 19(01):27–54, 2010.
- [129] J. M. Hammersley. The number of polygons on a lattice. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 57, pages 516–523. Cambridge University Press, 1961.
- [130] C. E. Soteros, D. W. Sumners, and S. G. Whittington. Linking of random p-spheres in $\mathbb{Z}^d$. *Journal of Knot Theory and Its Ramifications*, 8(01):49–70, 1999.
- [131] A. Rechnitzer. *Unpublished*, 2020.
- [132] M. Baiesi, E. Orlandini, A. L. Stella, and F. Zonta. Topological signature of globular polymers. *Phys. Rev. Lett.*, 106:258301, 2011.