SISSA

PAPER: CLASSICAL STATISTICAL MECHANICS, EQUILIBRIUM AND NON-EQUILIBRIUM

Bubble-bound state of triple-stranded DNA: Efimov physics in DNA with repulsion

To cite this article: Jaya Maji *et al J. Stat. Mech.* (2017) 073203

View the article online for updates and enhancements.

Bubble-bound state of triple-stranded DNA: Efimov physics in DNA with repulsion

Jaya Maji[1,4]**, Flavio Seno**[2]**, Antonio Trovato**[2] **and Somendra M Bhattacharjee**[3]

[1] Department of Physics, Indian Institute of Science Education and Research, Bhopal 462 066, India
[2] INFN, Dipartimento di Fisica e Astronomia 'Galileo Galilei', Università di Padova, Via Marzolo 8, 35131 Padova, Italy
[3] Institute of Physics, Bhubaneswar 751005, India
[4] Present address: Raja Peary Mohan College, Uttarparha, Hoogly 712258, India

E-mail: jmaji.phys@rpmcollege.org, flavio.seno@pd.infn.it, antonio.trovato@pd.infn.it and somen@iopb.res.in

Abstract. The presence of a thermodynamic phase of a three-stranded DNA, namely, a mixed phase of bubbles of two bound strands and a single one, is established for large dimensions ($d \geqslant 5$) by using exact real space renormalization group transformations and exact computations of specific heat for finite length chains. Similar exact computations for the fractal Sierpinski gasket of dimension $d < 2$ establish the stability of the phase in the presence of a repulsive three chain interaction. Although, for $d < 2$, cooperativity factors for bubbles or noncrossing conditions are needed for the melting transition, the mixed phase may exist even in absence of those. In contrast to the Efimov DNA, where three strands are bound though no two are bound, the mixed phase appears at temperatures less than the two chain melting temperature. Both the Efimov-DNA and the mixed phase are formed essentially due to the strand exchange mechanism.

Keywords: biopolymers, renormalisation group

Contents

1. Introduction

DNA serves as a primary unit of heredity and contains all the information necessary for the living systems. This complex object is made of two complementary strands via Watson–Crick base pairing. One important milestone of modern biology was the discovery of DNA double-helical structure [1]. The subsequent discovery of the triple helix has attracted great attention in medical sciences in view of its possible use in mapping chromosomes [2], inhibition of gene expressions [3, 4], gene therapy and gene targeting [5], inducing site-specific mutations (targeted mutagenesis) [6], interfering with DNA replication [7, 8] and other applications [9–11]. Triple helix is formed by a duplex binding with a single strand DNA, or RNA, or PNA, via Hoogsteen or reverse Hoogsteen hydrogen bonding for DNA [12, 13], or even by three RNA's [14, 15].

Recent theoretical studies identified two different types of triple-stranded DNA (tsDNA) states near the duplex melting point. We may recall that melting of DNA is the phenomenon of the temperature induced separation of double-stranded DNA into single strands. One of the two states of tsDNA is a loosely bound state of three strands when no two are bound [16–21], and the other one is a phase of bubbles of pairwise bound and a single strand below the melting point. Both the states are stabilized by strand exchange of a single strand, and, in a sense, produced by the bubble fluctuations

near the duplex melting point. The former one, occurring at or above the duplex melting point on the unbound side, resembles the well-known Efimov effect [22, 23] in quantum mechanics, and is called the *Efimov-DNA*. It is not a pure phase, but a continuation of the three chain bound state. In contrast, the other phase, which we may call a bubble-bound or a mixed phase, occurs on the bound side of a duplex, at temperatures below the duplex melting point. It was found as a genuine phase in low dimensional models with specially tuned interactions, and it differs from the Efimov-DNA as the chains are locally pairwise bound, but without any direct three chain contact.

Our aim in this paper is to explore the possibility of the bubble-bound phase, also called a mixed phase, in triple-stranded DNA in a wider context. We would call this a bubble-bound state as a single strand is always accompanied by a bubble of duplex DNA. This paper establishes the existence of the mixed phase for large dimensions, with $d = 5$ as an example. As already mentioned, this phase was predicted from fractal lattice studies with $d < 2$ [18]. We also study the role of three chain repulsion in these low dimensional system.

The Efimov effect originates from quantum fluctuations at the zero energy bound states of two attracting nonrelativistic particles. In a three particle system with pairwise critical potential, an effective long range $1/r^2$ attraction appears between two particles at distance r due to the wide excursions of the third particle in the classically forbidden regions [22–24]. Consequently, there are an infinite number of bound states. The original continuous scale invariance in the two particle problem changes to a discrete scale invariance characterizing the bound states. Recent experiments have more or less established the quantum Efimov effect in ultracold atoms, though the existence of the infinite number of states at the threshold is yet to be established [25–27][5]. The triplex bound state of DNA at the melting point of duplex DNA is the thermal analogue of the Efimov physics where the denatured bubbles play the role of classically forbidden paths in quantum mechanics. A polymer scaling analysis, using hyperscaling at the dsDNA melting point, recovers the inverse-square interaction very naturally. Such arguments show the importance of the fluctuations of bubbles near the melting point and the polymer correlations along the single strands of the bubbles [16]. Since then, the Efimov physics has been explored in various other systems, like Efimov-driven transition in many body systems [28], in quantum magnets [29], and in one dimensional systems under long range interactions [30]. Several studies have probed the importance of dimer-atom states near Efimov resonance [31], and possible topological origins of the Efimov physics [32]. In this scenario, triple-stranded DNA appears as a classical testing ground for Efimov physics, which can also be enriched by other thermal effects and relevant interactions. In fact, the polymer scaling analysis [16] is one of the simplest ways to see the emergence of the long range interaction at the heart of the Efimov physics. The bubble-bound phase we study in this paper is an analogue of the atom-dimer phase in the quantum problem. In the polymer context, such states appear as a stable thermodynamic phase.

---

[5] There are two aspects of the Efimov effect, the formation of the three particle bound state in the regime where two should not be bound, and the Efimov tower at the two-particle threshold. Of these, the tower is a consequence of the inverse-square interaction, but a bound state would be possible even for r^{-p} interactions with $p < 2$. We may add that in the context of DNA, Coulomb electrostatic repulsion is not important as it is screened out in ionic solvents.

To study the DNA problem in various dimensions, we adopt the model of three-stranded DNA on diamond hierarchical lattices and Sierpinski gaskets. These lattices are constructed in an algorithmic way by iterative replacement of bonds by a basic motif at each level. The discrete scaling symmetry makes these lattices affordable, in contrast to the regular lattices, for exact calculations or for renormalization group studies of different models [33–42]. Beside the understanding of complex physical system, hierarchical lattices also exhibit a lot of interesting mathematics. The method of construction of the chosen lattices allows one to express the partition functions as recursion relations which either can further be iterated for physical properties in the large lattice limit, or can be used to decimate for renormalization group (RG) flows and fixed points (fp). RG is considered a valuable tool to understand the emergent behavior of systems with diverging length scales. By integrating out the small length scale fluctuations and rescaling, the same system is regenerated but with renormalized parameters. The consequent RG transformations lead to fixed points and separatrices. The fixed points represent states of the system where it shows scale invariance under rescaling of lengths. A given system need not be at the fixed points but from the flow patterns of the parameters, one identifies the transitions and the phases of the system. Moreover, the transitions can be corroborated by finite size scaling analysis of finite length thermodynamic properties, like specific heat or energy, which can be computed exactly in a recursive manner.

In low dimensions, DNA melting does not occur unless a bubble weight factor σ is introduced or noncrossing conditions are imposed [18, 21]. With such weight factors σ, called cooperativity factors, assigned at junctions of bound and unbound states (Y-fork), the Efimov state and the mixed phase were studied in different classes of DNA models on the Sierpinski Gasket of dimension $d_f = \ln 3/\ln 2$. A detailed study of the Efimov state, using the fixed size transfer matrix approach, was done in [21] for $1 + 1$ dimensional Euclidean lattice without any three chain interaction. Here also the cooperativity factor was needed for a melting transition, but the occurrence of the Efimov state turned out to be sensitive to how the weights are assigned in a three chain system. A model of three noncrossing chains in $1 + 1$ dimensions, without any cooperativity factor but with a three chain interaction, was solved exactly in [43]. The phase diagrams obtained in [43] resemble partly the phase diagrams reported here, but without any mixed phase. In this paper we focus on the mixed phase which occurs on the bound side of the two chain melting. As a continuation of the higher dimensional studies to $d < 2$, we explored the possibilities of the existence of the bubble bound mixed phase in presence of a three chain repulsion for DNA models defined on the Sierpinski Gasket. Unlike $d > 2$ cases, here, thanks to σ, a quantitative criterion for the Efimov-DNA can be constructed. This is one of the important results of this section.

The paper is organized as follows. In section 2 we consider a few simplified polymer models to study the Efimov state and mixed phase of DNA in a lattice of dimension $d > 2$ using the renormalization group approach. The existence of a phase transition from the bubble-bound phase to the triple bound state is also shown by exact specific heat computations. In section 3 we discuss the model on a Sierpinski gasket of dimension $d < 2$ by the method of exact calculation. Here we extend the analysis of [18] by including a three body interaction. In section 4 we conclude that the mixed phase is well established from both the lattice models.

2. Model: $d > 2$

The native base pairing interaction of a DNA is best expressed as directed polymers on a lattice. Each monomer of the strands represents a collection of bases interacting with the monomer of same sequence index of the other strand as per the Poland–Scherega scheme.

The melting of thee-stranded DNA has been studied by real space renormalization group which can be implemented exactly on hierarchical lattices of dimensions $d > 2$ [16, 17, 38, 39, 44]. The procedure to construct the lattice is to start with a single bond, at generation $n = 0$, and then replace the bond by a diamond motif at $n = 1$ (see figure 1).

The dimension of the lattice is defined by $d = \ln(\lambda b)/\ln \lambda$, where λ is the length scale factor and b is the branching factor connecting the bottom and the top of the lattice as in figure 1. Parameter b can be tuned to change the dimension of the lattice. In our models, $\lambda = 2$.

Let us consider three directed polymers being laid from the bottom to the top of the lattice with no restriction on intersections and crossings. See figure 2. Two weight factors are needed for the DNA problem, viz., $y(= e^{\beta\epsilon})$ associated with two polymers sharing the same bond with energy $-\epsilon$ and $w(= e^{-\beta\eta})$, associated with three polymers sharing the same bond with an additional triplet energy η. The triplet interaction between monomers is attractive if $\eta < 0$, i.e. $w > 1$ or repulsive if $\eta > 0$, i.e. $w < 1$. Here $\beta = 1/T$ is the inverse temperature (with the Boltzmann constant $k_B = 1$), and all the pairwise interactions are taken to be the same.

Since temperature T is absorbed in y and w, for given interaction energies among the chains we get a temperature curve in the y-w plane, parameterized by T. As a result the T dependence of a given set of chains, i.e. for fixed ϵ, η, can be obtained from the intersection of such a curve with the phase transition lines in the y-w plane. In view of this, we keep y and w to be independent variables. A combination variable $X = wy^2$ is also found to be useful, as explained below.

2.1. Renormalization group equations

The RG recursion relations for the weights of the DNA are given by [16, 17] (see appendix)

$$y' = \frac{b - 1 + y^2}{b}, \tag{1a}$$

$$w' = \frac{(b-1)(b-2) + 3(b-1)y^2 + w^2 y^6}{b^2 y'^3}. \tag{1b}$$

These are the RG relations with w', y' denoting the renormalized values. The important point is that the renormalization of the two chain interaction is not affected by the three chain interaction.

The melting point of a duplex DNA is described by the unstable fixed point (fp) of equation (1a) at $y_c = b - 1$ while the high temperature unbound phase is given by the stable fixed point at $y = 1 (T = \infty)$. These two chain fp's lead to different possibilities for w as
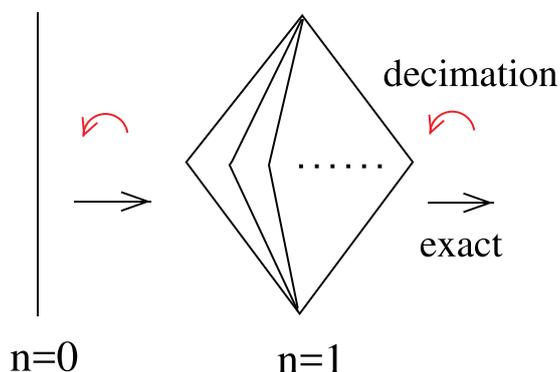
Figure 1. Schematic diagram of an hierarchical lattice. At each stage a bond is replaced by a basic motif (shown for $n = 1$). The left arrow denotes the direction of decimation of RG.
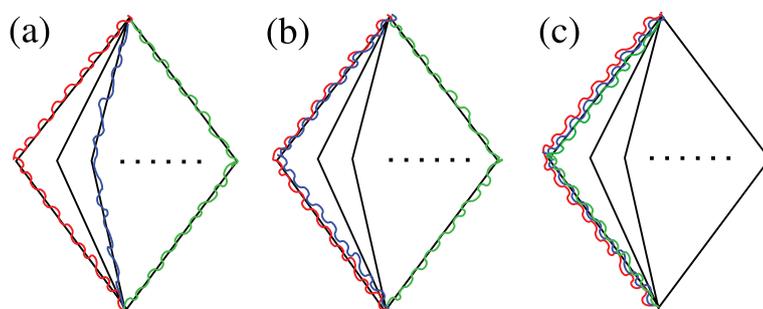

Figure 2. Different configurations of a triple-stranded DNA are shown by the wavy lines on a d dimensional lattice for general b. (a) Polymers are separate on the path and the total number of configurations is $b(b-1)(b-2)$. (b) Two polymers among three are together and the total number is $3b(b-1)$. (c) Three polymers are together and the total number is b.

$$
w^* = \begin{cases} 1, b^2 - 1, \infty & \text{for} \quad y = 1 \\ w_-, w_+, \infty & \text{for} \quad y = y_c \equiv b - 1, \end{cases} \tag{2a}
$$

with

$$
w_\pm = \frac{b^2 \pm \sqrt{4 - 24b + 32b^2 - 12b^3 + b^4}}{2(b-1)^3}, \tag{2b}
$$

being real for $b > b_c = 8.56...$. The dimension corresponding to this special b_c is $d_c = \ln(2b_c)/\ln 2$. Of these, the fp at $w_c = b^2 - 1$ for $y = 1$ is the three chain melting point due to a pure three body interaction. These fixed points for w were discussed previously in connection with the Efimov-DNA. We focus here on a new set of fp's that are found for low temperatures.

2.2. Zero temperature fixed points: a dilemma

In addition to the above-mentioned fp's, there is a stable fp at $y = \infty$ that describes the bound state of the duplex at zero temperature. In this limit, the RG equation, equation (1b), for w leads to a dichotomy, depending on how the zero temperature is

reached. For a straightforward $y \to \infty$ limit, there is an unstable fp at $w_\infty^* = b^{-1}$, separating the two stable fp's at 0 and ∞. As $b > 1$, $w_\infty^* < 1$; it is in the repulsive region. It might appear to represent a dissociation of a three chain bound state into a single strand and a bound pair at zero temperature because of large three chain repulsion. It is to be noted that $w = 0$ is like hard core repulsion, preventing overlap of three strands.

Curiously, the recursion relations also allow a different set of fixed points for the relative weight

$$X = wy^2.$$

In the $y \to \infty$ limit, the RG relation from equation (1b) becomes

$$X' = \frac{3(b-1)}{b} + \frac{1}{b}X^2, \quad (y \to \infty), \tag{3}$$

which has two fixed points

$$X_\pm = \frac{b}{2} \pm \frac{1}{2}\sqrt{b^2 - 12(b-1)}. \tag{4}$$

These fixed points are real for $b \geqslant b_X = 2(3+\sqrt{6}) = 10.898\,9794..$, or $b < 2(3-\sqrt{6}) = 1.1...$. Of these, the latter one is not meaningful and therefore not considered in this work, as $b > 2$ is required for melting.

A given system follows a path $w = y^{\eta/\epsilon}$ as T is varied, with $w \to 0$ or ∞ as $T \to 0$, depending on the sign of η. Therefore, the fixed point $w_\infty^* = 1/b$ does not play any role. We instead focus on the zero temperature fixed points for X.

2.2.1. Justification of X. To see why X is important, when the three chain weight is wy^3, let us calculate the energy of the states. If n_2, n_3 represent respectively pure two chain, and three chain (mutually exclusive) contacts per unit length, then $n_2 + n_3 \leqslant 1$. The total energy per unit length of DNA is

$$\frac{E}{N} = -n_2\epsilon + n_3(\eta - 3\epsilon). \tag{5}$$

On minimization,

$$\left.\frac{E}{N}\right|_{\min} = \begin{cases} \eta - 3\epsilon & \text{if } \eta < 2\epsilon \quad \text{for } n_2 = 0, n_3 = 1, \\ -\epsilon & \text{if } \eta > 2\epsilon \quad \text{for } n_2 = 1, n_3 = 0. \end{cases} \tag{6}$$

The zero temperature transition occurs when $\eta = 2\epsilon$ with the energy parameters as the variables. The high $\eta(> 0)$ phase consists of bubbles made of single chain and bound duplex, with nonzero entropy. Naively, if ΔS is the low temperature entropy difference per bond of these two phases, one may combine it with equation (6) to determine the free energy difference, $\Delta F = \Delta E - T\Delta S$. The continuity of the free energies at the transition, i.e. $\Delta F = 0$ then gives the transition temperature as $wy^2 \sim \exp(-|\Delta S|)$.

The above argument works for a first order transition as we see from equation (6). However local bubble formation in the triplex bound state at low temperatures softens the system. The ground state energy is independent of dimensions. As a result, there are two mutually exclusive possibilities, viz., (i) no transition, or (ii) a continuous transition. In absence of a good estimate of the bubble entropy, this simple argument does

not give a clue about the critical dimension for the transition ($d > d_X \approx 4.46$, where $d_X = \ln(2b_X)/\ln 2$, see below), but, in any case, it justifies the emergence of the combination variable $X = wy^2$ in the low temperature region, as noted in section 2.2.

X is relevant for a phase transition from a three chain bound state to a state of bubbles and single chain, while the three chain Boltzmann weight wy^3 plays a role in melting phenomenon of the Efimov state. The bound-mixed transition is not a straight forward peeling transition where the three chain bound state may split into a pair and a single one. If it were so, the transition would be equivalent to a two chain melting case because at very low temperatures a pair would be more-or-less bound acting like a single flexible polymer (in this model). The equivalent pairing interaction would be then $X = wy^2$, yielding a peeling-off temperature $X = b - 1$. This is not the case, as we see from equation (4). A comparison of equations (3) with (1a) shows the difference. The factor of 3 in the RG equation for X vis-a-vis equation (1a) for y shows the extra entropic contribution coming from strand exchange. This extra entropy makes the bubble-bound phase a unique phase for three-stranded DNA.

2.3. Phase diagrams for large *b*: $y > y_c$

2.3.1. From RG: separatrix. The existence of real fixed points of X for $y \to \infty$, allows us to draw the phase diagrams for $b > b_X$. This range of b corresponds to $d > 4.446....$ Based on equations (2a) and (4), we have three different situations. (i) $b < b_c = 8.56...$, (ii) $b_c < b < b_X$, and (iii) $b > b_X$. In case (i), there are no fixed points for X for $y \geqslant y_c$. The separatrix for the flow pattern in the y-w plane does not correspond to any transition since the three chain bound state is the only thermodynamic phase. Even though there is a pair of fixed point for $y = y_c$ in case (ii), still there is only the three chain bound phase. Note that the fixed point for w at $w_\infty^* = 1/b$ goes to infinity when expressed in terms of X.

The situation is different for $b > b_X$. The RG equations written in terms of X and $z = 1/y$ are

$$z' = \frac{bz^2}{(b-1)z^2 + 1} \approx bz^2 + O(z^4), \tag{7a}$$

$$X' = \frac{X^2 + 3(b-1) + (b-1)(b-2)z^2}{b[(b-1)z^2 + 1]}, \tag{7b}$$

which can be used to study the region $z < z_c = y_c^{-1} = (b-1)^{-1}$. In this low temperature regime, the unstable fixed points at $z = z_c$ and $z = 0$ are connected by a separatrix as shown in figure 3(a) for $b = 16$. In this case of $b = 16$, the fixed points are

$$X_\pm^{(c)} = 13.7184, 3.348\,31, \quad \text{at } y = y_c, \tag{8a}$$

$$X_\pm = 12.3589, 3.6411, \quad \text{at } y \to \infty. \tag{8b}$$

The global stable point X_- represents the mixed phase. The local slopes of the separatrix at the fixed points are consistent with the eigen-directions of the linearized versions of equations (7a) and (7b) around the fixed points. E.g. the horizontal tangent at X_+ follows from the absence of any linear term in equation (7a) for small z. The flow

along the separatrix is towards the zero temperature (or infinite y) fixed point. That the separatrix represents the phase boundary is established below by an exact computation for $b = 16$ which corresponds to $d = 5$.

It is interesting to note that under strong repulsion, the critical state at $y = y_c$ is represented by a stable fixed point.

The behavior of specific heat as measured by the derivative of free energy with respect to w, is controlled by the fixed point X_+ at $z = 0$. The relevant exponents are [17]

$$\nu = \frac{\ln 2}{\ln \frac{\mathrm{d}X'}{\mathrm{d}X}\big|_{X \to X_+}} = \frac{\ln 2}{\ln \left[1 + \sqrt{1 - 12 \frac{(b-1)}{b^2}} \right]}, \tag{9a}$$

$$\alpha = 2 - \nu, \tag{9b}$$

where ν describes the divergence of an appropriate length scale, while α describes the specific heat as $c_w \sim |X - X_m(z)|^{-\alpha}$, $X_m(z)$ being the transition point. The specific heat here is defined as

$$c_w = -\frac{1}{L_n} \, w \frac{\partial}{\partial w} w \frac{\partial \ln Q}{\partial w}, \tag{10}$$

keeping z constant. By construction, c_w is related to the fluctuations in the number of three chain contacts. Equation (9b) shows that there is a continuous transition for all $d > 4.46$ but no divergence in specific heat for $b < 13.4$, i.e. $d < 4.74$.

2.3.2. Exact computation of c_w. The specific heat or the fluctuations in the number of three contacts can be computed exactly for finite lengths in a recursive scheme. The recursion relations for the partition functions are given by

$$C_{n+1} = b^2 C_n, \tag{11a}$$

$$Z_{n+1}(y) = b Z_n(y)^2 + b(b-1)C_n^4, \tag{11b}$$

$$Q_{n+1} = b Q_n^2 + 3b(b-1)Z_n^2 C_n^2 \\ + b(b-1)(b-2)C_n^6, \tag{11c}$$

where C_n, Z_n, Q_n are the nth generation partition functions for single, double and triple strand cases. The initial conditions are $C_0 = 1, Z_0 = y, Q_0 = wy^3$. One may also write down the recursion relations for the derivatives. To be noted here that the three chain interaction w does not affect Z, and therefore the recursion relation for c_w (equation (10)) is simpler than other derivatives.

By iterating the recursion relations, c_w has been computed exactly up to $n = 27$. See figure 3(b) where c_w is plotted as a function of X for $y = 20$. The strong growth with size is an indication of a diverging specific heat. A finite size scaling form suggests that all these data points can be collapsed on to a single curve if plotted as $c_w L_n^{-p}$ vs $(X - X_m)L_n^q$, with $p = \alpha/\nu, q = 1/\nu$, where ν and α are given by equations (9a) and (9b). Here, $L_n = 2^n$. In this way $X_m(z)$ can be estimated by using the data collapse measure of [45]. The data collapse is shown in figure 3(c). The transition point sits nicely on the separatrix in figure 3(a).
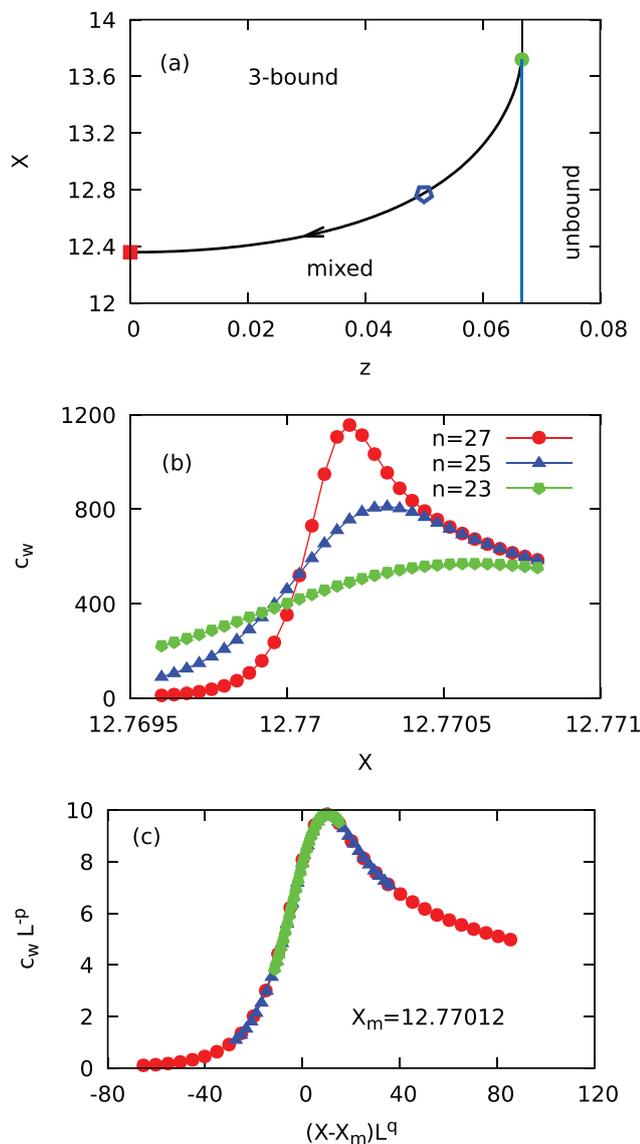
Figure 3. (a) The separatrix in the X-z plane, where $X = wy^2$ and $z = 1/y$. This is for $d = 5$ which corresponds to $b = 16$. The two chain melting point is $y_c = 15$. The relevant region shown is for $0 \leqslant z \leqslant z_c \equiv 1/y_c$. The solid curve is the numerically determined separatrix connecting the unstable fixed points at $z = z_c$ (green filled circle) and at $z = 0$ (red filled square). The open pentagon (blue) represents the transition point obtained from the specific heat data in (b). (b) Exact specific heat c_w for fixed $y = 20$, as a function of X for three different lengths of chains, $L_n = 2^n$ with $n = 23, 25, 27$. The size dependence indicates a diverging specific heat, and therefore, a continuous transition. (c) Finite size scaling: same data as in (b) but plotted with scaled variables, with $p = \alpha/\nu, q = 1/\nu$ and $X_m = 12.770\,12$. This value of X_m at $z = 0.05$ is shown by the open pentagon in (a).

2.4. Summary for $d > 2$

The end result is that for large $d > d_X$, there is now a new phase, the bubble-bound phase or the mixed phase, at low temperatures in presence of three chain repulsion. A possible form of this three-stranded DNA with pairwise bound but without three chain contacts is shown by the schematic diagram in figure 4. In the mixed phase two are bound with one free over a certain length scale of the chain, but the strand exchange mechanism leaves no one free completely. This phase undergoes a continuous transition to a completely bound state, where the two chain attraction overcomes the weak local three chain repulsion.

For easy reference we show all possible states schematically in figure 5 for various dimensions. Three typical cases shown are (a) $b = 8$ (i.e. $d = 4$), (b) $b = 9$ ($d = 4.1699$), and (c) $b = 16$ ($d = 5$). Figure 5(a) is similar to the $b = 4$ ($d = 3$) case of [17]. The phase diagram displays the Efimov region and the mixed phase in addition to the conventional bound and the unbound states. The Efimov-DNA has already been defined as the three chain bound state in situations where two chains would not have formed a bound state. The region in phase diagram where the Efimov-DNA may occur is called the Efimov region. For $w = 1$, this region will be for $y \leqslant y_c$ up to the melting point of the triplex. A three-chain bound state is possible in the region $w > w_c$ even for $y = 1$, i.e. in absence of any pair interaction ($\epsilon = 0$). Combining these, the Efimov region is defined as the region in the 3-chain bound phase, where two chains would not be bound or the three not bound by w alone. In other words, this is a state where we would not see a bound state if either a chain is removed, or there is no two body attraction (i.e. $\epsilon = 0$). The traditional Efimov case corresponds to $w = 1$, $y < y_c$. However, there is no well defined thermodynamic boundary for the Efimov region. Therefore drawing any boundary for the Efimov-DNA is subjective (see section 3.3 for a criterion applicable for $d < 2$). It is fair to say that in the w-y plane, the Efimov region lies inside the domain of the triplex phase with $w < w_c$ (pure three chain melting at $y = 1$) and $y < y_c$ (duplex melting). Such regions enclosing the Efimov states are marked as 'Efimov' in figure 5.

The mixed phase appears under the thin curve in figure 5(c) for $b > b_X$. The transition line between the unbound and the Efimov state is first order. The vertical melting line for the mixed to unbound phase and the mixed to bound phase transitions are both continuous. The transition from the mixed bubble-bound to the bound state is associated with diverging specific heat for high enough d.

We see that when there is an unstable fp at $y = y_c$, the separatrix connecting $(y = y_c, X = X_+^{(c)})$ and $(y = 1, X = w_c)$ defines the Efimov line (represented by the thick line). Thus, $d = 4$ is very special (figure 5(a)) among the three cases shown, where an Efimov-DNA occurs for any w at $y = y_c$. In the intermediate range of dimensions, say $4.0976... < d < 4.446...$, there is no mixed phase. The phase diagram (figure 5(b)) resembles figure 5(c) but without the mixed phase and the curved transition line. Curiouser and curiouser here is the difference in the melting behavior of the three chain bound state for $X > X_+^{(c)}$ and $X < X_+^{(c)}$. For $X > X_+^{(c)}$, the three chain bound state melts via the Efimov line, a first order melting [17], but for $X < X_+^{(c)}$, there is a continuous melting identical to the two chain melting problem.

Let us list the lower critical dimensions (d_{lc}) for the various phases (or transitions) we see in DNA:
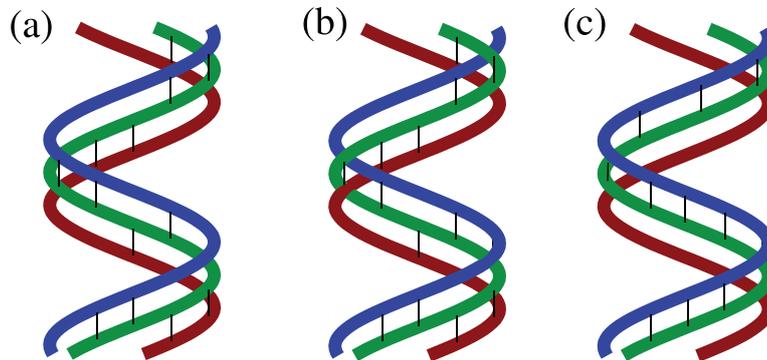
Figure 4. Schematic diagram of the mixed phase of three polymers. Per unit length of the chain two monomers are in contact leaving the third free. The pair interaction is shown by the vertical bonds. (a) Polymers cannot cross each other. (b) Polymers can cross each other. (c) Polymers cannot cross each other. Two bound and one free strands.

1. dsDNA melting: $d_{\mathrm{lc}} = 2$.
2. Pure tsDNA melting: $d_{\mathrm{lc}} = 1$.
3. Melting of Efimov-DNA: $d_{\mathrm{lc}} = 2$.
4. Existence of bubble-bound mixed phase: $d_{\mathrm{lc}} = d_X \approx 4.446$ (see below equation (4)).
5. Critical melting of tsDNA (under repulsion): $d_{\mathrm{lc}} = d_c \approx 4.0976$ (see below equation (2b)).

3. Model: $d < 2$

The mixed phase was observed in studies on low dimensional models [18], namely the Sierpinski gasket of fractal dimension $d = \ln 3 / \ln 2 < 2$. The melting in these low dimensional models is induced either by a noncrossing condition or by an extra weight for pair breaking, σ, called the cooperativity factor. This extra weight for pair-breaking (or better called Y-fork) allows the bubble entropy to be competitive with the bound state energy. It (σ) can be present in the noncrossing case too. The mixed phase was observed in several cases in the σ-y phase diagram. Our purpose is now to study the stability of the mixed phase with respect to the three chain repulsion. To appreciate its relevance, we mention two features, (i) the possibility of studying low d, and (ii) the possibility of investigating the role of σ. In fact, σ allows us to define an unambiguous criterion for the Efimov-DNA, resolving a difficulty faced for the $d > 2$ cases. By tuning σ to zero, bubble formation can be forbidden. As bubbles play a crucial role in Efimov physics, we may now compare the phase diagram of a nonzero σ with that of $\sigma = 0$. The region inbetween will be the Efimov region.

The Sierpinski gasket is constructed recursively. See figure 6 for generations $n = 1,\ 2$. The construction starts from a unit triangle, copies of which are glued together in the next generation as shown in the figure. This process is repeated iteratively in a self-similar way with the motif generated at the nth step to form a bigger lattice at the next step. As the size increases, so do the holes, making the lattice an object of $d < 2$.
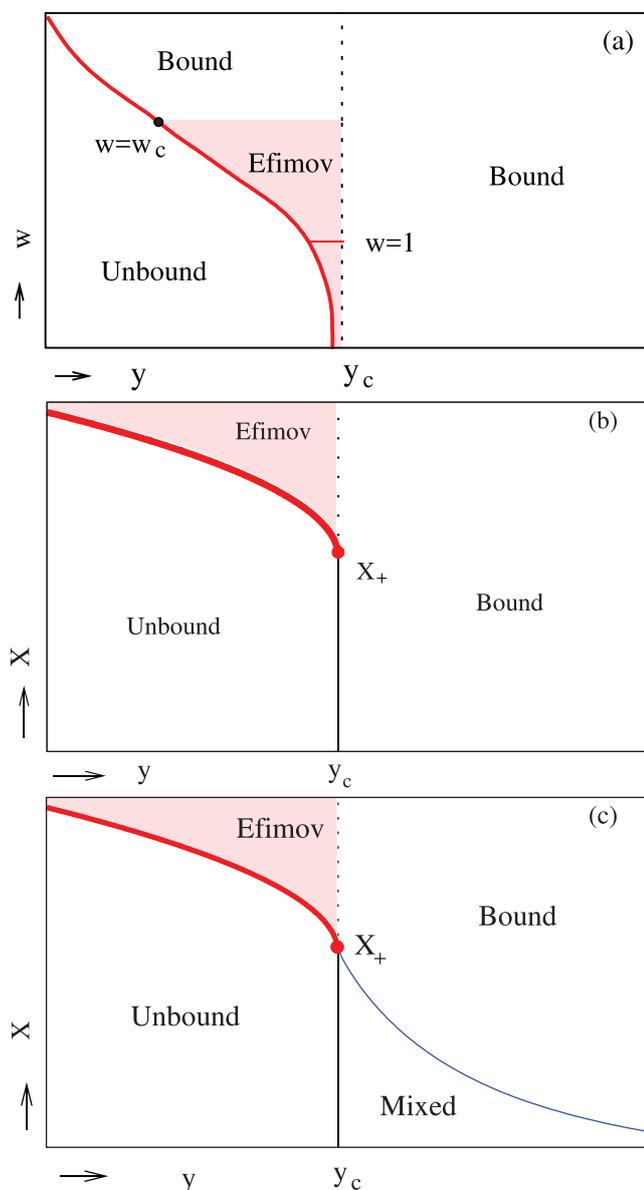
Figure 5. Schematic phase diagrams in the w-y or the X-y plane. (a) For $b = 8\,(d = 4)$. There are only two phases three-strand bound state and unbound DNA, separated by the solid line. This is a first order transition line. The $y = y_c$ dotted vertical line is the two chain melting line and does not exist for the three chain system. The filled circle on the solid line at $y = 1$, $w = w_c$ is the pure three chain melting point. The horizontal line $w = 1$ corresponds to the traditional Efimov case with pure two body interaction. The Efimov region is the region enclosed by (i) the $w = w_c$ line, (ii) the solid line, and (iii) the vertical $y = y_c$ line. (b) The X-y phase diagram for $b = 9\,(d = 4.1699)$. The Bound phase now melts via the Efimov state for larger X, but at $y = y_c$ for smaller X. This phase diagram is similar to that in [43]. (c) The X-y phase diagram for $b = 16(d = 5)$. There is now a mixed phase on the $y > y_c$ side. The thin lines denote continuous transitions, while the melting from the Efimov side (thick solid line) is first order. The dotted line, a relic of the two chain melting, does not exist in the three chain system. The two continuous lines and the first order melting line meet at the multicritical point at X_+ at $y = y_c$.

3.1. Chains on a Sierpinski gasket

The polymers are restricted to the non-horizontal bonds of a Sierpinski Gasket to keep DNA length same for all the strands. The DNA problem can be classified into two different cases according to the constraints on walks. In one case polymers cannot cross each other and in the other case polymers can cross each other at any length of the polymers. The triple-chain melting was discussed in [18], where weights σ_{ij} and σ_{ijk} were assigned at the vertex for bubble opening or closure, and y for sharing same bond by two strands. Here we define a more general model by assigning an extra weight w for three chains sharing a bond. As in section 2, if three strands share a bond, the weight is $y^3 w$. This model contains some of the models of [18] and it reproduces essentially all of the results discussed there, in appropriate limits.

3.2. Partition functions

To describe exactly all possible configurations of the three chain system, we need to introduce the following partition functions, a_{10}, a_{11}, a_{20}, a_{21}, a_{22}, a_{30}, a_{31}, a_{32}, a_{33}. Different possible polymer walks are shown in figure 7. The generating functions in terms of sum over all the configurations for the non-crossing case can be written as

$$A_{10} = a_{10}^2, \tag{12a}$$

$$A_{20} = a_{20}^2, \tag{12b}$$

$$A_{30} = a_{30}^2, \tag{12c}$$

$$A_{11} = a_{10}^2\, a_{11} + a_{11}^2, \tag{12d}$$

$$A_{21} = (a_{10}\, a_{11} + a_{01}\, a_{20})\, a_{21}, \tag{12e}$$

$$A_{31} = (a_{11}\, a_{20} + a_{01}\, a_{30})\, a_{31}, \tag{12f}$$

$$A_{22} = a_{11}\, a_{21}^2 + a_{20}^2\, a_{22} + a_{22}^2, \tag{12g}$$

$$A_{32} = a_{21}^2\, a_{31} + (a_{10}\, a_{22} + a_{02}\, a_{30})\, a_{32}, \tag{12h}$$

$$A_{33} = a_{22}\, a_{31}^2 + a_{11}\, a_{32}^2 + a_{30}^2\, a_{33} + a_{33}^2, \tag{12i}$$

with the initial conditions

$$
\begin{aligned}
&a_{10} = 1, a_{11} = 1, a_{20} = y, a_{21} = y, a_{22} = y^2, \\
&a_{30} = y^3 w, a_{31} = y^3 w, a_{32} = y^4 w, a_{33} = y^6 w^2.
\end{aligned}
\tag{13}
$$

Here we have set the bubble initiation factors $\sigma_{ij} = \sigma_{ijk} = 1$. This is a generalization of the TS1 case of [18].

The generating functions for the crossing case can be written as (rest are same as equation (12))

$$A_{22} = a_{11}\, a_{21}^2 + 2a_{20}^2\, a_{22} + a_{22}^2, \tag{14a}$$
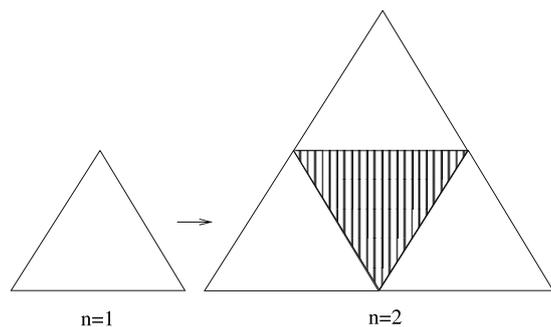
Figure 6. Schematic diagram of a Sierpinski Gasket for generation $n = 1$, 2. The basic motif is shown in (a). In next step shown in (b) three triangles are glued together with a forbidden region at the center.
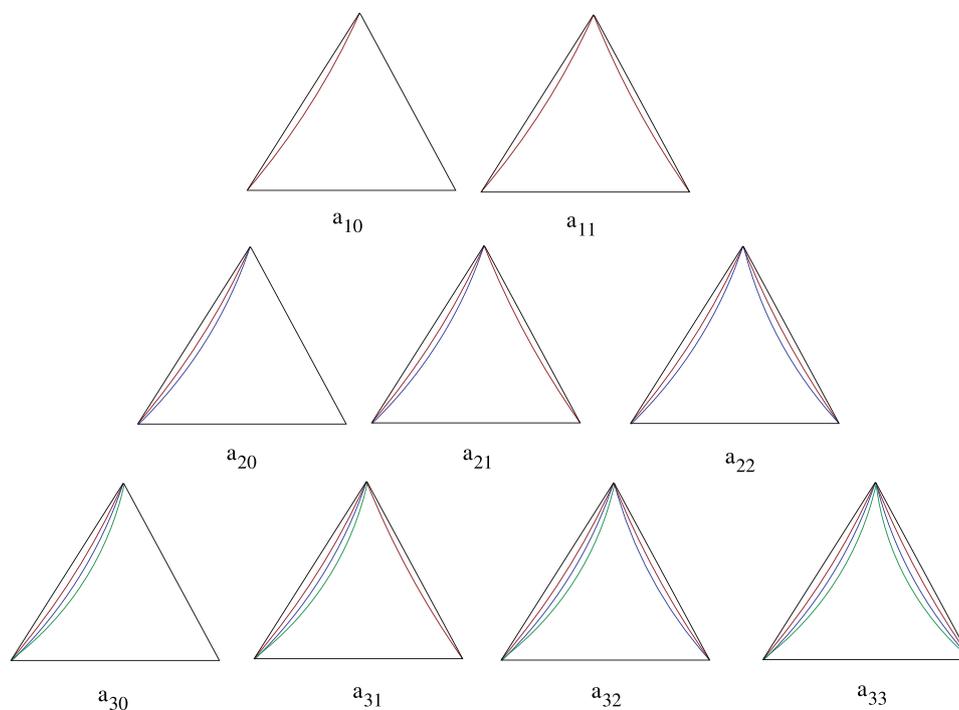

Figure 7. Possible configurations for single, double and triple chain systems. All possible configurations are represented by the partition functions a_{ij} ($i, j = 0$, 1, 2, 3).

$$A_{32} = 2a_{21}^2 \, a_{31} + (a_{10} \, a_{22} + a_{02} \, a_{30}) \, a_{32}, \tag{14b}$$

$$A_{33} = 3a_{22} \, a_{31}^2 + 3a_{11} \, a_{32}^2 + a_{30}^2 \, a_{33} + a_{33}^2, \tag{14c}$$

with the initial conditions shown in equation (13).

The total partition functions for the two chain and the three chain systems are given by

$$Z_2 = A_{11}^2 + A_{22}, \tag{15}$$

$$Z_3 = A_{11}^3 + A_{11}A_{22} + A_{33}, \tag{16}$$

with their logarithms giving the free energies. The two terms in equation (15) are for the unbound and for bound chain configurations. The first term of the two chain equation dominating over the second one results in a phase transition. A similar procedure of comparison has been adopted for the three chain system. The unbound, the bound, and the mixed phases appear over a wide range of temperatures. The mixed phase is represented by the middle term of equation (16). To obtain the phase diagram in w-y plane we look for the convergence or divergence of the following ratios

$$r_1 = \frac{A_{22}}{A_{11}^2}, \;\; r_2 = \frac{A_{33}}{A_{11}^3}, \;\; \text{and} \;\; r_3 = \frac{A_{33}}{A_{11}A_{22}}. \tag{17}$$

The phase boundaries separating the phases can be obtained from condition $r_i = 1$, for any $i = 1,\, 2,\, 3$.

3.3. Phase diagrams

3.3.1. Non-crossing case. The phase diagram for the non-crossing case is shown in figure 8. All the three phases, three-bound, three-unbound and mixed, appear here. The vertical line $y = y_c$ corresponds to the two chain melting and it is also the melting line for the mixed phase. On the bound side, $y > y_c$, there is a triplex-mixed phase transition line. The topology of the phase diagram is similar to figure 5(c), but here all are first order lines. The three transition lines meet at the triple point at y_c. The dashed line, $w = 1/y$ verifies the results of TS1 model of [18]. TS1 is defined as the model of noncrossing walks with $y_{12} = y_{23} = y_{31} = y$, $\sigma_{ij} = \sigma$, and $w = 1/y$. A sequence of transitions

$$\text{triplex} \leftrightarrow \text{mixed} \leftrightarrow \text{denatured}$$

occurs in this case.

3.3.2. Crossing case. For $y > y_c$ when chains are supposed to be in pairs, the three chain repulsion plays an essential role to produce the mixed phase by stopping three monomers contact at a time. For the non-crossing case each of the three phases (bound, unbound, and mixed) occurs for a wide range of temperatures whereas for the crossing case there is no two chain melting at any finite temperature. See figure 9. This is because the bubble entropy in low dimensions is not enough to induce a melting. As a result the DNA strands remain bound at all temperatures for any arbitrarily weak short range pair attraction. Nevertheless the situation can be changed if the three chain repulsive interaction is incorporated. A transition from the bound to the mixed state is possible by tuning the three chain repulsive force among chains.

Furthermore transitions can be induced in the crossing case if we introduce the bubble initiation or closure factor σ in equation (13), e.g. by considering the initial conditions $a_{21} = y\sigma$, $a_{31} = y^3w\sigma^2$, $a_{32} = y^4w\sigma^2$, as in the TS2 model of [18].

With $\sigma = 0$ when the bubble formation is suppressed, the chains will be either open or bound, i.e. once the chains are unbound reunion is not possible along the length of chains. This is the Y-fork model. In the Y-fork model the crossing and the non-crossing case do not have any difference as the strand exchange is not allowed for $\sigma = 0$. We refer to [21] for discussions on the Euclidean lattice results.
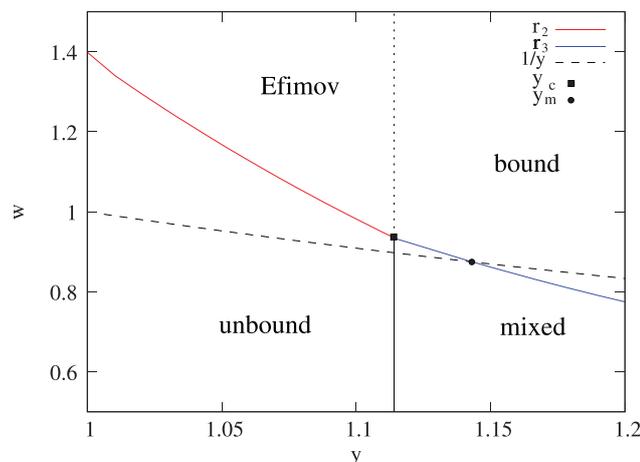
Figure 8. Phase diagram in the w-y plane for the non-crossing case with $\sigma = 1$. The Efimov (red) and the mixed (blue) phase boundaries intersect the vertical line $y = y_c$. The dashed line is for the TS1 model.

By equating the partition function of the bound state to the partition function of the unbound state, the duplex melting turns out to be

$$A_{11}y^N = A_{11}^2 \implies wy_0 = A_{11}^{1/N}, \tag{18}$$

where $y_0 = y_c(\sigma = 0)$. It is known that $y_0 = 1.264\,08.....$ We may use the same logic for the triplex melting

$$w_t = \frac{A_{11}^{2/N}}{y^3} = \frac{y_0^2}{y^3} \tag{19}$$

This relation fits the numerically obtained points in figure 10, suggesting that the phase is a triple chain bound state and not necessarily the Efimov-DNA. This should be the case because with $\sigma = 0$ there are no bubbles. This point is further elaborated below.

Using similar logic as above, the triplex to mixed transition occurs at

$$w_m = \frac{A_{11}^{1/N}}{y^2} = \frac{y_0}{y^2}. \tag{20}$$

where $N = 2^{n+1}$ is the length of a polymer. The arguments used here are very similar to those in section 2.2.1, and show the roles played by wy^3 and wy^2 in different transitions.

All the above three transitions occur at the same temperature $y = y_0$ if we choose $w = 1/y$ in a Y-fork model. It is also apparent from figure 10 that the three chain phase boundary and the mixed phase boundary meet at $w_0 = 1/y_0$ for $\sigma = 0$.

Bubbles are essential for the Efimov-DNA and the bubble-bound mixed phase. Therefore, Efimov-DNA may occur for $\sigma \neq 0$ but definitely not at $\sigma = 0$. The triplex melting line w_t given by equation (19) is from the tightly bound to the unbound phase, without any bubble. Incidentally, the Efimov state is a continuation of the three chain bound phase, and not a distinct thermodynamic phase; there is no phase boundary to protect it, except for melting. We may then identify the Efimov region for $0 < \sigma < 1$ as the region $w < w_t$ up to the corresponding melting line and $y < y_c(\sigma)$ (see the inset in figure 11). This is the region where the bubbles contribute most. A similar situation
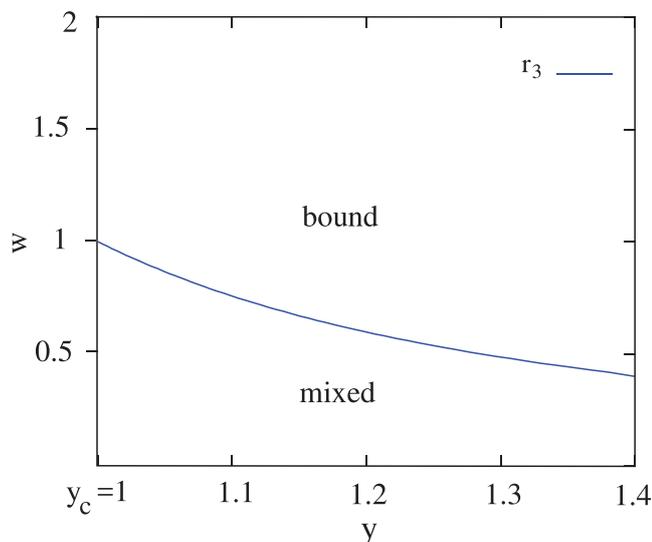
Figure 9. Phase diagram in the w-y plane for the crossing case with $\sigma = 1$. Two chain melting is at $y = y_c$. The phase boundary is for the transition from the bound to the mixed phase. No two chain or pure three chain melting at finite temperature as $y_c = 1$ ($w_c = y_c^{-1} = 1$).
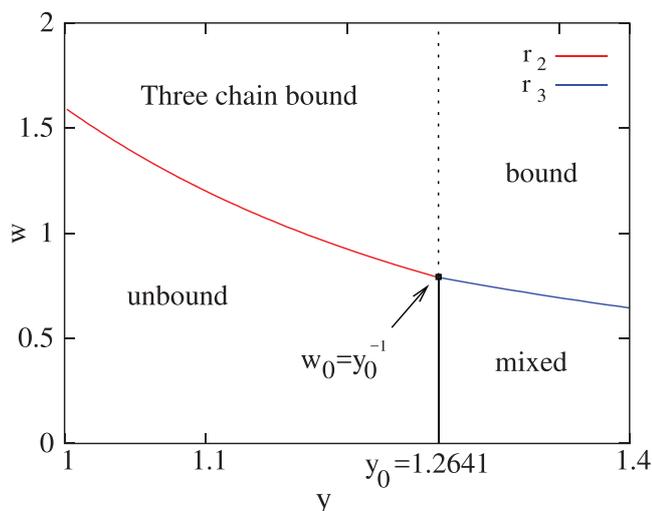

Figure 10. Phase diagram in the w-y plane for the crossing case with $\sigma = 0$. All the phases occur in this case like in the non-crossing case. Three phases (three-chain bound, all bound and the mixed phase by continuation) coexist at the triple point ($y_0 = 1.264..$, $w_0 = y_0^{-1}$).

arises in the context of the bubble-bound mixed phase, whose existence is also at stake at $\sigma = 0$. In the $\sigma = 0$ limit, a transition takes place at w_m (equation (20)) to a phase where any two will be bound throughout and one free (see figure 4). Absence of bubbles strictly implies no strand exchange. As a result, the transition is actually like peeling off one chain from the three. We may still call this a mixed phase by continuation because, by symmetry, any one can be unbound.

The projection of the w-y-σ phase diagram for the crossing case is depicted in figure 11 for different σ ($=1, 0.8, 0.6, 0$) in the w-y plane. The melting lines for the three
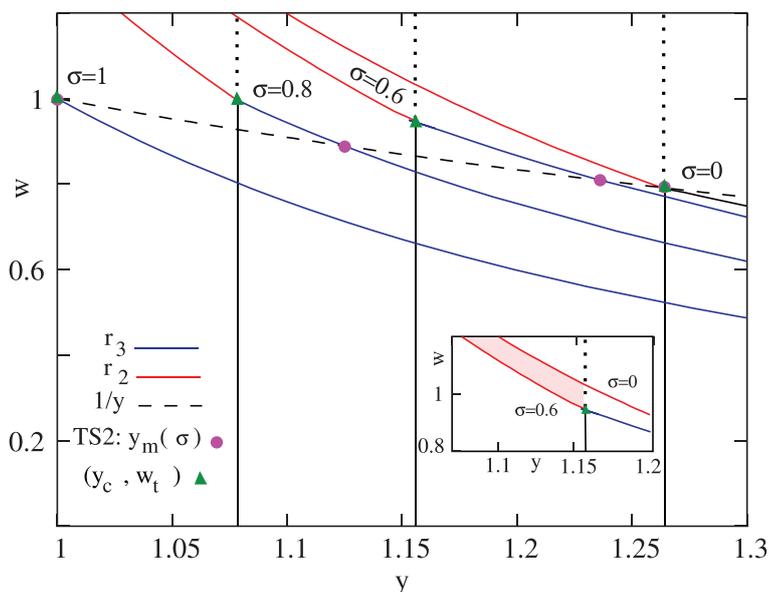
Figure 11. Crossing case. Phase diagram in the $w - y$ plane for different $\sigma(=1, 0.8, 0.6, 0)$. For each σ, the thin red line is the phase boundary for the unbound to the Efimov region of the triplex bound state, whereas the blue phase boundary (thick line) is for the bound to the mixed state. These two boundaries meet at the filled triangles on the vertical lines at the two chain melting $y = y_c(\sigma)$, which is also the melting line of the mixed phase. Filled triangles are the triple points. The dashed curve $w = 1/y$ intersects the mixed phase boundaries at $y_m(\sigma)$ (shown by filled circles), and reproduces the phase diagram of TS2 model [18] in the σ-y plane. The triangle and the circle coincide only at $\sigma = 0$ and $\sigma = 1$. Inset: in the inset the shaded part denotes the Efimov region for $\sigma = 0.6$. This region is enclosed by two Efimov transition lines corresponding to $\sigma = 0$ and $\sigma = 0.6$. See text.

chain and the mixed phase meet at $w_t(\sigma) = w_m(y_c)$ for any given value of σ. Model TS2 of [18] can be recovered exactly from this model. The curve $w = 1/y$ intersects the mixed phase boundaries for different values of $(\sigma, y_c(\sigma))$. These data certainly reproduce the σ-y phase diagram of the TS2 model in [18].

4. Conclusion

The role of dimensionality of the underlying lattices on the emergence of the mixed or the bubble-bound phase of a triple-stranded DNA has been the focus of this paper. The mixed phase was first discussed in the context of a class of DNA models on low dimensional fractal lattices, where a melting transition is induced by a bubble initiation factor that suppresses bubble entropy. We showed that the phase remains stable even with three body repulsion. With native DNA pair interaction and three chain repulsion, we further showed that such a phase is thermodynamically stable on the bound side of a duplex if dimensionality is large ($d > 4.5$). We established a diverging specific

heat for $d = 5$ from both renormalization group and finite size scaling analysis of exact computations.

When a bubble-bound phase exists, the topology of the phase diagram remains the same for both higher and low dimensional models. However the nature of the transitions are different. In general, the transition from the mixed bubble-bound phase to the three chain bound state is continuous for $d > 4.5$, but for $d < 2$, all transitions are first order.

A few comments on dimensionality are in order. The evidence is that the mixed phase is present for $d < 2$ and for $d > d_c = 4.46...$. This is true even at $\sigma = 1$ (figure 9). We may generalize to say that there exists a d^* such that the mixed phase is seen for $d < d^*$ and $d > d_c$, with $\ln 3/\ln 2 < d^* < \ln 6/\ln 2$, although we cannot evaluate d^* in the present study. This 'reentrant' behaviour as a function of d is probably associated with the different order of the mixed-bound and mixed-unbound transitions seen in different dimensions. This calls for more elaborate studies on systems with both σ and w in spaces whose dimensions can be tuned finely over a bigger range than studied here.

Appendix. Derivations of equations (1a) and (1b)

The RG recursion relation for double stranded DNA is given by

$$Z_{n-1}(y') \propto Z_n(y), \tag{A.1}$$

where left (right) hand side of the equation is the $(n-1)$th (nth) generation partition function, following the decimation as shown in figure 1. The partition functions are given by equations (11a) and (11b). The RG transformation for y and its renormalized value y' can then be written as (see figure 1)

$$y' = A\{b(b-1) + by^2\}, \tag{A.2}$$

with the constant A determined by the condition that infinite temperature case corresponds to $y = y' = 1$. This gives $A = b^{-2}$. Hence equation (1a).

The RG transformation for the three-chain case can also be written in the same spirit as in the two-chain case as

$$Q_{n-1}(y', w') \propto Q_n(y, w) \tag{A.3}$$

but with y' determined by equation (A.2). For the single step of figure 1, writing the renormalized three strand weight as $w'y'^3$, we get

$$w'y'^3 = B\{b(wy^3)^2 + 3b(b-1)y^2 + b(b-1)(b-2)\}, \tag{A.4}$$

with B to be determined such that at infinite temperature, $y = w = 1$, we get $y' = w' = 1$. This sets the proportionality constant as $B = b^{-3}$. Hence equation (1b).

References

[1] Sinden R R 2012 *DNA Structure and Function* (New York: Academic)
[2] Moser H E and Dervan P B 1987 *Science* **238** 645
[3] Grigoriev M, Praseuth D, Guieysse A L, Robin P, Thuong N T, Helene C and Harel-Bellan A 1993 *Proc. Natl Acad. Sci.* **90** 3501
[4] Tu G C, Cao Q N and Israel Y 1995 *J. Biol. Chem.* **270** 28402
[5] Knauert M and Glazer P M 2001 *Hum. Mol. Genet.* **10** 2243

[6] Wang G, Levy D D, Seidman M M and Glazer P M 1995 *Mol. Cell. Biol.* **15** 1759
[7] Diviacco S, Rapozzi V, Xodo L, Helene C, Quadrifoglio F and Giovannageli C 2001 *FASEB J.* **15** 2660
[8] Guo M *et al* 2015 *J. Biol. Chem.* **290** 5174
[9] Jain A *et al* 2008 *Biochimie* **90** 1117
[10] Duca M *et al* 2008 *Nucl. Acids Res.* **36** 5123
[11] Conde J *et al* 2016 *Nat. Mater.* **15** 353
[12] Rich A 2006 *J. Biol. Chem.* **281** 7693
[13] Li Y, Syed J and Sugiyama H 2016 *Cell Chem. Biol.* **23** 1325
[14] Holland J A and Hoffman D W 1996 *Nucl. Acids Res.* **24** 2841
[15] Frank-Kamenetskii M D 1999 *Triple Helix Forming Oligonucleotides* ed C Malvy *et al* (Boston: Kluwer)
[16] Maji J, Bhattacharjee S M, Seno F and Trovato A 2010 *New J. Phys.* **12** 083057
[17] Maji J and Bhattacharjee S M 2012 *Phys. Rev.* E **86** 041147
[18] Maji J, Bhattacharjee S M, Seno F and Trovato A 2014 *Phys. Rev.* E **89** 012121
[19] Pal T, Sadhukhan P and Bhattacharjee S M 2013 *Phys. Rev. Lett.* **110** 028105
[20] Pal T, Sadhukhan P and Bhattacharjee S M 2015 *Phys. Rev.* E **91** 042105
[21] Mura F, Bhattacharjee S M, Maji J, Masetto M, Seno F and Trovato A 2016 *J. Low Temp. Phys.* **185** 102
[22] Efimov V 1970 *Phys. Lett.* B **33** 563
[23] Efimov V 1971 *Sov. J. Nucl. Phys.* **12** 589
[24] Braaten E and Hammer H W 2006 *Phys. Rep.* **428** 259
[25] Kraemer T *et al* 2006 *Nature* **440** 315
[26] Zaccanti M *et al* 2009 *Nat. Phys.* **5** 586
[27] Pires R, Ulmanis J, Häfner S, Repp M, Arias A, Kuhnle E D and Weidemüller M 2014 *Phys. Rev. Lett.* **112** 250404
[28] Piatecki S and Krauth W 2014 *Nat. Commun.* **5** 3503
[29] Nishida Y, Kato Y and Batista C D 2013 *Nat. Phys.* **9** 93
[30] Moroz S, D'Incao J P and Petrov D S 2015 *Phys. Rev. Lett.* **115** 180486
[31] Langmack C, Smith D H and Braaten E 2012 *Phys. Rev.* A **86** 022718
[32] Horinouchi Y and Ueda M 2015 *Phys. Rev. Lett.* **114** 025301
     Horinouchi Y and Ueda M 2016 *Phys. Rev.* A **94** 050702
[33] Stasio M D, Seno F and Stella A L 1992 *J. Phys. A: Math. Gen.* **25** 3891
[34] Gefen Y, Mandelbrot B B and Aharony A 1980 *Phys. Rev. Lett.* **45** 855
[35] Knezevic M and Vannimenus J 1986 *Phys. Rev. Lett.* **56** 1591
[36] Elezovic-Hadzic S and Zivic I 2013 *J. Stat. Mech.* P02045
[37] Derrida B, De Seze L and Itzykson C 1983 *J. Stat. Phys.* **33** 559
[38] Mukherji S and Bhattacharjee S M 1995 *Phys. Rev.* E **52** 1930
[39] Bhattacharjee S M 2005 *Statistics of Linear Polymers in Disordered Media* ed B K Chakrabarti (Amsterdam: Elsevier)
[40] Hinczewski M and Berker A N 2006 *Phys. Rev.* E **73** 066126
[41] Lyra M L *et al* 2014 *Phy. Rev.* E **89** 052133
[42] Chakrabarti A and Bhattacharyya B 1996 *Phys. Rev.* B **54** R12625
     Chakrabarti A 1996 *J. Phys.: Condens. Matter* **8** L99
[43] Tabbara R, Owczarek A L and Rechnitzer A 2016 *J. Phys. A: Math. Theor.* **49** 154004
[44] Haddad T A S, Andrade R F S and Salinas S R 2004 *J. Phys. A: Math. Gen.* **37** 1499
[45] Bhattacharjee S M and Seno F 2001 *J. Phys. A: Math. Gen.* **34** 6375