



**Title:**

**Cold denaturation of RNA secondary structures with loop entropy and quenched disorder**

**Author(s):**

Flavio Iannelli, Yevgeni Mamasakhlisov, and Roland R. Netz

Document type: Preprint

Terms of Use: Copyright applies. A non-exclusive, non-transferable and limited right to use is granted. This document is intended solely for personal, non-commercial use.

**Citation:**

"Cold denaturation of RNA secondary structures with loop entropy and quenched disorder  
Flavio Iannelli, Yevgeni Mamasakhlisov, and Roland R. Netz  
Phys. Rev. E 101, 012502 ; <https://doi.org/10.1103/PhysRevE.101.012502>"

# Cold denaturation of RNA secondary structures with loop entropy and quenched disorder

Flavio Iannelli\*

*Humboldt-Universität zu Berlin, Institut für Physik, Newtonstraße 15, 12481 Berlin*

Yevgeni Mamasakhlisov

*Department of Molecular Physics, Yerevan State University,  
1 Alex Manougian Street, Yerevan 0025, Armenia*

Roland R. Netz

*Fachbereich Physik, Freie Universität Berlin, 14195 Berlin, Germany*

We study the folding of RNA secondary structures with quenched sequence randomness by means of the constrained annealing method. A thermodynamic phase transition is induced by including the conformational weight of loop structures. In addition to the expected melting at high temperature, a cold melting transition appears. Our results suggest that the cold denaturation of RNA found experimentally is in fact a continuous phase transition triggered by quenched sequence disorder. We calculate both hot and cold melting critical temperatures for the competing energy scenario between favorable and unfavorable base pairs and present a phase diagram as a function of the loop exponent and temperature.

## I. INTRODUCTION

Ribonucleic acids are biopolymers that are crucial to all living systems, they process and transmit genetical informations and take part in many important cellular activities [1]. The RNA primary structure is a chain sequence which consists of four bases U, A, G and C, while the secondary structure is the listing of base pairings occurring in the more complex tertiary structure. An interesting phenomena emerging at low temperatures is the folding of these molecules into their native conformations. Like in the case of proteins, the folding of RNA structures is crucial for understanding their biological functions and has been vastly studied, grounded mostly in the idea of hierarchical folding [2]. In this scheme the primary structure alone determines the folding mechanism and therefore the secondary structure forms independently of the tertiary structure. Such a great simplification makes it possible to exactly compute the secondary structure partition function of given RNA sequences. The molecule folding can be qualitatively and quantitatively addressed via the experimentally observable helicity degree defined as the average fraction of paired bases, which increases when lowering the temperature in the standard scenario. Here we will focus on a two-letter alphabet binding energy model with symbols chosen from the subset {U,A}, in the spirit of [3, 11, 17, 23, 27] which physically corresponds to the hydrophilic-hydrophobic model for protein folding [13]. Although this approximation is an oversimplification, it has proven to reproduce in a reasonable fashion the thermodynamics of these molecules and it is able to capture the essential physics behind the folding as well as the glass transition [23]. The so-called

Watson-Crick base pairs UA, and GC, are the most stable hydrogen bonds. Since each base is essentially planar and its conformations are limited, every RNA secondary structure is defined by a list of pairings  $(i, j)$  with each position appearing at most once. In addition for any two base pairings  $(i, j)$  and  $(k, l)$ , we only consider nested base pairings, where  $i < k < l < j$ , and independent base pairings, where  $i < j < k < l$ . A third possibility would be to include pseudoknots, which are rare in real RNA, where one has  $i < k < j < l$ , but they can be excluded as a first approximation [2, 29]. This choice defines a hierarchical structure for the RNA conformation so that a recursive equation for the partition function can be used. Furthermore, since the exclusion of pseudoknots discards the configurations that are not planar, when considering disordered sequences the system features frustration [22, 23]. When the molecule is not in the native state, i.e. when there is a high number of unpaired bases, also loops play a crucial role and it is not rare that these are plentiful even at low temperatures. The configurational entropy contribution  $\Delta S_n^{loop} \sim k_B \ln n^{-c}$  for loops of length  $n$  is characterized by the universal loop exponent  $c$  [21]. For ideal polymers, which are modeled as simple random walks, the exponent is  $c_{RW} = D/2$  where  $D$  is the spatial dimensionality. Instead, for self-avoiding walks one finds in three dimensions  $c_{SAW} \approx 1.76$ , which even increases further in real polymers since  $c$  depends also on the number of helical strands emerging from loops [4]. For the model with no disorder, homopolymeric RNA, in the range  $2 < c \lesssim 2.479$  a phase transition from the folded to the unfolded state, usually referred to as the molten phase [17], is known to occur when the temperature increases up to the melting point  $T_m(c)$  [5] and the  $c$ -dependent critical exponents have been analytically obtained [12]. A finite loop exponent has been shown to significantly improve salt-dependent RNA folding compared

---

\* iannelli.flavio@gmail.com

with experiments [20].

In this paper we address the influence of a loop exponent  $c \neq 0$  on the behavior of RNA secondary structures with random sequences where the disorder is quenched, i.e. fixed. This allows us to characterize a generic RNA molecule and in fact most of the results obtained here apply also to DNA. A very interesting phenomenon that characterizes these polymers and proteins is the *cold denaturation* [6, 31, 33, 34]. The denaturation of proteins and polymers rising the temperature is a consequence of the increase in configurational entropy. Denaturation when lowering the temperature is usually interpreted in terms of hydrophobic interactions. Experimentally, denaturation can be inferred by the presence of peaks in the specific heat which physically follow from an abrupt increase of the system entropy [7–10]. In this paper we give an alternative explanation for cold denaturation in terms of quenched disorder which itself weakens the secondary structure formation at low temperature. The double peak behavior of  $C_V$  turns out to be associated with two different melting temperatures of the RNA secondary structure.

## II. DISORDERED RNA

### A. The model

For a sequence  $h$  of length  $N$  we define the base pairing matrix  $\mathcal{S}$ , which completely determines the secondary structure, as the  $N \times N$  symmetric matrix with components  $s_{i,j}$  equal to unity if  $(i, j)$  are paired and vanishing otherwise. If stack energies are neglected, the Hamiltonian for a given sequence configuration can be written as a sum over non-repeated base pairs

$$\mathcal{H}(\mathcal{S}, h) = \sum_{(i,j) \in \mathcal{S}} \epsilon_{i,j} = \sum_{1 \leq i < j \leq N} s_{i,j} \epsilon_{i,j}. \quad (1)$$

Here we take the simplest non trivial pairing energy function as the sum of a constant and a disorder term in the spirit of [27] as

$$\epsilon_{i,j} = \epsilon_0 + \epsilon h_i h_j, \quad (2)$$

which can be easily generalized to a four-letter alphabet RNA. The sign of the constant  $\epsilon_0$  defines the nature of the background interaction between nucleotides. If  $\epsilon_0 > 0$  the interaction is repulsive and attractive otherwise. The second term is the product of two independent variables of the form of a spin glass model for neural networks [15], multiplied by a constant energy  $\epsilon$ . We assign Ising variables to each base along the chain so that  $h_i = +1$  if  $i$  is the nucleotide U and  $h_i = -1$  if it is A. Contrary to the base pairing matrix elements  $s_{i,j}$ , which are free to evolve within the dynamics of the system, the site sequence variables  $h_i$  are frozen and not free to rearrange to minimize the total energy. Having fixed the sign of the background interaction  $\epsilon_0$ , the absolute value of the

ratio  $\epsilon/\epsilon_0$  is the only relevant parameter characterizing the system behavior, which is a direct consequence of the two-state model adopted here. For each sequence  $h$  we take  $h_i$  as quenched random independent and identically distributed variables so that the probability distribution factorizes as

$$\mathcal{P}(h) = \prod_{i=1}^N \rho(h_i). \quad (3)$$

This construction, which is analytically more manageable, is supported by the fact that no strong correlations are found in the base type occurrence [23]. Defining the probability of finding the base U as  $p \equiv \rho(h_i = +1)$ , the probability distribution factors can be written as

$$\rho(h_i) = p\delta(h_i - 1) + (1 - p)\delta(h_i + 1). \quad (4)$$

Due to symmetry one only needs to explore the parameter range  $0 < p < 0.5$ .

### B. Partition function

The partition function of a given sequence is the sum over all allowed realizations of the base pairing matrix

$$Z_N(h) = \sum_{\{\mathcal{S}\}} e^{-\beta \mathcal{H}(\mathcal{S}, h)}, \quad (5)$$

where  $\beta = (k_B T)^{-1}$  and  $\{\mathcal{S}\}$  denotes the set of all secondary structures without pseudoknots for the given sequence  $h$ . The free energy is obtained by performing the quenched average, denoted by  $\overline{(\dots)}$ , of the disordered free energy

$$f(h) = -\frac{1}{\beta N} \ln Z_N(h), \quad (6)$$

over the disorder distribution eq. (3), yielding

$$\overline{f(h)} = \sum_{\{h\}} \mathcal{P}(h) f(h) = -\frac{1}{\beta N} \overline{\ln Z_N(h)}. \quad (7)$$

For sufficiently large chains the physical properties of the system do not depend on the specific disorder realisation  $\{h\}$  and the free energy self-averages [22]

$$\lim_{N \rightarrow \infty} f(h) = \overline{f(h)}. \quad (8)$$

Numerically the partition function is usually obtained via the recursive equation of the restricted partition function [16]

$$Z_{i,j+1} = Z_{i,j} + \sum_{k=i}^j w_{k,j+1} Z_{i,k-1} Z_{k+1,j}, \quad (9)$$

as illustrated in Fig. 1. On the right-hand side the first term corresponds to the probability that base  $j + 1$  is not paired, and the summation term corresponds to the

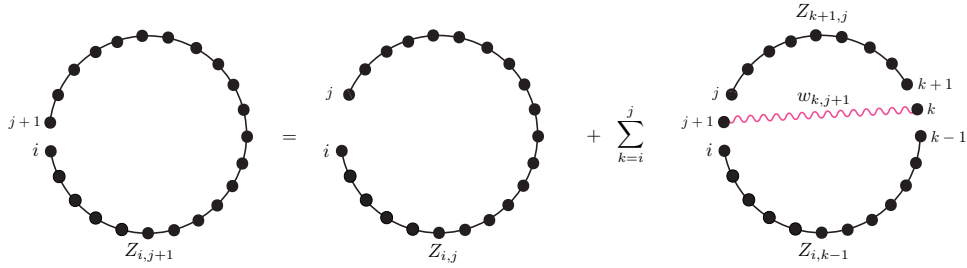


FIG. 1. (Color online) Hierarchical recursive scheme for the partition function eq. (9). The subchain partition function from base  $i$  to base  $j + 1$  is the sum of the partition function from base  $i$  to base  $j$  and the partition functions related to all nested or independent pairings formed with base  $j + 1$  by any base  $k \in (i, j)$ .

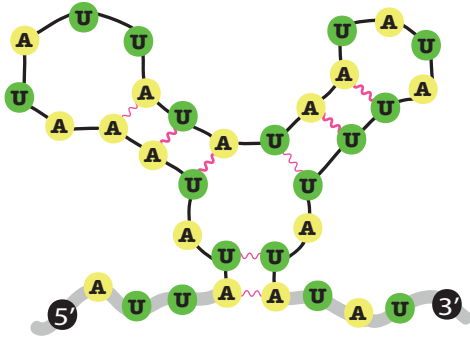


FIG. 2. (Color online) Secondary structure representation with stacks and loops. Nucleotides are the green and yellow dots for U and A respectively. The wavy red lines identify the hydrogen bonds for favorable (thicker line) and unfavorable (thinner line) base pairs and the black solid lines the nested backbone links. In thick grey are the non-nested backbone links which are accounted by the upper index  $M$  in the partition function eq. (12). For this structure going from the 5'– to the 3'–end the number of total links is  $N = 34$  out of which  $M = 8$  are the non-nested.

probability associated to all possible nested or independent pairings given that position  $k$  forms a pairing with position  $j + 1$  with statistical weight

$$w_{k,j+1} \equiv \exp(-\beta\epsilon_{k,j+1}). \quad (10)$$

This recursive equation allows to compute the exact partition function  $Z_N = Z_{1,N}$  without pseudoknots in a time of order  $\mathcal{O}(N^3)$ , starting with the boundary conditions  $Z_{i,i} = Z_{i,i-1} = 1, \forall i$ .

Einert *et al.* have shown how to take into account loops in the recursive equation by including the statistical weight

$$v_n = n^{-c}, \quad (11)$$

for each loop consisting of  $n$  links [24] and summing over the restricted partition functions of strands terminated by a helix. If  $Z_{i,j}^M$  denotes the partition function of a polymer going from monomer  $i$  to monomer  $j$  with  $M \leq j - i$  unlooped links, see Fig. 2, given the initial conditions

$Z_{i,i}^M = \delta_M$ ,  $Z_{i,i-1}^M = \delta_{M+1}$ , and  $Z_{i,j}^{-1} = \delta_{j+1-i}$ , the recursive partition function reads [19]

$$\tilde{Z}_{i,j+1}^{M+1} = \tilde{Z}_{i,j}^M + \sum_{k=i}^j w_{k,j+1} \tilde{Z}_{i,k-1}^M \tilde{Z}_{k,j+1}^0, \quad (12)$$

as illustrated in Fig. 3, with computational time  $\mathcal{O}(N^4)$ . Here  $\tilde{Z}_{i,j}^M \equiv Z_{i,j}^M / u_M$  is the partition function rescaled by the statistical weight of  $M$  non-looped links  $u_M$ , which will be not considered from this point onward, and

$$\tilde{Z}_{k,j+1}^0 = \sum_{l=-1}^{j-k-1} \tilde{Z}_{k+1,j}^l v_{l+2} \quad (13)$$

is the partition function of the arbitrary substrand that is terminated by a helix.

### C. Helicity degree

The most relevant quantity that characterizes the conformation of the molecule is the helicity degree defined as

$$\theta = \frac{2}{N} \sum_{i<j} \overline{\langle s_{i,j} \rangle} = \frac{2}{N} \overline{\langle \sum_{i<j} s_{i,j} \rangle} = \frac{2}{N} \overline{\langle |\mathcal{S}| \rangle}, \quad (14)$$

where  $\langle \dots \rangle$  denotes the average over the canonical ensemble and

$$|\mathcal{S}| \equiv \sum_{i<j} s_{i,j} \quad (15)$$

is the number of paired bases in the structure  $\mathcal{S}$ . Since  $\overline{\langle |\mathcal{S}| \rangle} \in [0, N/2]$ , the helicity degree is a function of the temperature in the interval  $[0, 1]$  with  $\theta = 1$  if every base is paired, corresponding to the native state, and  $\theta = 0$  if no base is paired. The helicity is a good measure of the order of the molecule conformation, since it gives the statistical weight of paired bases and thus can be used to quantify to what extent the molecule is folded. This can also be expressed in terms of the free energy by noting

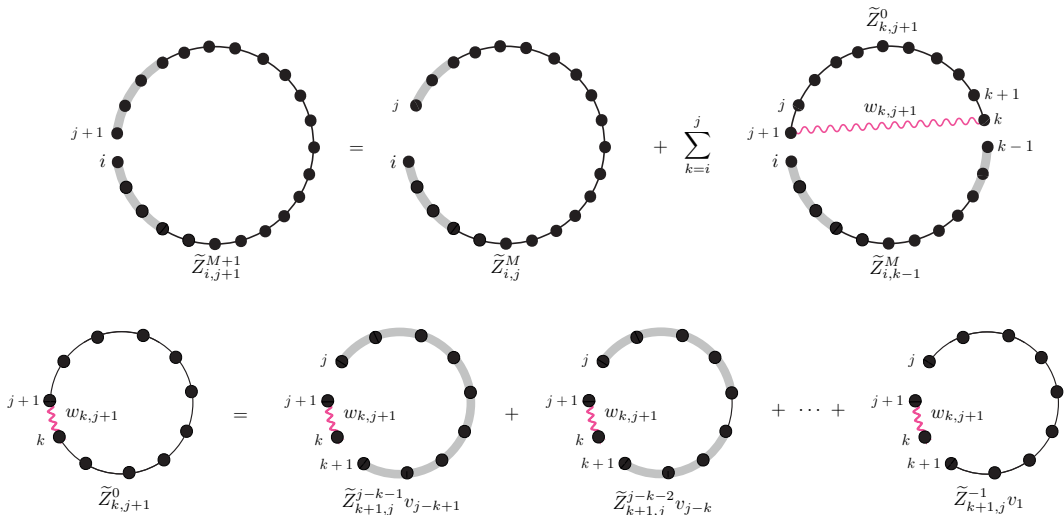


FIG. 3. (Color online) Recursion scheme for the canonical partition function eq. (12). The partition function of a strand ranging from  $i$  to  $j+1$  with  $M+1$  non-nested backbones (thick grey lines) is computed from the strand ranging from  $i$  to  $j$  with  $M$  non-nested backbones by adding a nested base at position  $j+1$  and considering all possible pairings with a base  $k \in (i, j)$ . Each of these pairings defines a structure which has zero non-nested backbones, i.e. an arbitrary substrand that is terminated by a helix. The explicit diagrams for the latter are showed in the second row and are obtained by considering the sum over all non-nested backbones with associated statistical weight given by eq. (11).

that

$$\begin{aligned} \theta &= \frac{2}{N} \frac{1}{Z_N(h)} \sum_{\{\mathcal{S}\}} e^{-\beta \epsilon_0 |\mathcal{S}|} e^{-\beta \epsilon \sum_{i < j} s_{i,j} h_i h_j} |\mathcal{S}| \\ &= \frac{2}{N} \left( -\frac{\partial}{\partial(\beta \epsilon_0)} \right) \ln \sum_{\{\mathcal{S}\}} e^{-\beta \epsilon_0 |\mathcal{S}|} e^{-\beta \epsilon \sum_{i < j} s_{i,j} h_i h_j} \\ &= 2 \frac{\partial}{\partial(\beta \epsilon_0)} \beta f(h). \end{aligned} \quad (16)$$

Numerically the helicity can be estimated in the absence of a loop free energy using the probability of base pair formation  $P_{i,j}$  between nucleotides  $i$  and  $j$  [17]

$$P_{i,j} = \langle s_{i,j} \rangle, \quad \Rightarrow \quad \theta = \frac{2}{N} \sum_{i < j} P_{i,j}, \quad (17)$$

where the binding probability is obtained from the recursive equation as [28]

$$P_{i,j} = e^{-\beta \epsilon_{i,j}} \frac{Z_{i,j}^{int} Z_{i,j}^{ext}}{Z_{1,N}}. \quad (18)$$

Here  $Z_{i,j}^{int}$  is given by the partition function of the internal sequence  $(i+1, \dots, j-1)$ , while  $Z_{i,j}^{ext}$  is the partition function of the external sequence  $(1, \dots, i-1, j+1, \dots, N)$ . The latter can be computed by extending the recursion relation to the duplicated sequence  $(1, \dots, N, N+1, \dots, 2N)$  as  $Z_{j+1, N+i-1}$  so that

$$P_{i,j} = e^{-\beta \epsilon_{i,j}} \frac{Z_{i+1, j-1} Z_{j+1, N+i-1}}{Z_{1, N}}. \quad (19)$$

The exact enumeration of 30 random sequences, through the partition function with no loops eq. (9), and the

quenched average of the helicity degree are shown in Fig. 4, where we compare the attractive and repulsive background energy scenarios for the disordered two state model.

### III. HOMOPOLYMER RESULTS

Most of the known results for RNA belong to the special case of homopolymer models which we shall review in this section. De Gennes was the first to obtain an expression for the the canonical partition function starting from the singularity analysis of the generating function using a propagator formalism [26]. By setting  $\epsilon = 0$  in eq. (1) the energy associated to each pairing becomes site independent, i.e.  $\epsilon_{i,j} = \epsilon_0, \forall(i, j)$ , thus making the energy of the structure  $\mathcal{S}$  depending only on the number of paired bases  $|\mathcal{S}|$ .

#### A. Folded RNA without loops

Since the restricted partition function is translationally invariant, in the no-loop scenario we can write  $Z_{i,j} = Q_{j-i+1}$  as the partition function of a homopolymer of length  $j-i+1$  which depends only on the difference between position  $i$  and position  $j$ . To decouple the summation it is useful to introduce the z-transform of  $Q_N$

$$Q(z) = \sum_{N=1}^{\infty} z^{-N} Q_N. \quad (20)$$

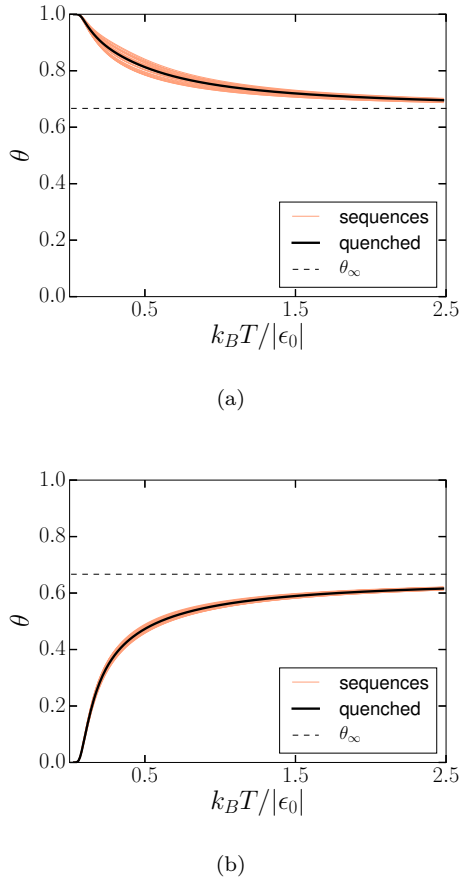


FIG. 4. (Color online) Helicity degree for (a)  $\epsilon_0 = -1$ ,  $\epsilon = 0.5|\epsilon_0|$  and for (b)  $\epsilon_0 = +1$ ,  $\epsilon = -0.5|\epsilon_0|$  in the no-loop scenario with probability of U base occurrence  $p = 0.75$ . Each red line corresponds to the helicity of a single RNA sequence realization of the disorder. The quenched average (black line) is obtained from the exact computation of the partition function for 30 random sequences of length  $N = 50$ . The black dashed line is the asymptote  $\theta_\infty$  for  $T \rightarrow \infty$  of eq. (25).

The canonical partition function  $Z_N = Q_N$  is then obtained by back transforming the positive root of the resulting equation for  $Q(z)$  and in the limit of large  $N$  a saddle point approximation yields [17]

$$Z_N = \sum_{\{S\}} w^{|S|} \sim \xi(w) N^{\alpha-1} z_0^{-N}, \quad (21)$$

where  $\alpha = -1/2$  and  $z_0 = 1/(1 + 2\sqrt{w})$ , while  $\xi(w)$  is a scaling function of the homopolymeric weight

$$w = \exp(-\beta\epsilon_0). \quad (22)$$

In this scenario the free energy assumes the same scaling for all temperatures with universal pre factor  $3/2$ , originally obtained by de Gennes [26], characteristic of the folded state for polymers and therefore no phase transition takes place. From eq. (21) it becomes possible to express the helicity degree as a function of the statistical

weight in a very intuitive form. Indeed by writing the partition function as  $Z_N = \sum_{\{S\}} e^{|\mathcal{S}| \ln w}$ , we have

$$\theta^{hom} = \frac{2}{N} \langle |\mathcal{S}| \rangle = \frac{2}{N} \frac{1}{Z_N} \frac{\partial}{\partial \ln w} Z_N = \frac{2}{N} \frac{\partial \ln Z_N}{\partial \ln w}. \quad (23)$$

Then using  $w \partial / \partial w = \partial / \partial \ln w$ , for  $N \gg 1$ , the helicity in the folded state takes the simple form

$$\begin{aligned} \theta^{hom} &= \frac{2}{N} w \frac{\partial \ln Z_N}{\partial w} \\ &= \frac{2}{N} w \frac{\partial}{\partial w} \left( \ln \xi(w) - \frac{3}{2} \ln N + N \ln(1 + 2\sqrt{w}) \right) \\ &\approx \frac{2\sqrt{w}}{1 + 2\sqrt{w}}, \end{aligned} \quad (24)$$

which asymptotically approaches a constant value

$$\theta_\infty = \lim_{T \rightarrow \infty} \theta^{hom} = \lim_{\beta \rightarrow 0} \frac{2e^{-\beta\epsilon_0/2}}{1 + 2e^{-\beta\epsilon_0/2}} = \frac{2}{3}. \quad (25)$$

## B. RNA folding with loops

As in the no-loops scenario for homopolymers we set  $w_{k,j} = w$  so that eq. (12) reads

$$\tilde{Q}_{N+1}^{M+1} = \tilde{Q}_N^M + w \sum_{k=0}^N \sum_{l=-1}^{N-k-1} \frac{\tilde{Q}_{k-1}^M \tilde{Q}_{N-k-1}^l}{(l+2)^c}, \quad (26)$$

where the rescaled homopolymeric partition function  $\tilde{Q}_N^M$  describes a polymer with  $N$  links and  $M$  non-looped links with  $-1 \leq M \leq N$ . In absence of external forces the canonical partition function, which includes loop structures, is obtained by summing over all backbones as  $Z_N^{loop} = \sum_{M=0}^{\infty} \tilde{Q}_N^M$  and the grand-canonical partition function follows as

$$\mathcal{Z}^{loop}(z) = \sum_{N=0}^{\infty} z^N Z_N^{loop} = \sum_{N=0}^{\infty} \sum_{M=0}^{\infty} z^N \tilde{Q}_N^M, \quad (27)$$

where  $z$  is the fugacity. By performing the double sum  $\sum_{N=0}^{\infty} z^N \sum_{M=0}^{\infty}$  on both sides of eq. (26), after rearranging indices one obtains [12]

$$\mathcal{Z}^{loop}(z) = \frac{\kappa(w, z)}{1 - z\kappa(w, z)}, \quad (28)$$

where

$$\kappa(w, z) \equiv 1 + w \sum_{l=-1}^{\infty} \sum_{N=l}^{\infty} \frac{z^{N+2} \tilde{Q}_N^l}{(l+2)^c} \quad (29)$$

is the grand-canonical partition function of RNA structures with zero non-nested backbone links, i.e. structures which consist of just one nucleotide or structures where the terminal bases are paired. For  $|z\kappa| < 1$  the grand-canonical partition function eq. (28) can be expanded into a geometric series as  $\mathcal{Z}^{loop} = \sum_{M=0}^{\infty} z^M \kappa^{M+1}$ , and comparing the coefficients of the power series with eq.

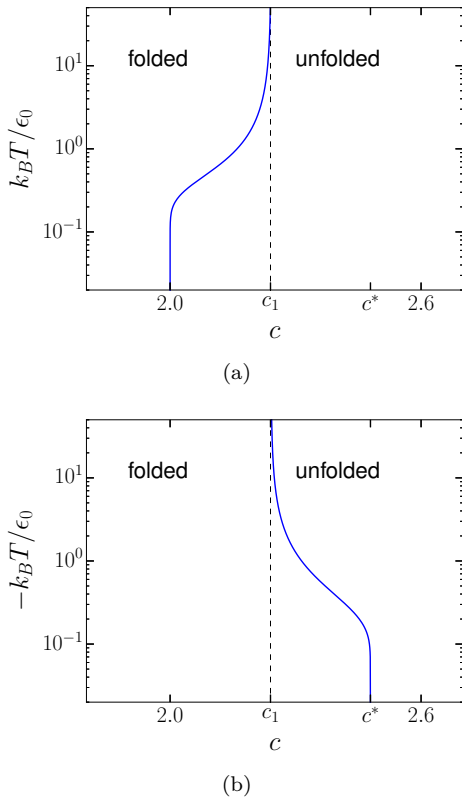


FIG. 5. (Color online) Phase diagram of homopolymeric RNA in the  $T - c$  plane featuring the unfolded and folded phase for repulsive (a) and attractive (b) base pair interaction energy. The critical lines are obtained by solving  $w_m(c) = w$  and in both cases they diverge for  $c = c_1$ . For  $c \leq 2$  the molecule is always folded and at  $c = c^* \approx 2.479$  the critical weight  $w_m(c)$  diverges so that for  $c > c^*$  no folded phase can exist for both attractive and repulsive scenarios.

(27) leads to  $\sum_{N=M}^{\infty} z^N \tilde{Q}_N^M = z^M \kappa^{M+1}$ . Using this relation, eq. (29) can be written as

$$\kappa(w, z) - 1 = \lambda(z, \kappa(w, z)), \quad (30)$$

where

$$\lambda(z, \kappa(w, z)) \equiv \frac{w}{\kappa} \text{Li}(c, z\kappa(w, z)), \quad (31)$$

and  $\text{Li}(c, x) \equiv \sum_{n=1}^{\infty} x^n n^{-c}$  is the polylogarithm [18]. This relation yields the first constitutive equation of the homopolymers theory with loop entropy. Going back to the canonical ensemble from eq. (28) the partition function takes the general form of eq. (21) but with  $\alpha$  and  $z_0$  not determined univocally. In fact, contrary to the no-loop scenario, now the grand-canonical partition function features two relevant singularities. These are the single pole  $z_0 = z_p$ , where the denominator of eq. (28) vanishes, and the branch point  $z_0 = z_b$  of the function  $\kappa(w, z)$ , characteristic of the unfolded and folded phase respectively.

For  $z < z_b$  at least one real solution of eq. (30) exists, while exactly at  $z = z_b$  the two solutions eventually

merge and the slope of  $\lambda(z, \kappa(w, z))$  at the tangent point  $z_b$  equals unity. Imposing this condition, eq. (30) yields

$$\kappa_b^2 = w [\text{Li}(c-1, z_b \kappa_b) - \text{Li}(c, z_b \kappa_b)], \quad (32)$$

where we use the short notation  $\kappa_b \equiv \kappa(w, z_b)$ . This relation together with eq. (30) univocally determines the branch singularity  $z_b$ , if it exists. By a first order expansion of  $\kappa(w, z)$  near the branch point the canonical partition function scales as

$$Z_N^{\text{loop}} \sim \xi_b(w) N^{-3/2} z_b^{-N}, \quad (33)$$

which leads for the free energy to the logarithmic  $N$ -contribution with universal prefactor  $3/2$ , in agreement with the no-loop scenario eq. (21). This is therefore the partition function describing homopolymeric RNA with loop structures in the folded phase. From eq. (33) the helicity degree eq. (23) follows as [12]

$$\theta_b \approx \frac{2w}{N} \frac{\partial \ln z_b^{-N}}{\partial w} = -\frac{2w}{z_b} \frac{\partial z_b}{\partial w} = \frac{2\text{Li}(c, z_b \kappa_b)}{\text{Li}(c-1, z_b \kappa_b)}. \quad (34)$$

Instead the simple pole  $z_p$  is determined, together with eq. (30), by

$$z_p \kappa_p = 1, \quad (35)$$

where  $\kappa_p \equiv \kappa(w, z_p)$ . By inserting this into eq. (30) yields an explicit expression for the pole singularity as

$$z_p = 2 \left( 1 + \sqrt{1 + 4w\zeta(c)} \right)^{-1}, \quad (36)$$

where we use the Riemann zeta function  $\zeta(c) = \text{Li}(c, 1) = \sum_{n=1}^{\infty} n^{-c}$ . In this scenario the partition function scales as

$$Z_N^{\text{loop}} \sim \xi_p(w) z_p^{-N}, \quad (37)$$

which in contrast to the branch point does not lead to the logarithmic  $N$ -contribution for the free energy, since the singularity exponent is  $\alpha = 1$ , and describes the thermodynamics of homopolymeric RNA above the melting critical point. In this phase the helicity degree takes the form

$$\theta_p \approx -\frac{2w}{z_p} \frac{\partial z_p}{\partial w} = 1 - \frac{1}{\sqrt{1 + 4w\zeta(c)}}. \quad (38)$$

Since the critical behavior of the system is characterized by the singularity that is closest to the origin in the complex  $z$  plane, a phase transition is possible only if a critical fugacity  $z_m$  and a critical weight  $w_m$  exist such that  $z_m = z_b(w_m) = z_p(w_m)$ . Then, at the critical point the three constitutive equations (30), (32) and (35) have to hold simultaneously. Using equations (35) and (36) a closed form expression can be given for the critical weight as a function of the loop exponent only

$$w_m(c) = \frac{\zeta(c-1) - \zeta(c)}{(\zeta(c-1) - 2\zeta(c))^2}. \quad (39)$$

This expression, which determines the critical line in the homopolymeric phase diagram, defines an upper bound for the range of the loop exponent  $c$  in which a phase transition can occur. This is given by the universal value  $c^* \approx 2.479$  at which the denominator of eq. (39) vanishes and  $w_m$  diverges so that only the pole singularity can exist and the molecule is always unfolded. Furthermore from the definition of polylogarithm it follows that the derivative with respect to  $\kappa$  of the function  $\lambda(z, \kappa)$  converges only for  $c > 2$ , and since at  $z = z_b$  this has to be equal to unity, this sets also a lower bound to the critical region in  $c$ . Depending on the nature of the interaction, repulsive if  $\epsilon_0 > 0$  and attractive otherwise, the behavior of the statistical weight of base pair formation  $w = \exp(-\beta\epsilon_0)$  can be smaller or greater than unity. Then one can define the additional universal value of the loop exponent

$$c_1 \approx 2.241, \quad (40)$$

defined by

$$\zeta(c_1 - 1) - \zeta(c_1) = (\zeta(c_1 - 1) - 2\zeta(c_1))^2. \quad (41)$$

From this argument one concludes that only for  $2 < c < c^*$  a phase transition can occur between the folded and the unfolded phase, determined by  $z_b$  and  $z_p$  respectively, since only then both singularities can coexist. The corresponding phase diagram for attractive and repulsive interaction is obtained by solving  $w_m(c) = w$ , see Fig. 5.

#### IV. CONSTRAINED ANNEALING WITH LOOP ENTROPY

##### A. Outline of the method

Instead of the standard replica approach used for spin glasses [22], to compute the average over the disorder we use the constrained annealing approximation [14]. The basic idea is to perform an annealed average, where the annealed free energy is defined as

$$f^a = -\frac{1}{\beta N} \ln \overline{Z_N(h)}, \quad (42)$$

with the random variables  $\{h\}$  coupled to appropriate constraints  $\{\mu\}$ , which are functions of some disorder self-averaging variables [22]. The values of the constraints, which assume the form of Lagrange multipliers, that for  $N \gg 1$  maximize the thermodynamic potential

$$f^{ca}(\mu) = -\frac{1}{\beta N} \ln Z_N^{ca}(\mu), \quad (43)$$

are those that select the realizations with a correct value of the disorder intensive variables and at the same time that minimize the difference between the quenched free energy eq. (7) and the annealed free energy eq. (42), in which the disorder variables  $\{h\}$  are free to evolve as the dynamical degrees of freedom  $\{s\}$ . Thus  $f^{ca}(\mu)$  improves

the lower bound estimation of the Jensen's inequality [32] given by  $f^a$  for the quenched free energy so that

$$\overline{f(h)} \geq f^{ca}(\mu) \geq f^a \quad \forall \mu. \quad (44)$$

To construct the partition function [14]

$$Z_N^{ca}(\mu) = \overline{Z_N(h) e^{-N\mu\alpha(h)}}, \quad (45)$$

we define a function of the disordered sequence which self-averages to zero as

$$\alpha(h) \equiv \frac{1}{N} \sum_{i=1}^N [h_i - (2p - 1)]. \quad (46)$$

It follows immediately from

$$\bar{h} \equiv \sum_{\{h\}} \mathcal{P}(h) h = \sum_{h=\pm 1} \rho(h) h = 2p - 1, \quad (47)$$

that  $\alpha(h) \rightarrow 0$  for  $N \rightarrow \infty$ . As we will show next, since the Hamiltonian (1) is separable [25] as we consider site instead of link random variables, with this choice of  $\alpha(h)$  in the thermodynamic limit the disorder terms can be averaged independently. To see this we write eq. (5) as

$$Z_N(h) = \sum_{\{S\}} e^{-\beta\epsilon_0|\mathcal{S}|} \prod_{k < l}^N e^{-\beta\epsilon_{s_k, l} h_k h_l}, \quad (48)$$

so that the constrained annealing partition function eq. (45) becomes

$$Z_N^{ca}(\mu) = e^{N\mu(2p-1)} \sum_{\{S\}} e^{-\beta\epsilon_0|\mathcal{S}|} \Pi(\mu), \quad (49)$$

with

$$\Pi(\mu) \equiv \sum_{\{h\}} \prod_{i=1}^N \rho(h_i) e^{-\mu h_i} \prod_{k < l}^N e^{-\beta\epsilon_{s_k, l} h_k h_l}. \quad (50)$$

At this point the key observation is that since in the product we have contributions different from unity only when  $s_{k,l} = 1$  and each base can only participate in at most one base pair, in the average over the disorder we get a product of  $|\mathcal{S}| = \sum_{i < j} s_{i,j}$  times the factor  $e^{-\beta\epsilon h_k h_l}$ . This is equivalent to saying that every disorder term can be averaged independently, which follows as a direct consequence of the mutual independence of the sequence disorder variables  $h_i$ . Thus explicating the summation we obtain

$$\begin{aligned} \Pi(\mu) &= \left( \sum_{h=\pm 1} \rho(h) e^{-\mu h} \right)^{N-2|\mathcal{S}|} \times \\ &\times \left( \sum_{h, h'=\pm 1} \rho(h) \rho(h') e^{-\mu h} e^{-\mu h'} e^{-\beta\epsilon h h'} \right)^{|\mathcal{S}|} \\ &= \Omega^N(\mu) \left( \frac{\Upsilon(\mu)}{\Omega^2(\mu)} \right)^{|\mathcal{S}|}, \end{aligned} \quad (51)$$



where we have defined the two auxiliary quantities  $\Omega(\mu) \equiv pe^{-\mu} + (1-p)e^{\mu}$  and  $\Upsilon(\mu) \equiv e^{-\beta\epsilon}[p^2e^{-2\mu} + (1-p)^2e^{2\mu}] + e^{\beta\epsilon}2p(1-p)$ . This can be written in a more compact form if we define a new constant interaction energy

$$\bar{\epsilon}(\mu) \equiv -\frac{1}{\beta} \ln \frac{\Upsilon(\mu)}{\Omega^2(\mu)}, \quad (52)$$

where all the information relative to the disorder average is now included in the parameter on which to perform the variation  $\mu$ . Then eq. (49) reduces to

$$Z_N^{ca}(\mu) = e^{N\mu(2p-1)} \Omega^N(\mu) Z_N^{hom}(\mu), \quad (53)$$

where  $Z_N^{hom}(\mu) = \sum_{\{\mathcal{S}\}} [w^{ca}(\mu)]^{|\mathcal{S}|}$  is a homopolymeric partition function associated with the constrained annealing weight

$$w^{ca}(\mu) = \exp[-\beta\epsilon^{ca}(\mu)], \quad (54)$$

with pair interaction energy

$$\epsilon^{ca}(\mu) = \epsilon_0 + \bar{\epsilon}(\mu). \quad (55)$$

From its definition in the limit of vanishing disorder strength  $\epsilon \rightarrow 0$  one simply recover the homopolymeric statistical weight of eq. (22) with constant interaction energy  $\epsilon^{ca} = \epsilon_0$ . The maximisation with respect to  $\mu$  of

$$\beta f^{ca}(\mu) = -\mu(2p-1) - \ln \Omega(\mu) - \frac{1}{N} \ln Z_N^{hom}(\mu) \quad (56)$$

is achieved by imposing

$$\left. \frac{\partial}{\partial \mu} f^{ca}(\mu) \right|_{\mu=\tilde{\mu}} = 0. \quad (57)$$

This condition yields the value  $\tilde{\mu}(\beta\epsilon_0, \beta\epsilon)$  for which

$$f^{ca}(\tilde{\mu}) = \max_{\mu} f^{ca}(\mu) \approx \overline{f(h)}, \quad (58)$$

while the annealed free energy is obtained by setting  $\mu = 0$  in  $Z_N^{ca}(\mu)$ .

## V. CRITICAL BEHAVIOUR

The behaviour of the specific heat and helicity degree has been addressed recently by Hayrapetyan *et al.* [11] by solving eq. (57), where eq. (21) for  $Z_N^{hom}$  is used, with results showing a very good agreement between the quenched and constrained annealing averages. By changing the disorder strength  $\epsilon$ , they showed that in the double peak structure of  $C_V$  obtained for  $0.5 < p < 1$  and  $\epsilon_0 \neq 0$ , the low temperature jump is more pronounced than the high temperature one when  $\theta$  decreases from its maximum value as  $T \rightarrow 0$  and vice versa. Instead, a partition function with  $\epsilon_0 = 0$  yields a single peak structure for all values of the disorder strength  $\epsilon$ .

In this paper we account for loop entropy by using the scalings equations (33) and (37) with corresponding free energies

$$\beta f_{[b,p]}^{ca}(\mu) \approx -\mu(2p-1) - \ln \Omega(\mu) + \ln z_{[b,p]}(\mu), \quad (59)$$

where

$$z_p(\mu) = 2 \left( 1 + \sqrt{1 + 4w^{ca}(\mu)\zeta(c)} \right)^{-1}. \quad (60)$$

Thus a thermal phase transition is triggered by the different nature of the two singularities in the homopolymeric partition function and is therefore expected to explain physically the unusual drop of  $\theta$  at low temperatures found in [11]. By keeping implicit the expression for  $Z_N^{hom}$ , eq. (57) yields

$$\begin{aligned} 0 &= 2p-1 + \left[ \frac{\partial \ln \Omega(\mu)}{\partial \mu} + \frac{1}{N Z_N^{hom}(\mu)} \times \right. \\ &\quad \left. \times \sum_{\{\mathcal{S}\}} |\mathcal{S}| (w^{ca}(\mu))^{|\mathcal{S}|-1} \frac{\partial}{\partial \mu} w^{ca}(\mu) \right]_{\mu=\tilde{\mu}} \\ &= 2p-1 + \left[ \frac{\partial \ln \Omega(\mu)}{\partial \mu} + \frac{\langle |\mathcal{S}| \rangle}{N} \frac{\partial \ln w^{ca}(\mu)}{\partial \mu} \right]_{\mu=\tilde{\mu}} \\ &= 2p-1 + \left[ \frac{\partial \ln \Omega(\mu)}{\partial \mu} + \frac{\theta^{ca}(\mu)}{2} \frac{\partial}{\partial \mu} \ln \frac{\Upsilon(\mu)}{\Omega^2(\mu)} \right]_{\mu=\tilde{\mu}} \end{aligned} \quad (61)$$

where we have used equations (23), (52) and (55), and where

$$\begin{aligned} \theta^{ca}(\mu) &= 2 \frac{\partial}{\partial(\beta\epsilon_0)} \ln \left( \frac{1}{1 + 2\sqrt{w^{ca}(\mu)}} \right) \\ &= \frac{-2}{1 + 2\sqrt{w^{ca}(\mu)}} \left( \frac{1}{\sqrt{w^{ca}(\mu)}} \frac{\partial w^{ca}(\mu)}{\partial(\beta\epsilon_0)} \right) \\ &= \frac{2\sqrt{w^{ca}(\mu)}}{1 + 2\sqrt{w^{ca}(\mu)}}. \end{aligned} \quad (62)$$

### A. Cold melting

Once the solution of eq. (61) is known, the folded phase singularity  $z_b(\tilde{\mu})$  can be obtained by solving the two equations

$$\begin{cases} \tilde{\kappa}_b^{ca} = \frac{w^{ca}(\tilde{\mu})}{\kappa_b^{ca}} \text{Li}(c, z_b(\tilde{\mu}) \tilde{\kappa}_b^{ca}) + 1 \\ \tilde{\kappa}_b^{ca} = \frac{w^{ca}(\tilde{\mu})}{\kappa_b^{ca}} \left[ \text{Li}(c-1, z_b(\tilde{\mu}) \tilde{\kappa}_b^{ca}) - \text{Li}(c, z_b(\tilde{\mu}) \tilde{\kappa}_b^{ca}) \right] \end{cases} \quad (63)$$

where  $\tilde{\kappa}_b^{ca} \equiv \kappa(w^{ca}(\tilde{\mu}), z_b(\tilde{\mu}))$ . In the unfolded phase  $z_p(\tilde{\mu})$  is instead determined by the explicit expression eq. (60) for  $\mu = \tilde{\mu}(\beta\epsilon_0, \beta\epsilon)$ . Then the free energies in the folded and unfolded phases are computed using eq. (59). The critical weight  $w_m(c)$  defined by eq. (39) is a monotonically increasing function of the loop exponent  $c$  in the interval  $2 < c < c^*$  which defines the critical region

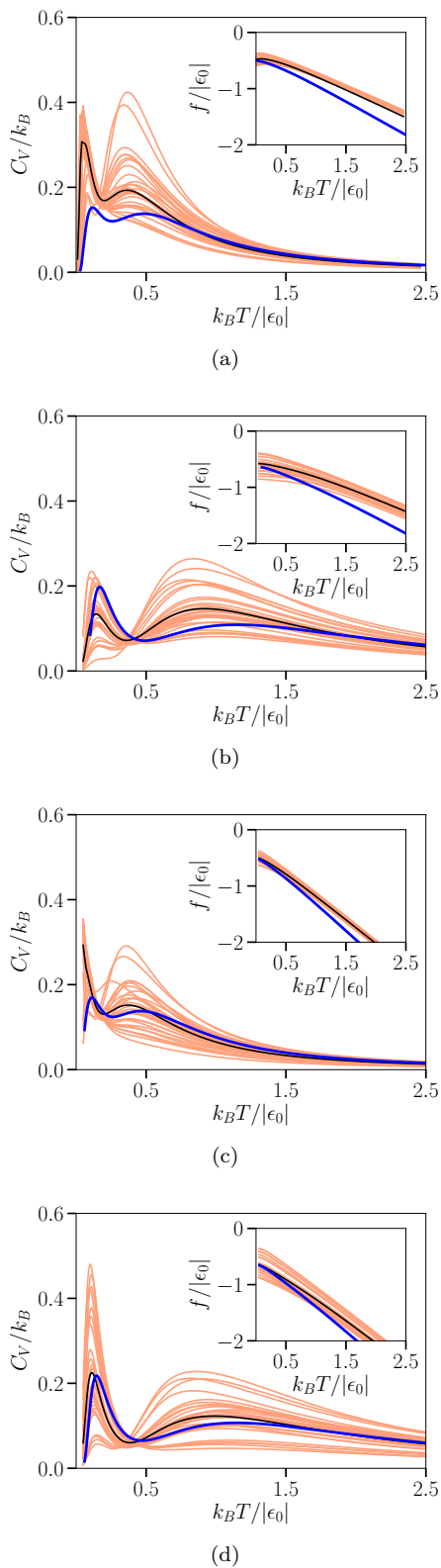


FIG. 6. (Color online) Folded phase specific heat with  $c = c_{RW} = 1.5$  (a,b) and without loop entropy  $c = 0$  (c,d). The quenched average (black line) is obtained from 30 random sequences (red lines) with  $N = 50$ ,  $p = 0.75$ , (a)  $\epsilon = 0.5|\epsilon_0|$  and (b)  $\epsilon = 1.5|\epsilon_0|$ . In blue the constrained annealing average. Insets: the corresponding free energies with quenched and constrained annealing averages from which the specific heat have been obtained.

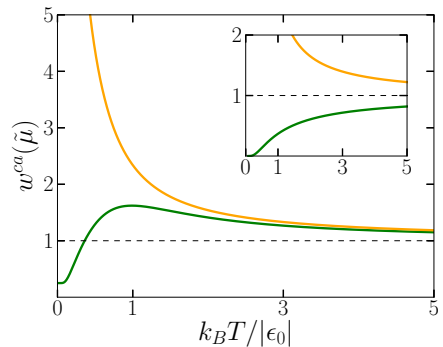


FIG. 7. (Color online) Constrained annealing weight in the competitive scenario ( $\Lambda < 1$ ) for  $p = 0.75$ ,  $\epsilon = 1.5|\epsilon_0|$  (green line) and in the non-competitive scenario ( $\Lambda > 1$ ) for  $p = 0.75$ ,  $\epsilon = 0.5|\epsilon_0|$  (orange line) as a function of the temperature. Inset: Homopolymeric weight  $w = e^{-\beta\epsilon_0}$  with  $\epsilon_0 > 0$  (green line) and  $\epsilon_0 < 0$  (orange line), for repulsive and attractive regimes respectively.

for homopolymeric RNA. For  $w > w_m$  the molecule is always folded, governed by  $z_b$ , and unfolded otherwise, governed by  $z_p$ . Then the critical temperature for the disordered model can be estimated numerically by imposing  $w^{ca}(\tilde{\mu}) = w_m(c)$  for fixed  $c$ , yielding the critical value of the Lagrange multipliers  $\tilde{\mu}_m(c)$ .

In Fig. 6 we compare quenched and constrained annealing averages in the folded phase for (a,b)  $c = c_{RW} = 1.5$  which qualitatively reproduce the behavior for  $c = 0$  in (c,d), where a double peak structure was also found in the specific heat. Here, the partition function of each sequence is computed with the loop recursive equation (12) and the constrained annealing free energy is  $f_b^{ca}(\tilde{\mu})$  of eq. (59) from which the specific heat follows as

$$C_V^{ca}(\tilde{\mu}) = \frac{k_B T}{N} \frac{\partial^2 T \ln Z_N^{ca}(\tilde{\mu})}{\partial T^2} = -T \frac{\partial^2 f^{ca}(\tilde{\mu})}{\partial T^2}. \quad (64)$$

The quenched and constrained annealing free energies show a good agreement for low temperatures while increasing  $T$  a gap arises. This energy gap however seems to increase linearly in  $T$  and does not affect the specific heat qualitatively. For  $\epsilon_0 < 0$ , which account for an attractive background interaction, we also allow the helicity to reach its maximum value when  $T \rightarrow 0$ . For the following analysis it is useful to separate the two main energetic regimes depending on whether quenched disorder has an effective relevance in the global behavior.

The fact that  $\tilde{\mu}$  is a function of  $\beta\epsilon_0$  and  $\beta\epsilon$  implies that we have two possible scenarios depending on the value of

$$\Lambda \equiv |\epsilon_0/\epsilon|. \quad (65)$$

If  $\Lambda > 1$  there is no competition between the favorable UA and unfavorable AA/UU pairs because in this case the sign of the energy of each pairing is determined only by  $\epsilon_0$  and therefore is constant. We will call this the non-competitive regime. However when  $\Lambda < 1$  the effect of quenched disorder becomes relevant since there is an

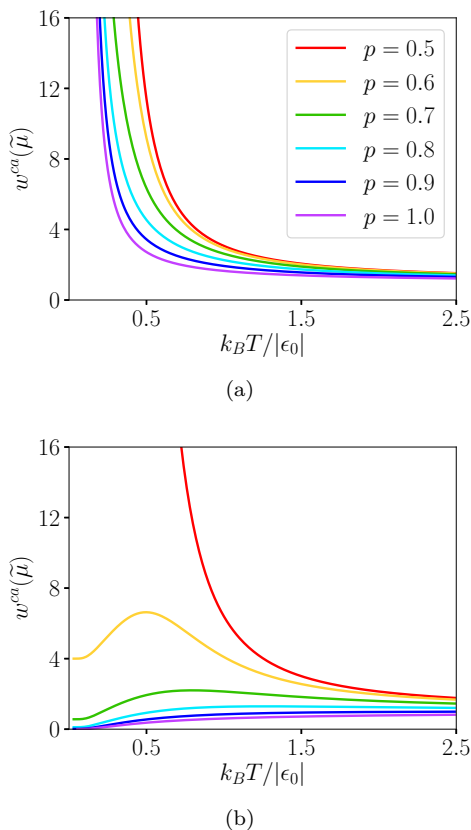


FIG. 8. (Color online) Statistical weight at various probabilities  $0.5 \leq p \leq 1.0$  in the constrained annealing approach (a) non-competitive scenario with  $\Lambda = 2$  and (b) competitive scenario with  $\Lambda = 2/3$ .

effective energetic competition between different pairings with a change of sign in the energy associated to favorable and unfavorable pairs. This case shall be referred to as the competitive regime. With  $\epsilon_0 < 0$ , although in both regimes  $\epsilon_0 - \epsilon < 0$ , in the competitive regime  $\epsilon_0 + \epsilon > 0$  and  $w^{ca}(\tilde{\mu})$  exhibits a global maximum as well as a global minimum in the low temperature range independently from the specific values of  $p$ ,  $\epsilon_0$  and  $\epsilon$ , see Fig. 7. As we will show next the presence of a global maximum in the statistical weight of base pairings is ultimately connected with the behavior of the helicity degree. When  $\Lambda < 1$  the competition between favorable and unfavorable base pairs results in the global maximum, for  $0.5 < p < 1.0$ , in  $w(\tilde{\mu})$ , see fig. 8. From the comparison between specific heat and helicity with the respective constrained annealing weights for both the competitive and non-competitive regimes it emerges that it is this competition that triggers the cold melting of the secondary structure which is described by the abrupt drop of the helicity at  $T^* \approx 0.5|\epsilon_0|/k_B$  for different probabilities of U base occurrence  $p$  with  $c = c_{RW} = 1.5$ , see Fig. 9. The first peak of the specific heat located in the range  $0 < T < T^*$  becomes more pronounced with respect to the case  $\Lambda > 1$ . Indeed each specific heat

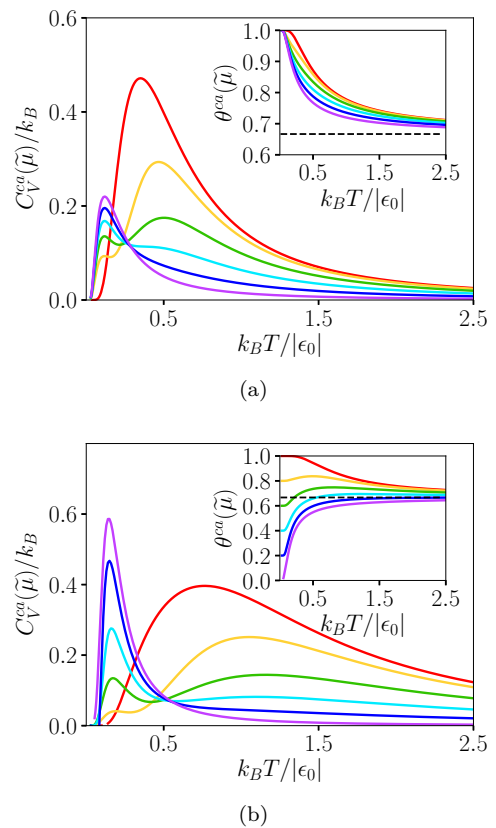


FIG. 9. (Color online) Specific heat for  $c = c_{RW} = 1.5$  for the same range of probabilities and energy parameters as in Fig. 8 in the constrained annealing approach for the non-competitive scenario (a)  $\Lambda = 2$  and competitive scenario (b)  $\Lambda = 2/3$ . Insets: Corresponding helicity degrees where the dashed black line defines the asymptotic value of eq. (25) in the high temperature molten phase.

peak corresponds to the gradual melting of the RNA secondary structure, the second of which is related to the usual hot melting. This behavior is also reproduced by setting  $\epsilon_0 = 0$ , i.e. by keeping a single degree of freedom in the Hamiltonian, although the specific heat in that scenario shows a single peak. While for  $0.5 < p < 1.0$  and  $\epsilon_0 \neq 0$  the specific heat features always two peaks, also in the limiting cases  $p = 0.5$  and  $p = 1.0$ , which correspond to annealed and homopolymeric case respectively, it exhibits a single peak in each of the two different temperature regions.

## B. Global phase diagrams

In the competitive regime in addition to  $c_1$ , defined by eq. (41), it is useful to define  $c_{max}$  and  $c_{min}$  respectively as

$$w_m(c_{max}) \equiv \max_T w^{ca}(T), \quad (66)$$

$$w_m(c_{min}) \equiv w^{ca}(T = 0). \quad (67)$$

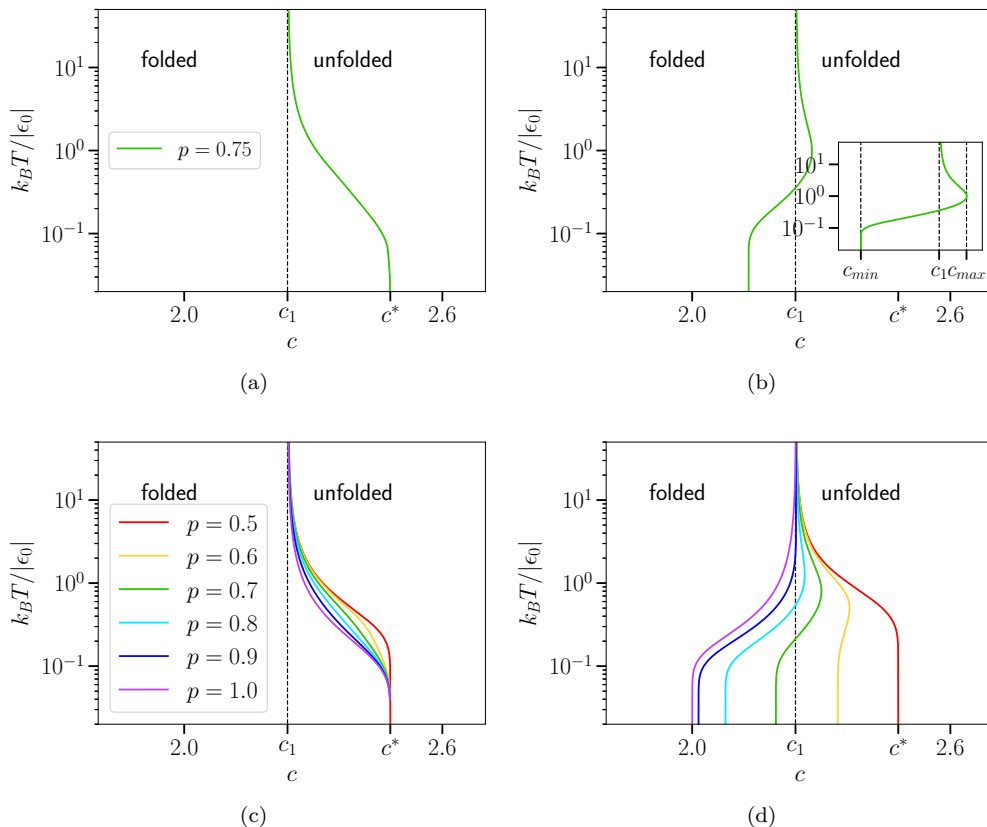


FIG. 10. (Color online) (a) Phase diagram for  $\Lambda = 2$  and  $p = 0.75$  in the  $T - c$  plane. The critical line on which hot melting takes place is contained in the range  $c_1 < c < c^*$  of  $w_m(c)$  so that for  $c < c_1$  the molecule is always folded. (b) Phase diagram for  $\Lambda = 2/3$  and  $p = 0.75$  in the  $T - c$  plane. For  $c < c_{min}$  the molecule is always folded since  $w_m(c) < w^{ca}(T)$ ,  $\forall T$ . For  $c_{min} < c < c_{max}$  there is only hot melting while for  $c > c_{max}$  the molecule is always unfolded since  $w_m(c) > w^{ca}(T)$ ,  $\forall T$ . In the range  $c_1 < c < c_{max}$  the double intersection between  $w_m(c)$  and  $w^{ca}(T)$  determines also the onset of cold melting. Lower inset: Close up of the reentrant melting cross line. In (c) and (d) are shown the complete phase diagrams for  $\Lambda = 2$  and  $\Lambda = 2/3$  for the range of probabilities from red to blue  $p = 0.5$ ,  $p = 0.6$ ,  $p = 0.7$ ,  $p = 0.8$ ,  $p = 0.9$  and  $p = 1.0$ .

Then from the behaviour of  $w^{ca}$  in the range  $0.5 < p < 1.0$  there are only four separate cases for the state of the molecule, depending on the value of the loop exponent. In fact there can be a double, single or no intersection at all between the values of  $w^{ca}$  and  $w_m(c)$  for each value of the loop exponent  $c$ . If  $p > p^*$ , where  $p^*$  is defined by

$$w_m(c_{min}(p^*)) = 1, \quad (68)$$

the most interesting case appears for  $c_1 < c < c_{max}$ , where  $c_{max}$  depends on  $p$ ,  $\epsilon_0$  and  $\epsilon$  while  $c_1 \approx 2.241$  is universal. Analogously if  $p < p^*$  one has  $c_{min}(p) > c_1$  and the interesting range becomes  $c_{min} < c < c_{max}$  as can be seen from the  $p$ -dependence of  $w^{ca}$  in Fig. 8. In this scenario  $w_m(c)$  intersects  $w^{ca}(\tilde{\mu})$  at two different temperatures  $T_{cm}$  and  $T_{hm}$ . In addition to the high temperature melting also the cold melting found in [11] takes place and manifests itself as a proper thermodynamic phase transition, where each phase is characterized by the relevant singularity of the homopolymeric grand-canonical partition function. By contrast in the non-competitive regime, as for the homopolymer, only the hot melting transition

takes place. By fixing the background interaction  $\epsilon_0 < 0$  and the probability at  $p = 0.75$ , we solve the equation

$$w^{ca}(\tilde{\mu}) = w_m(c), \quad (69)$$

which yields the values  $\tilde{\mu}_m(\beta\epsilon_0, \beta\epsilon, c)$  of the critical line in the phase diagram of Fig. 10 (a,b) corresponding to  $\Lambda = 2$  and  $\Lambda = 2/3$  for  $\epsilon = 0.5|\epsilon_0|$  and  $\epsilon = 1.5|\epsilon_0|$  respectively. The reentrant melting point is defined by the crossing with the asymptote  $c = c_1$  in the case  $\Lambda < 1$  (b). In figures 10 (c) and (d) we show the global phase diagram for the complete range of the probabilities  $0.5 \leq p \leq 1.0$  in the competitive and non-competitive scenarios.

The  $c$ -dependence in the folded phase of the specific heat is very subtle for both the disordered model, where there is a two peak structure, and the homopolymeric model where  $C_V$  features only one peak, see Fig. 11 (a,b). A similar non significant spread is obtained for the helicity behavior in Fig. 11 (c,d), where for the folded phase we use  $\theta_b$  defined by eq. (34) with  $\kappa_b = \tilde{\kappa}_b^{ca}$  and  $z_b = z_b(\tilde{\mu})$ , while in the unfolded phase  $\theta_p$  defined by eq. (38) with  $w = w^{ca}(\tilde{\mu})$ . The cold and the hot melting

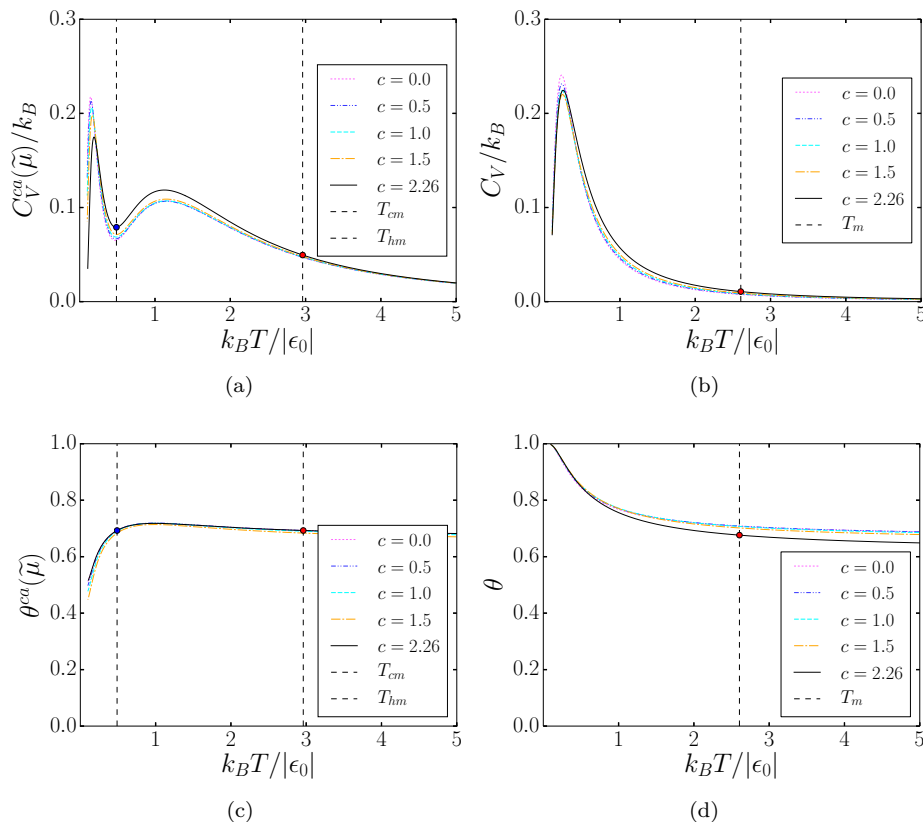


FIG. 11. (Color online) Constrained annealing specific heat in (a) the disordered model for  $p = 0.75$  and  $\epsilon = 1.5|\epsilon_0|$  resulting in  $\Lambda = 2/3$  and (b) the corresponding homopolymeric model for  $\epsilon_0 < 0$  at various values of the loop exponent  $c$  in the folded phase. In the disordered model the two critical temperatures  $T_{cm} \approx 0.488|\epsilon_0|/k_B$  (blue circle) and  $T_{hm} \approx 2.961|\epsilon_0|/k_B$  (red circle) correspond to the value of the loop exponent  $c = 2.26$  while for the homopolymer there is only  $T_m \approx 2.605|\epsilon_0|/k_B$ . In (c) and (d) is shown the corresponding helicity.

temperatures are estimated numerically for  $c = 2.26$ ,  $p = 0.75$ , and  $\epsilon = 1.5|\epsilon_0|$  as  $T_{cm} \approx 0.488|\epsilon_0|/k_B$  and  $T_{hm} \approx 2.961|\epsilon_0|/k_B$  respectively. Quite remarkably the critical point for the reentrant transition lies almost exactly in the valley formed by the two peaks of the specific heat relative to the two different melting transitions.

## VI. CONCLUSIONS

In a two-letter RNA model without loop entropies, where no phase transition occurs, the heat capacities were shown to exhibit two peaks for the case of a quenched random sequence [11], indicating hot as well as cold melting. For homopolymeric RNA, on the other hand, it is known that a finite loop entropy lead to a phase transition between a folded and an unfolded RNA state for a small range of the loop exponent  $c$ . In the present paper we combine the main features of these two models and consider a two-letter RNA model with quenched randomness and in the presence of a finite loop entropy. As a main result, we show that for a small range of the loop exponent  $c$  two phase transitions are encoun-

tered with changing the temperature, i.e. the folded state is only stable at intermediate temperatures. This might be related to the experimental observation of cold melting of RNA. With the present work by combining loop entropy with quenched randomness, which introduces energetic competition between different matching in the base pairs, we are able reproduce the cold denaturation phenomena and to describe it in the language of statistical mechanics and phase transitions. Only for competing energies between favorable and unfavorable base pairs this transition occurs, as a result of the weakening of the secondary structure formation due to quenched randomness in RNA sequences, as well as loop penalties which account for the existence of two relevant singularities in the grand-canonical partition function. The connection between the competitive regime and cold denaturation is investigated by comparing the constrained annealing weight and helicity with the specific heat peaks in the low temperature range. We argue that this transition is continuous and more specifically, since it is triggered by the same conformational effect of the homopolymeric counterpart, of order  $n$ , where  $n$  is determined by  $(c - 2)^{-1} - 1 < n < (c - 2)^{-1}$  [12]. For the examined

case with  $c = 2.26$  the reentrant melting transition is of third order and would become visible only by looking at higher order derivatives of the specific heat. Finally, a

particularly interesting direction for future works is to investigate the connection between cold denaturation and the glass transition for RNA molecules.

- 
- [1] R. F. Gesteland, T. R. Cech and J. F. Atkins, *The RNA World*, (Cold Spring Harbor Laboratory Press, 2005)
- [2] I. Tinoco, O. C. Uhlenbeck and M. D. Levine, *Nature* **230**, 362 (1971)
- [3] P. G. Higgs, *Physical Review Letters* **76**, 704 (1996)
- [4] B. Duplantier, *Physical Review Letters* **57**, 941 (1986)
- [5] Y. Kafri, D. Mukamel and L. Peliti, *Physical Review Letters* **85**, 4988 (2000)
- [6] P. L. Privalov, *Critical Reviews in Biochemistry and Molecular Biology* **25**, 281 (1990)
- [7] P. J. Mikulecky and A. L. Feig, *Journal of the American Chemical Society* **124**, 890 (2002)
- [8] P. J. Mikulecky and A. L. Feig, *Nucleic Acids Research* **32**, 3967 (2004)
- [9] A. V. Finkelstein, *Protein Physics*, (Academic Press, 2002)
- [10] F. Sedlmeier, D. Horinek and R. R. Netz, *Journal of Chemical Physics* **134**, 055105 (2011)
- [11] G. N. Hayrapetyan, F. Iannelli, J. Lekscha, V. F. Morozov, R. R. Netz and Y. Sh. Mamasakhlisov, *Physical Review Letters* **113**, 068101 (2014)
- [12] T. R. Einert, P. Näger, H. Orland and R. R. Netz, *Physical Review Letters* **101**, 048103 (2008)
- [13] K. A. Dill, *Biochemistry* **24**, 1501 (1985)
- [14] M. Serva and G. Paladin, *Physical Review Letters* **70**, 105 (1993)
- [15] D. J. Amit and H. Gutfreund, *Physical Review A* **32**, 1007 (1985)
- [16] J. McCaskill, *Biopolymers* **29**, 1105 (1990)
- [17] R. Bundschuh and T. Hwa, *Physical Review E* **65**, 031903 (2002)
- [18] G. E. Andrews, R. Askey and R. Roy, *Special Functions*, (Cambridge University Press, 1999)
- [19] T. R. Einert, H. Orland and R. R. Netz, *European Physical Journal E* **34**, 55 (2011)
- [20] T. R. Einert, and R. R. Netz, *Biophysical Journal* **100**, 2745 (2011)
- [21] P-G. de Gennes, *Scaling Concepts in Polymer Physics*, (Cornell University Press, 1979)
- [22] M. Mezard, G. Parisi and M.A. Virasoro, *Spin Glass Theory and Beyond*, (World Scientific, 1988)
- [23] A. Pagnani, G. Parisi and F. Ricci-Tersenghi, *Physical Review Letters* **84**, 2026 (2000)
- [24] D. Poland and H. A. Scheraga, *Journal of Chemical Physics* **45**, 1464 (1965)
- [25] J. Bascle, T. Garel and H. Orland, *Journal de Physique I France* **3**, 259 (1992)
- [26] P-G. de Gennes, *Biopolymers* **6**, 715 (1968)
- [27] R. Bundschuh and T. Hwa, *Physical Review Letters* **83**, 1479 (1999)
- [28] C. Monthus and T. Garel, *Physical Review E* **75**, 031103 (2007)
- [29] R. Nussinov and A.B. Jacobson, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6309 (1980)
- [30] M. Mézard and A. Montanari, *Information, Physics, and Computation* (Oxford University Press, 2009)
- [31] P. De Los Rios and G. Caldarelli, *Physical Review E* **63**, (2001) 031802
- [32] M. Mezard and A. Montanari, *Information, physics, and computation*, (Oxford University Press, 2009)
- [33] P.L. Privalov, Yu. V. Griko, S. Yu. Venyaminov, and V.P. Kutysenko, *J. Mol. Biol.* **190**, 487 (1986)
- [34] N.C. Pace and C. Tanford, *Biochemistry* **7**, 198 (1968)