Path integral method for DNA denaturation

Marco Zoli[*]

Dipartimento di Fisica, Universitá di Camerino, Camerino I-62032, Italy
(Received 30 October 2008; revised manuscript received 1 March 2009; published 29 April 2009)

The statistical physics of homogeneous DNA is investigated by the imaginary time path integral formalism. The base pair stretchings are described by an ensemble of paths selected through a macroscopic constraint, the fulfillment of the second law of thermodynamics. The number of paths contributing to the partition function strongly increases around and above a specific temperature T_c^*, whereas the fraction of unbound base pairs grows continuously around and above T_c^*. The latter is identified with the denaturation temperature. Thus, the separation of the two complementary strands appears as a highly cooperative phenomenon displaying a smooth crossover versus T. The thermodynamical properties have been computed in a large temperature range by varying the size of the path ensemble at the lower bound of the range. No significant physical dependence on the system size has been envisaged. The entropy grows continuously versus T while the specific heat displays a remarkable peak at T_c^*. The location of the peak versus T varies with the stiffness of the anharmonic stacking interaction along the strand. The presented results suggest that denaturation in homogeneous DNA has the features of a second-order phase transition. The method accounts for the cooperative behavior of a very large number of degrees of freedom while the computation time is kept within a reasonable limit.

I. INTRODUCTION

Path integral methods provide a powerful tool to investigate the properties of nonlinear dynamical systems with retarded interactions. Hamiltonian models with nonlocal electron-phonon couplings, used in polymer science [1], have been studied by such methods. In this paper the path integral approach [2] is applied to analyze the temperature-driven DNA denaturation which occurs when the double stranded molecule separates into two coils while, at lower T, only localized openings exist [3]. Modeling of DNA melting has been motivated since long by the need to understand the transcription mechanism in which the double helix opens locally to allow a reading of the genetic code. In this regard, the Poland-Scheraga model [4] has been path breaking and still makes the base of several investigations [5–10]. Despite extensive analytical and numerical work [11–16], establishing character and nature of the melting phase transition, whether first or second order, remains a challenging task [17]. A significant advance toward a comprehension of the DNA dynamics was made after Peyrard and Bishop (PB) introduced a one-dimensional (1D) Hamiltonian model [18] which recognized the role of nonlinearities in the molecule [19] and reduced its great complexity to two essential interactions: (i) a nonlinear coupling between the two bases in a pair connected by hydrogen bonds, (ii) a *harmonic* stacking potential between adjacent bases along the strand. While solutions based on the transfer integral method [20] predicted a smooth denaturation transition both for homogeneous and heterogeneous DNA [21], extensions of the PB model [22,23] which include *anharmonic* stacking interactions showed that the denaturation can be sharp, occurs at lower T than in the harmonic case, and it is indeed a thermodynamic transition [24]. The latter statement does not contradict general theorems on the impossibility of phase transitions in 1D

systems [25] as the PB Hamiltonian contains an on-site potential depending on unbounded transverse stretchings. Moreover, in the thermodynamic limit, the energy of a domain wall between two states of the molecule is infinite and an ordered state is then possible [26].

As the anharmonic PB model is assumed as fundamental tool, it is understood that the emphasis of the present investigation is purely on modeling the unbinding of the two strands versus temperature. This is done by capturing the essentials of the interactions consistently with the spirit of the theory of critical phenomena.

The PB Hamiltonian model is outlined in Sec. II focusing on those properties which make reliable to attack the problem by path integral techniques. Section III describes the method while the results concerning the character of the denaturation transition are discussed in Sec. IV. Some conclusions are drawn in Sec. V.

II. HAMILTONIAN MODEL

To begin, I consider a fragment of homogeneous DNA whose strands are represented by a set of point masses corresponding to the nucleotides. The Hamiltonian [22] for a chain of N base pairs (bps) with mass μ reads

$$H = \sum_{n=1}^{N} \left[\frac{\mu \dot{y}_n^2}{2} + V_S(y_n, y_{n-1}) + V_M(y_n) \right],$$

$$V_S(y_n, y_{n-1}) = \frac{K}{2} g(y_n, y_{n-1})(y_n - y_{n-1})^2,$$

$$g(y_n, y_{n-1}) = 1 + \rho \exp[-\alpha(y_n + y_{n-1})],$$

$$V_M(y_n) = D[\exp(-a y_n) - 1]^2, \tag{1}$$

where y_n, the transverse stretching for the n-bp, measures the relative pair separation from the ground-state position and

*marco.zoli@unicam.it

continuously grows from 0 (closed bp) to ∞ (open bp after denaturation occurs). As the hydrogen bond can be compressed with respect to the equilibrium, y_n takes also negative values with a lower cutoff accounting for the repulsive interaction of the phosphate groups. The longitudinal displacements are neglected in the model as their amplitudes are much smaller than the transverse stretchings [3,27]. D and a, setting depth and width of the Morse potential $V_M(y_n)$, are site independent for *homopolymer* DNA. In the stacking potential $V_S(y_n, y_{n-1})$, the factor g depends on the sum of the stretchings of two adjacent bases. $K = \mu \nu^2$ with ν being the harmonic phonon frequency. For $\rho = 0$ in Eq. (1), the statistical mechanics can be worked out exactly through the transfer operator method [20], which, in continuum approximation and strong K limit, defines a pseudo-Schrödinger equation for a particle in a Morse potential [28]. The latter has a discrete spectrum with localized eigenfunctions at $T < T_c$ being

$$T_c = \frac{2\sqrt{2KD}}{K_B a}, \tag{2}$$

the temperature at which the lowest eigenvalue vanishes and even the ground state merges into the continuum. K_B is the Boltzmann constant. For $T > T_c$ only delocalized states exist. Thus, even the simplest harmonic stacking model shows that a transition occurs and T_c is identified as the temperature at which the strands separate. Taking typical values for DNA parametrization [3,29] such as $D = 30$ meV, $a = 4.2$ Å^{-1} and $K = 60$ meV Å^{-2}, one gets $T_c = 331$ K. Anharmonicity ($\rho > 0$) has a special meaning in DNA stacking. While the analytical form of g may not be unique in order to reproduce the denaturation transition [30], the key property of the stacking potential is the following: whenever either one of the bps is stretched over a distance larger than α^{-1}, the hydrogen bond breaks and the electronic distribution around the two pair mates is modified. Accordingly, the stacking coupling (along each strand) between neighboring bases in Eq. (1) drops from $K(1+\rho)$ to K. Then, also the next bp tends to open as both bases are less closely packed along their respective strands. Here is the *cooperative* character which underlines the formation of a region with open bps, a bubble whose size increases with T. In heterogeneous DNA, also the length of the sequence may affect the probability to open a bubble which is quantified in polymer network theories by the *cooperativity parameter* σ: for sequences of intermediate length (up to $\sim 10^4$ bps), a small σ ($\sim 10^{-4} - 10^{-5}$) generally suppresses bubble formation while large portions of the helix unbind close to the melting making the transition highly cooperative [31–33]. For long sequences, σ is larger and small bubbles may form already below the transition which accordingly is expected to be less cooperative.

When the stacking decreases the vibrational mode between the two bases softens, thus reducing its contribution to the free energy. Then, as confirmed by molecular-dynamics simulations [28], the denaturation onset is signaled by a phonon mode softening which, in general [34,35], may point to the occurrence of a phase transition. For these reasons, anharmonic stacking models well account for cooperativity effects in the formation of open domains. Being the latter a collective phenomenon, computational methods have to include hundreds of base pairs in order to study the dynamics even of a molecule fragment and this requires considerable computational power. These facts suggest that models at intermediate scales are important and motivate the idea to investigate the DNA denaturation by path integral techniques.

III. PATH INTEGRAL METHOD

The path integral formalism defines the time (t) evolution amplitude between two points, say "a" and "b," as a sum over *all* histories along which a system can evolve in going from "a" to "b." Each history is weighed by a phase factor, the exponential of the action associated to a given path [36]. At finite temperature, the thermal properties of the system can be derived by weighing the contributions by the particle paths $x(\tau)$ running along the imaginary time axis τ, after an analytic continuation is performed: $\tau = it$. Accordingly, in the statistical formalism, the Euclidean action $A[x(\tau)]$ replaces the mechanical canonical action and the partition function is an integral in the path phase space. Each path is weighed by a probability factor $\exp\{-A[x(\tau)]\}$. Only closed paths contribute to the statistical partition function, the integration being a trace integration [37]. In the calculations, not *all* histories can be accounted for and, given the specific problem, one has to select the suitable class of paths which mainly contribute (are expected to contribute) to the physical properties.

For the model in Eq. (1), the path integral description naturally follows by replacing the bps transverse stretching by a one dimensional path x which is continuous function of τ. Accordingly, the space interaction are mapped onto the time scale as

$$y_n \rightarrow x(\tau), \quad y_{n-1} \rightarrow x(\tau'), \quad \tau' = \tau - \Delta\tau, \tag{3}$$

with $\tau \in [0, \beta]$ and β being the inverse temperature. While only adjacent bases are taken in Eq. (1), the stacking interactions range can be varied in the path integral model by tuning the retardation $\Delta\tau$ in Eq. (3). The discrete lattice nature of the Hamiltonian is maintained by the replacement in Eq. (3). In fact, for any β, the computation is made convergent by taking a large but finite number N_τ of points in the τ range. The path $x(\tau)$ is further expanded in Fourier series with cutoff M_F

$$x(\tau) = x_0 + \sum_{m=1}^{M_F} [a_m \cos(\omega_m \tau) + b_m \sin(\omega_m \tau)] \tag{4}$$

and $\omega_m = 2m\pi / \beta$. As $x(0) = x(\beta)$, periodic boundary conditions hold for the present method analogously to those imposed for the 1D finite chain described by Eq. (1). The heart of the matter lies in summing over a suitable set of paths such that the computation of the thermodynamical properties becomes sufficiently accurate. As explained hereafter, the selection of the paths ensemble mainly contributing to the partition function is done on the base of physical constraints. By Eq. (3), the index n maps onto a single τ value and, by Eq. (4), for any τ an ensemble of path coefficients is taken in the computation.

The imaginary time partition function [2] is given by

$$Z = \oint \mathfrak{D}x(\tau)\exp\{-A[x(\tau)]\},$$

$$A(x(\tau)) = \int_0^\beta d\tau \left\{ \frac{\mu}{2}\dot{x}(\tau)^2 + V[x(\tau)] \right\},$$

$$\oint \mathfrak{D}x(\tau) \equiv \frac{1}{\sqrt{2}\lambda_\mu} \int dx_0 \prod_{m=1}^{M_F} \left(\frac{m\pi}{\lambda_\mu} \right)^2 \int da_m \int db_m, \quad (5)$$

where $A[x(\tau)]$ is the Euclidean action for a particle in the potential $V[x(\tau)]$. $\mathfrak{D}x(\tau)$ is the measure of integration which normalizes the free particle action

$$\oint \mathfrak{D}x(\tau)\exp\left[-\int_0^\beta d\tau \frac{\mu}{2}\dot{x}(\tau)^2 \right] = 1 \quad (6)$$

and $\oint$ denotes integration over closed particle trajectories consistently with periodic boundary conditions holding for paths in Eq. (4).

λ_μ is the thermal wavelength whose form depends on the model whether quantum or classical. DNA denaturation occurs at temperatures for which a classical model applies. Then, the time derivative $\dot{y}_n$ [Eq. (1)] maps onto the imaginary time derivative $\dot{x}(\tau)$ [Eq. (5)], the proper replacement being: $d/dt \rightarrow (\nu\beta)d/d\tau$, hence, $\lambda_\mu = \sqrt{\pi/\beta K}$. Note that the pseudo-Schrödinger equation is also solved in a classical framework with $\hbar$ replaced by $(\nu\beta)^{-1}$.

Applying Eq. (3) to the Hamiltonian in Eq. (1), Z in Eq. (5) transforms into

$$Z = \frac{1}{\sqrt{2}\lambda_\mu} \int_{-U_0}^{U_0} dx_0 \prod_{m=1}^{M_F} \left[\frac{1}{\pi} \int_{-U}^{U} ds_m \int_{-U}^{U} dt_m \right.$$

$$\left. \times \exp(-s_m^2 - t_m^2) E(x_0, s_m, t_m) \right],$$

$$E(x_0, s_m, t_m) = \exp\left[-\int_0^\beta d\tau V_{\Delta\tau}(x) \right],$$

$$V_{\Delta\tau}(x) = V_M(x) + V_S(x, x'),$$

$$V_M(x) = D[\exp(-ax) - 1]^2,$$

$$V_S(x, x') = \frac{K}{2}g[x, x'](x - x')^2,$$

$$g[x, x'] = 1 + \rho \exp[-\alpha(x + x')],$$

$$x \equiv x(\tau); \quad x' \equiv x(\tau'),$$

$$s_m^2 \equiv \frac{m^2\pi^3 a_m^2}{\lambda_\mu^2}; \quad t_m^2 \equiv \frac{m^2\pi^3 b_m^2}{\lambda_\mu^2}, \quad (7)$$

where the measure has been written in a form suitable for computation by introducing the cutoffs U_0 and U on the Fourier coefficients integrals [38]. By inserting the expansion in Eq. (4), the left-hand side of Eq. (6) transforms into a product of Gaußian integrals. Normalization of the latter gives the mathematical criterion to set U_0 and U through the conditions

$$\frac{1}{\sqrt{2}\lambda_\mu} \int_{-U_0}^{U_0} dx_0 = 1,$$

$$\frac{1}{\sqrt{\pi}} \int_{-U}^{U} ds_m \exp(-s_m^2) = 1. \quad (8)$$

While U_0 is set by the first in Eq. (8), U can be determined after a series expansion for the Gaußian integral [39] is applied to the left-hand side in the second of Eq. (8). However, there is no *a priori* reason why the mathematical cutoffs had to fulfill also the physical requirements of the problem in Eq. (7). In fact, as extensively discussed below, too negative Fourier coefficients would induce large negative path amplitudes and diverging V_M which, in turn, would yield zero contribution to Z. Such paths, although allowed by Eq. (8), have therefore to be discarded as their inclusion in the computation would produce a decreasing entropy versus T. While this event cannot be accepted on general physical grounds, it also provides clue to solving consistently the cutoff problem: imposing the second law of thermodynamics, I solve the path integral in Eq. (7) selecting at any T the suitable path ensemble for the transverse stretchings. Let us see in detail the technicalities of the method.

Besides the five *model potential* parameters, namely, D, a, K, ρ, and α, the path integral method contains some intrinsic input parameters. The latter are (a) the cutoff M_F, (b) the retardation $\Delta\tau$, (c) the number of points N_τ for the $d\tau$ integral in Eq. (7), (d) the number of integration points over the Fourier coefficients x_0, s_m, and t_m.

(a) $M_F = 1$ suffices to make the calculation convergent in the sense that Z would not change by taking $M_F = 2$. This result can be understood as follows: the inclusion of higher Fourier components leads to larger (absolute) values for the path amplitudes but V_M *takes care of suppressing paths that leave the reasonable boundaries*. While there is no reason to include paths whose fate is that of being suppressed, a proper ensemble of paths can be built by taking a sufficiently high number of integration points for the first Fourier component in a sufficiently broad range. This ensures that the program samples a large portion of the phase space in which all the relevant path amplitudes are accounted for. Incidentally, the ($M_F = 1$) strategy also permits to save a great amount of computing time.

(b) This investigation is restricted to adjacent bps interactions along the strand. Accordingly, τ and τ' in Eq. (3) are first neighbors in the discrete imaginary time lattice and the retardation is $\Delta\tau = \beta/N_\tau$.

(c) $N_\tau = 100$ suffices to make the action numerically stable. This value also sets the number of bps as N_τ coincides with N in Eq. (1) due to the mapping in Eq. (3). While $N = 100$ is a (minimal) reliable choice for a cooperative phenomenon such as denaturation, it turns out that larger N_τ (N) would not change anything in the computation of Eq. (7). At
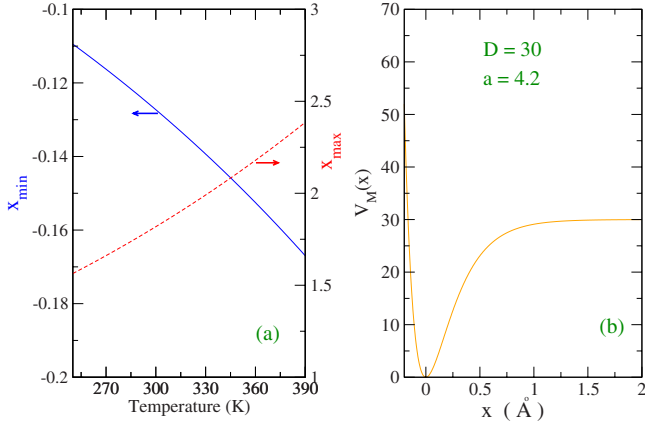
FIG. 1. (Color online) (a) Lower and upper cutoffs (in Å) for the paths, representing the base pairs transverse stretchings, versus temperature; (b) Morse potential V_M (in meV) versus path amplitude x. D is in meV and a in Å^{-1}.

this stage, the reader may wonder whether and how the path integral method can account for those finite-size effects [40,41], which are considered relevant in studies of the DNA properties. The fact is that for the present method, N is not the exclusive tuning parameter in order to set the system *size*, the latter being rather determined by the total number of paths allowed for the N-bps. Thus, although N is kept constant, the system *size* may still increase/decrease by considering a larger/smaller path ensemble which eventually works as the *true scaling parameter* in the path integral method. Here, we see the deep meaning of the method underlied by the mapping in Eq. (3). The n-bp maps onto an effective ensemble of paths N_{eff} taken at a specific τ. The latter may be viewed in itself as an ensemble of base pairs. As this holds for any of the N_τ points, the *true scaling parameter* turns out to be $N_\tau \cdot N_{eff}$.

(d) Taking one Fourier component, the program evaluates at any T the contribution to Z by three integrals over Fourier coefficients in Eq. (7). If, for a given number of integration points, the increasing entropy constraint is fulfilled then the *good paths* are included in the computation. The number of *good paths* for a single τ is N_{eff}. The total number of *good paths* over which computations are carried out is $N_\tau \cdot N_{eff}$.

Clearly, the computing time depends on the number of paths required to stabilize the system. Many paths should contribute to the thermodynamical properties consistently with the previous observations regarding the cooperativity behavior in the double helix. Such number is expected to grow markedly around and above the transition. This would be a consistent hallmark for the reliability of the path integral approach to the DNA Hamiltonian.

IV. DISCUSSION

The issue of the maximum (absolute) path amplitudes is key in understanding the double helix denaturation [42] and it has to be discussed consistently with the available experimental data. As put forward above, the properties of the potential V_M, plotted in Fig. 1(b), suggest which path ensemble

is physically meaningful.

Hydrogen bond stretchings are generally broken for lengths ~ 2 Å. For such separation, the potential is flat with a plateau value $D \sim 30$ meV and the force between the bases vanishes. Accordingly, larger amplitudes would yield no contribution. Then, the bp breaking energy D is $\sim K_B T$ above room temperature. This justifies the parameter values used in Fig. 1(b). Largely positive paths would sample the flat portion of V_M and add constant terms to the action which do not affect the free-energy derivatives.

On the other side, steric hindrance processes prevent $x(\tau)$ from being too negative. Paths $x(\tau) \sim -0.14$ Å also experience a value $D \sim 30$ meV while larger negative paths encounter the hard-core repulsive side of V_M. Here the latter grows exponentially, thus yielding a large action which, in turn, makes a vanishing contribution to Z.

These are the intuitive reasons to set a finite range which inspire the computational method described in Sec. III. Moreover, it is physically plausible that such range had to be modulated by temperature effects, allowing for larger path amplitudes at higher T. Saying T_c^* the denaturation temperature, the path displacements may reasonably vary in the range $x(\tau) \in [x_{min}, x_{max}]$ with $x_{min} \sim -0.14$ Å and $x_{max} \sim 2$ Å around the crossover. The functions $x_{min} = -0.14 \exp[(T - T_c^*)/T_c^*]$ and $x_{max} = 2 \exp[(T - T_c^*)/T_c^*]$ provide suitable (although not unique) expressions to account for linear temperature dependence in a window around T_c^*, whereas larger deviations occur outside that window. x_{min} and x_{max} are plotted in Fig. 1(a) taking for T_c^* the value given by Eq. (2). However, this assumption does not imply that the melting occurs precisely at T_c. In fact the true denaturation temperature (found below) does not coincide with the above given T_c which stems from an analytic harmonic model. Moreover, the specific choice for upper and lower bounds has not to imply in principle that a transition exists nor it has to force the transition to take place at T_c^*. Certainly, if the denaturation occurs, a significant fraction of bps is expected to be ~ 2 Å. This feature will be hereafter investigated.

Thus, after setting a temperature value, the program searches for those paths $\{x(\tau), x(\tau')\}$ which lie inside the range $[x_{min}, x_{max}]$ for any $\tau \in [0, \beta]$ and adds their contribution to Z. Such a check is done for any set of Fourier coefficients which in fact defines a specific path. Had the entropy not to have a positive derivative, the computation should be rerun for a new *number of integration points* which, in turn, leads to rebuild the ensemble of Fourier coefficients and to repeat the procedure to filter the suitable paths. It is seen numerically that the restriction due to x_{max} can be lifted without affecting the physical properties whereas the lower bound x_{min} is required in the computational method.

As a final output N_{eff} is obtained versus T as shown in Fig. 2(a) for the same values D, a, K as in Figs. 1 and for three stacking parameters ρ with $\alpha = 0.35$ Å^{-1} [29]. While all plots begin with $N_{eff} \simeq 8000$ at $T = 250$ K, their slopes display a remarkable dependence on ρ and, as a general feature, N_{eff} is higher for stronger anharmonicities. In particular, for the case $\rho = 2$, N_{eff} increases at $T \sim 315$ K. This feature is mirrored in the plot [Fig. 2(b)] of the specific heat which is a significant hallmark for denaturation as it is proportional to the differential melting curve [43]. For $\rho = 2$, a peak is found
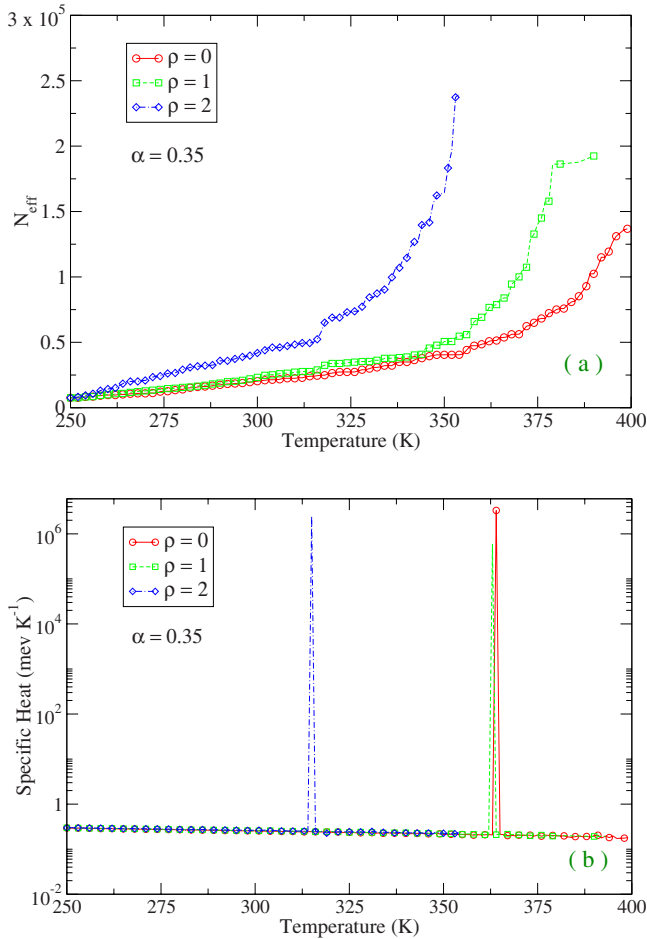
FIG. 2. (Color online) (a) Effective number of paths contributing to the partition function for three stacking parameters. α in Å^{-1}. (b) Specific heat computed via Eq. (7).
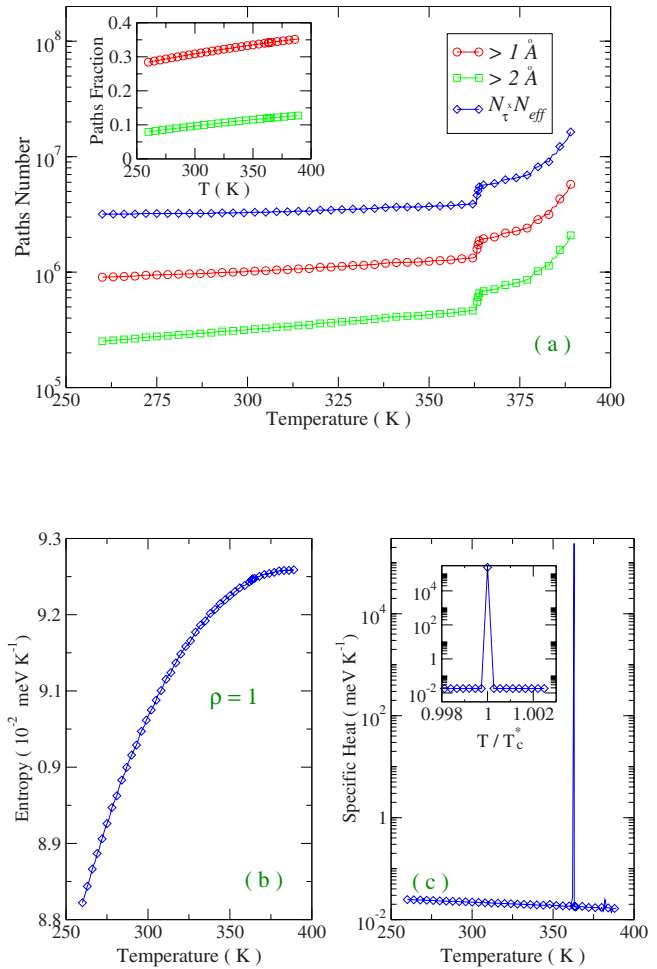

FIG. 3. (Color online) Anharmonic stacking case $\rho=1$. $N_{\text{eff}}=31792$ at $T=260$ K is assumed. T axis starts with a 10 K shift with respect to Fig. 2 in order to distinguish among the three plots corresponding to different ρ values. (a) Paths larger than 1 Å, 2 Å and *total number* of paths ($N_\tau \cdot N_{\text{eff}}$) are plotted. Inset: ratios of paths whose amplitude is larger than 1 Å and 2 Å, respectively, over $N_\tau \cdot N_{\text{eff}}$. (b) Entropy versus temperature. (c) Specific heat versus temperature. Inset: Magnification of the specific heat peak with a temperature resolution of 0.1 K.

at the same T and in general, the specific heat displays a sharp increase when the degree of cooperativity (measured by N_{eff}) markedly grows. The peak shifts toward lower T by increasing ρ. This finding, shared by previous studies [22], is physically meaningful as systems with larger anharmonic couplings should have a higher number of pairs which break at lower T.

A. Phase transition

While the results displayed so far witness by themselves reliability and consistency of the method, they also induce to face the unresolved issue regarding the order of the denaturation transition. Two questions need to be addressed:

(1) How does the entropy behave?

(2) To which extent do the thermodynamical properties depend on the number of paths included in the calculation? Or, how does the macroscopic system scales versus N_{eff}?

Taking the intermediate plot, $\rho=1$ in Fig. 2, I have thus progressively increased N_{eff} over the value $N_{\text{eff}}=9920$ at $T=260$ K in Fig. 2(a). While the T axis in the latter starts from 250 K, the 10 K shift upwards allows one to distinguish among the three plots. $T=260$ K is then the lower bound in the T window considered hereafter. Once the starting N_{eff} is

set at the lower bound, the thermodynamical properties are evaluated at higher T according to the procedure detailed above.

In Fig. 3, the calculation begins with $N_{\text{eff}}=31792$, a value larger by a factor three than in Fig. 2. Figure 3(a) plots: (i) the total number of paths versus T starting with $\sim 3.2 \times 10^6$ at $T=260$ K; (ii) the number of paths whose amplitude is larger than 1 Å; (iii) the number of paths whose amplitude is larger than 2 Å. All curves increase versus T and show a kink at T_c^*. This feature is more pronounced and evident than in Fig. 2, confirming that denaturation is indeed a highly cooperative phenomenon. While paths >1 Å already sample the plateau of V_M where the bps unbind, paths >2 Å certainly belong to broken pairs. Accordingly, the latter become more numerous around and above T_c^*. Thus, approaching the transition, the system incorporates a larger number of paths with an increasing fraction of broad path amplitudes. As such
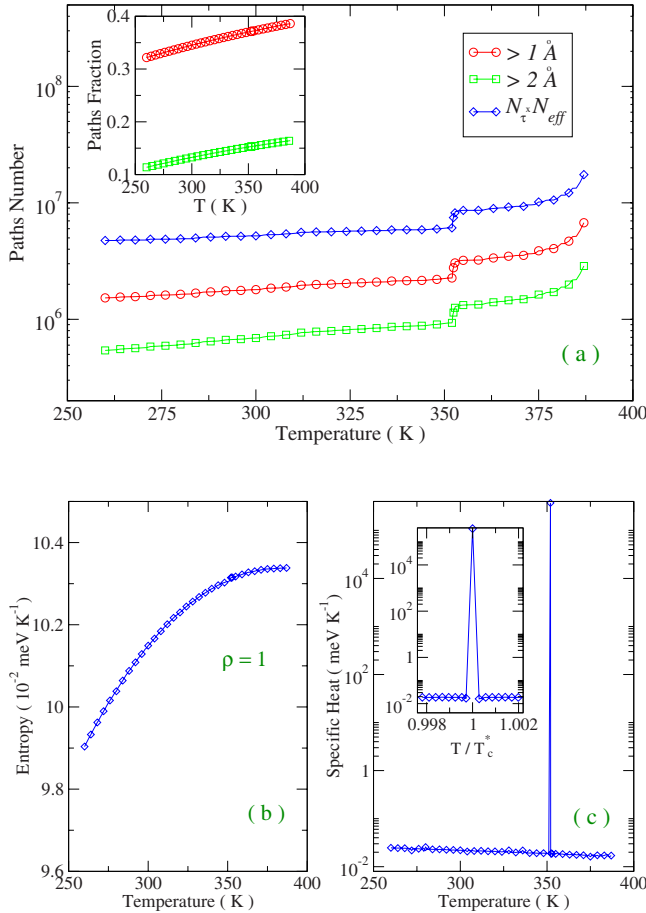
FIG. 4. (Color online) As in Fig. 3 but with N_{eff}=47494 at T=260 K.

fractions are considered reliable indicators of the order of the denaturation transition, quantitative estimates are greatly relevant. As the inset makes evident, the fractions of paths >1 Å and >2 Å, respectively (normalized over $N_\tau \cdot N_{eff}$), grow continuously versus T. Around T_c^*, one third of all paths are larger than 1 Å: it means that either they are broken or about to break. We see no steplike increase at T_c^* and this behavior is fully consistent with the continuous entropy growth shown in Fig. 3(b). The slight irregularity appearing in the entropy plot at T_c^* is responsible for the remarkable peak in the specific heat plotted in Fig. 3(c). The peak remains pinned at T_c^*=363 K as in Fig. 2(b). The inset shows an enlargement with a temperature resolution which appreciates 0.1 K. Finer partitions may be taken further increasing computing time.

The results presented so far point to the occurrence of a smooth crossover at T_c^*. While evidence is emerging for classifying the transition as of second order, the careful reader may argue that the aforementioned question (2) is not yet fully answered. I have then further increased the starting N_{eff}: all the obtained curves look like those presented in Fig. 3 with, at most, shifts of a few degrees in the location of T_c^*. Figure 4 shows the results obtained for N_{eff}=47 494 at T=260 K: in this case T_c^*=352 K which makes the largest observed shift with respect to Figs. 2 and 3. Note that ~8 ×10^6 paths participate to the computation around T_c^*, with

37% being larger than 1 Å. Remarkably, ~18×10^6 paths are accounted for at the largest T here considered.

Some features shown in the plots and drawn from the whole numerical work are in order:

(i) The entropy is always a continuously increasing function of T whatever the starting value for N_{eff} may be.

(ii) At T_c^*, the entropy always displays a slightly more pronounced enhancement although there is no finite melting entropy which would point to a first-order phase transition [24]. The small entropy kink is instead responsible for the peak in the specific heat. The crossover is smooth, the phase transition appears to be of second order.

(iii) The entropy values are larger by including more paths in the computation.

(iv) Above T_c^*, the entropy goes on growing but with smaller gradient consistently with the fact that, as the strands separation proceeds, the probability to add entropic effects is reduced.

(v) The location of T_c^* slightly varies with increasing N_{eff} but the changes are not significant oscillating in the range $T_c^* \simeq [350-365$ K] for ρ=1.

While different choices for the stiffness parameters in the PB Hamiltonian are not expected to change these qualitative physical features, it may be instead worth investigating by path integrals also Hamiltonian models with different stacking potentials both for homogeneous and heterogeneous DNA [44].

V. CONCLUSION

The homogeneous DNA denaturation has been studied in the path integral formalism by mapping the essential interactions of the Peyrard-Bishop Hamiltonian onto the imaginary time scale. While the nonlinearities in the stacking interactions are recognized as a key feature of the Hamiltonian, the path integral method permits to deal with the highly cooperative character of the denaturation by incorporating a very high number of degrees of freedom. I have thought of the transverse stretchings for the base pairs as time dependent paths: in the imaginary time formalism, this amounts to incorporate the temperature effects in the base pairs displacements. The path amplitudes are taken consistently with the shape of the Morse potential modeling the hydrogen bonds. Accordingly, I have set up a computational method to monitor, versus temperature, the paths ensemble which mainly contributes to the partition function, hence to the macroscopic equilibrium properties of the system. A temperature window has been considered in which denaturation is expected to occur. Given the size of the paths ensemble at the lower bound of such window, the second law of thermodynamics has been invoked to build the ensemble at any larger T. The method works self-consistently selecting the minimum number of paths required (at any T) to pursue the goal, that is the growing entropy constraint. There is no need to add further paths to the partition function once such goal is achieved: I feel that a *minimum effort principle* should hold. Instead, the system physical properties may depend on the initial ensemble size and care has been taken to unravel this issue. Thus, the partition function and its derivatives have

been computed by varying the boundary condition but no significant variation in the physical behavior has been observed. It is understood that the initial size should be large enough to allow for the appearance of cooperative effects. Remarkably, the ensemble size strongly grows around and above the denaturation temperature. The denaturation of the complementary strands always shows up as a smooth crossover whose location along the T axis may depend on the stiffness of the stacking interaction. Focusing on the system with moderate anharmonic stacking, it has been found that (i) the entropy grows continuously versus T, (ii) the specific heat displays a peak precisely at the same temperature for which the path ensemble size sharply grows.

These results suggest that the denaturation in homogeneous DNA is a highly cooperative phenomenon with the hallmarks of a second-order phase transition. Consistently,

the fraction of paths sampling the plateau of the Morse potential grows continuously also around the denaturation temperature. No steplike increase has been observed in the paths fractions which describe the unbinding of the base pairs. Eventually, I wish to emphasize that the thermodynamical constraint applied on the system does not force *a priori* the entropy to grow continuously. While the system autonomously selects at any T the total number of paths to fulfill such constraint, both a step discontinuity and a continuous behavior may in principle appear. The obtained plots are therefore not an artifact of the method.

After testing the feasibility of the path integral approach to a large size system as homogeneous DNA, I feel that the presented formalism may be further improved/extended to account for bubble formations in heterogeneous models with inhomogeneous sequences.

[1] M. Zoli, in *Polarons in Advanced Materials*, edited by A. S. Alexandrov, Springer Series in Materials Science (Springer, Berlin/Canopus, Bristol, 2007).

[2] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).

[3] R. M. Wartell and A. S. Benight, Phys. Rep. **126**, 67 (1985).

[4] D. Poland and H. Scheraga, J. Chem. Phys. **45**, 1456 (1966).

[5] M. Y. Azbel, Phys. Rev. A **20**, 1671 (1979).

[6] Y. Kafri, D. Mukamel, and L. Peliti, Phys. Rev. Lett. **85**, 4988 (2000).

[7] E. Carlon, E. Orlandini, and A. L. Stella, Phys. Rev. Lett. **88**, 198101 (2002).

[8] T. Ambjörnsson, S. K. Banik, O. Krichevsky, and R. Metzler, Phys. Rev. Lett. **97**, 128105 (2006).

[9] H. C. Fogedby and R. Metzler, Phys. Rev. Lett. **98**, 070601 (2007).

[10] M. Baiesi and R. Livi, J. Phys. A: Math. Theor. **42**, 082003 (2009).

[11] M. E. Fisher, J. Chem. Phys. **45**, 1469 (1966).

[12] A. Hanke, M. G. Ochoa, and R. Metzler, Phys. Rev. Lett. **100**, 018106 (2008).

[13] J. Rudnick and T. Kuriabova, Phys. Rev. E **77**, 051903 (2008).

[14] S. J. Rahi, M. P. Hertzberg, and M. Kardar, Phys. Rev. E **78**, 051910 (2008).

[15] S. Whitelam, S. Pronk, and P. L. Geissler, Biophys. J. **94**, 2452 (2008).

[16] F. de los Santos, O. Al Hammal, and M. A. Munoz, Phys. Rev. E **77**, 032901 (2008).

[17] C. Richard and A. J. Guttmann, J. Stat. Phys. **115**, 925 (2004).

[18] M. Peyrard and A. R. Bishop, Phys. Rev. Lett. **62**, 2755 (1989).

[19] E. W. Prohofsky, Phys. Rev. A **38**, 1538 (1988).

[20] D. J. Scalapino, M. Sears, and R. A. Ferrel, Phys. Rev. B **6**, 3409 (1972).

[21] D. Cule and T. Hwa, Phys. Rev. Lett. **79**, 2375 (1997).

[22] T. Dauxois, M. Peyrard, and A. R. Bishop, Phys. Rev. E **47**, R44 (1993).

[23] Y. L. Zhang, W. M. Zheng, J. X. Liu, and Y. Z. Chen, Phys. Rev. E **56**, 7100 (1997).

[24] N. Theodorakopoulos, T. Dauxois, and M. Peyrard, Phys. Rev. Lett. **85**, 6 (2000).

[25] L. van Hove, Physica **16**, 137 (1950).

[26] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Addison-Wesley, Reading, MA, 1958).

[27] J. A. Mac Cammon and S. C. Harvey, *Dynamics of Proteins and Nucleic Acids* (Cambridge University Press, Cambridge, 1988).

[28] M. Peyrard, Nonlinearity **17**, R1 (2004).

[29] A. Campa and A. Giansanti, Phys. Rev. E **58**, 3585 (1998).

[30] M. Joyeux and S. Buyukdagli, Phys. Rev. E **72**, 051902 (2005).

[31] R. Blossey and E. Carlon, Phys. Rev. E **68**, 061911 (2003).

[32] A. Bar, Y. Kafri, and D. Mukamel, Phys. Rev. Lett. **98**, 038103 (2007).

[33] A. Krueger, E. Protozanova, and M. D. Frank-Kamenetskii, Biophys. J. **90**, 3091 (2006).

[34] J. A. Krumhansl and J. R. Schrieffer, Phys. Rev. B **11**, 3535 (1975).

[35] M. Zoli, J. Low Temp. Phys. **151**, 1205 (2008).

[36] R. P. Feynman, Rev. Mod. Phys. **20**, 367 (1948).

[37] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, Polymer Physycs and Financial Markets* (World Scientific, Singapore, 2004).

[38] M. Zoli, Phys. Rev. B **71**, 205111 (2005).

[39] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products* (Academic, New York, 1965).

[40] J. Palmeri, M. Manghi, and N. Destainville, Phys. Rev. E **77**, 011913 (2008).

[41] S. Buyukdagli and M. Joyeux, Phys. Rev. E **76**, 021917 (2007).

[42] T. Dauxois and M. Peyrard, Phys. Rev. E **51**, 4027 (1995).

[43] B. Coluzzi, Phys. Rev. E **73**, 011911 (2006).

[44] S. Buyukdagli and M. Joyeux, Phys. Rev. E **77**, 031903 (2008).