

Theoretical investigation of finite size effects at DNA melting

Sahin BUYUKDAGLI and Marc JOYEUX^(#)

Laboratoire de Spectrométrie Physique (CNRS UMR 5588),

Université Joseph Fourier - Grenoble 1,

BP 87, 38402 St Martin d'Hères, FRANCE

We investigated how the finiteness of the length of the sequence affects the phase transition that takes place at DNA melting temperature. For this purpose, we modified the Transfer Integral method to adapt it to the calculation of both extensive (partition function, entropy, specific heat, etc) and non-extensive (order parameter and correlation length) thermodynamic quantities of finite sequences with open boundary conditions, and applied the modified procedure to two different dynamical models. We showed that rounding of the transition clearly takes place when the length of the sequence is decreased. We also performed a finite-size scaling analysis of the two models and showed that the singular part of the free energy can indeed be expressed in terms of an homogeneous function. However, both the correlation length ξ and the average separation between paired bases $\langle y \rangle$ diverge at the melting transition, so that it is no longer clear to which of these two quantities the length L of the system should be compared. Moreover, Josephson's identity is satisfied for none of the investigated models, so that the derivation of the characteristic exponents which appear, for example, in the expression of the specific heat, requires some care.

^(#)email : Marc.JOYEUX@ujf-grenoble.fr

PACS numbers: 87.14.Gg, 05.70.Jk, 87.15.Aa, 64.70.-p

I. INTRODUCTION

Real systems manifesting critical behavior have necessarily finite volume. However it is well-known that the finiteness of the system size lets the critical singularities disappear and smears out the phase transition (see for example [1-5]). The obvious argument which may reconcile these two aspects is that, as the finite size of the system is increased and passes through a critical value which characterizes the border of the thermodynamical domain, the thermodynamical limit is approximately reached and consequently the critical singularities manifest themselves. Thus it is crucial to thoroughly understand the evolution of singularities with respect to the system size and estimate the critical size above which the thermodynamical limit is attained.

An efficient tool to analyze the volume dependence of a critical phenomenon is the *finite size scaling theory*. Besides providing information on the rounding of critical singularities and the shift in the critical point, this theory is also an alternative way to determine critical exponents characterizing the phase transition.

Finite size scaling theory was developed by Fisher and Barber in the early seventies [6]. During the last thirty years it has been applied to various systems exhibiting both first and second order phase transitions. Among hundreds of subjects we can mention the study of finite size effects at first order transitions by gaussian approximation [7-8], the Gibbs ensemble [9], five dimensional Ising model [10-11], percolation models [12-13], stochastic sandpiles [14], six-dimensional Ising system [15], Baxter-Wu model [16], two dimensional anisotropic Heisenberg model [17]...

In nature, the majority of phase transitions are sharp and discontinuous first order transitions. Experimental UV absorption spectra of diluted DNA solutions reveal that the DNA melting transition belongs to this class. A widely used dynamical non-linear DNA model was proposed twenty years ago by Peyrard and Bishop [18]. This model involving harmonic interaction terms between successive base pairs was later improved by the contribution of Dauxois [19] and the new model (DPB model) yields a sharp transition. We recently proposed an alternative DNA model (JB model) which respects the finiteness of the stacking energy, and showed that this model also exhibits a sharp first order phase transition [20]. Then in [21] we showed that for both models the generalized homogeneity assumption is not respected so that Josephson's identity (also known as the hyperscaling relation) is not valid.

We tentatively explained this fact by the divergence of the order parameter at the critical point.

The goal of this article is to investigate the sequence length dependence of the DNA melting transition. This is an important point since experiments dealing with DNA molecules are carried out with various sequence lengths. To this end we employ a modified transfer integral method adapted to finite chains with open boundary conditions. The two hamiltonian DNA models to be studied, i.e. the DPB and the JB models, are briefly described in Sec. II. Section III deals with the transfer matrix theory for finite linear chains and Sec. IV is devoted to the finite size scaling analysis which leads to a better understanding of finite size effects.

II. NON-LINEAR HAMILTONIAN MODELS FOR DNA

The Hamiltonians of the two DNA models whose critical behaviour is studied in this article are of the form

$$H = \sum_{n=1}^N \left\{ \frac{p_n^2}{2m} + V_M(y_n) + W(y_n, y_{n-1}) \right\} \quad (1)$$

where y_n is the transverse stretching of the hydrogen bond between the n th pair of bases, while the one-particle Morse potential term

$$V_M(y_n) = D (1 - e^{-ay_n})^2 \quad (2)$$

models the binding energy of the same hydrogen bond. The choice of the nearest-neighbor interaction potential $W(y_n, y_{n-1})$ is crucial since the type of the transition, which is a collective effect, depends primarily on its form. The DPB model [19] assumes that the stacking energy is of the form

$$W(y_n, y_{n-1}) = \frac{K}{2} (y_n - y_{n-1})^2 [1 + \rho e^{-\alpha(y_n + y_{n-1})}] . \quad (3)$$

This non-linear stacking interaction has the particularity of having a coupling constant which drops from $K(1 + \rho)$ to K as the critical point is approached. This decreases the rigidity of the DNA chain close to the dissociation and yields a sharp, first-order transition.

This interaction potential still has the inconvenience that the stacking energy diverges when two paired bases separate. Taking into account the finiteness of the interaction between adjacent bases, we proposed a potential of the form [20]

$$W(y_n, y_{n-1}) = \frac{\Delta H}{2} \left[1 - e^{-b(y_n - y_{n-1})^2} \right] + K_b (y_n - y_{n-1})^2 \quad (4)$$

which, contrary to the model (3), depends only on the distance between base pairs. The small harmonic term, whose constant K_b is 2000 times smaller than the parameter K of the DPB model, was introduced in order to take into account the stiffness of the phosphate-sugar backbone.

Numerical values of the parameters are those of Refs. [19,20], that is $D = 0.03 eV$, $a = 4.5 \text{ \AA}^{-1}$, $\alpha = 0.35 \text{ \AA}^{-1}$, $K = 0.06 eV \text{ \AA}^{-2}$, $\rho = 1$ for the DPB model, and $D = 0.04 eV$, $a = 4.45 \text{ \AA}^{-1}$, $\Delta H = 0.44 eV$, $K_b = 10^{-5} eV \text{ \AA}^{-2}$ and $b = 0.10 \text{ \AA}^{-2}$ for the JB model.

III. TRANSFER INTEGRAL METHOD FOR FINITE CHAINS WITH OPEN BOUNDARY CONDITIONS

A. The partition function and extensive thermodynamic quantities

Let us define the Transfer Integral (TI) kernel according to

$$K(y_n, y_{n-1}) = \exp[-\beta \{V_M(y_n)/2 + V_M(y_{n-1})/2 + W(y_n, y_{n-1})\}] \quad (5)$$

where $\beta = (k_B T)^{-1}$ is the inverse temperature. In the following analysis we deal only with sequences having open boundary conditions. Then the partition function of the system can be expressed as

$$Z = \int dy_1 dy_2 \cdots dy_N e^{-\beta V_M(y_1)/2} K(y_2, y_1) K(y_3, y_2) \cdots K(y_N, y_{N-1}) e^{-\beta V_M(y_N)/2}. \quad (6)$$

The TI method consists in expanding the kernel of Eq. (5) in an orthonormal basis

$$K(y_n, y_{n-1}) = \sum_i \lambda_i \Phi_i(y_n) \Phi_i(y_{n-1}) \quad (7)$$

where the $\{\Phi_i\}$ and $\{\lambda_i\}$ are the eigenvalues and eigenvectors of the integral operator and satisfy

$$\int dx K(x, y) \Phi_i(x) = \lambda_i \Phi_i(y). \quad (8)$$

This integral equation was solved by diagonalizing the symmetric TI operator $K(x, y)$ on a regularly spaced grid defined between $y_{min} = -200/a$ and $y_{max} = 4000/a$ with $1/a$ intervals. Numerical integrations were performed on the same grid.

In this study we extended the transfer matrix approach for open chains developed in [22] to adapt it to the calculation of the order parameter $\langle y \rangle$ and the correlation length ξ . Let us first consider extensive thermodynamic quantities. By introducing

$$a_i = \int dy e^{-\beta V_M(y)/2} \Phi_i(y) \quad (9)$$

and by substituting the kernel expansion of Eq. (7) into Eq. (6), we get

$$Z = \sum_i a_i^2 \lambda_i^{N-1}. \quad (10)$$

Determination of the partition function then allows the computation of extensive quantities of the system such as the free energy, the entropy and the specific heat :

$$\begin{aligned} F &= -k_B T \ln(Z) \\ S &= -\frac{\partial F}{\partial T} \\ C_V &= -T \frac{\partial^2 F}{\partial T^2}. \end{aligned} \quad (11)$$

In the thermodynamical limit $N \rightarrow \infty$ the major contribution to the partition function arises from the largest eigenvalue λ_1 and in this limit it is reasonable to drop the eigenvalues with $i \geq 2$. Nevertheless, we will consider large DNA molecules as well as small ones. Consequently as many eigenvalues as possible must be taken into account in numerical computations. From the practical point of view, it was found that considering the first 400 eigenvalues is enough to insure numerical convergence of the results presented below.

B. The order parameter and the correlation length

The order parameter of DNA melting transition is the mean separation of the bases averaged over the sites of the sequence :

$$\langle y \rangle = \frac{1}{N} \sum_{n=1}^N \langle y_n \rangle. \quad (12)$$

In order to reduce $\langle y_n \rangle$ to a form depending only on the eigenvalues and eigenvectors of the TI operator, we first write it as

$$\langle y_n \rangle = \frac{1}{Z} \int dy_1 dy_2 \cdots dy_N y_n e^{-\beta V_M(y_1)/2} K(y_2, y_1) K(y_3, y_2) \cdots K(y_N, y_{N-1}) e^{-\beta V_M(y_N)/2}. \quad (13)$$

Substituting Eq. (7) into Eq. (13) and defining

$$\begin{aligned} b_i &= \int dy e^{-\beta V_M(y)/2} \Phi_i(y) y \\ Y_{ij}^{(1)} &= \int dy \Phi_i(y) y \Phi_j(y), \end{aligned} \quad (14)$$

we obtain

$$\langle y_1 \rangle = \langle y_N \rangle = \frac{1}{Z} \sum_i a_i b_i \lambda_i^{N-1} \quad (15)$$

and

$$\langle y_n \rangle = \frac{1}{Z} \sum_{i,j} a_i a_j Y_{ij}^{(1)} \lambda_i^{n-1} \lambda_j^{N-n} \quad (16)$$

for $n \neq 1, N$. By evaluating the geometric summation that appears in Eq. (12) when replacing the $\langle y_n \rangle$ by their expressions in Eqs. (15)-(16), we finally get

$$\langle y \rangle = \frac{2}{ZN} \sum_i a_i b_i \lambda_i^{N-1} + \frac{1}{ZN} \sum_{i,j} a_i a_j Y_{ij}^{(1)} \lambda_i^{-1} \lambda_j^N \frac{r_{ij}^2 - r_{ij}^N}{1 - r_{ij}} \quad (17)$$

with $r_{ij} = \lambda_i/\lambda_j$. The computation of the correlation length ξ proceeds along similar lines although it is more elaborate. A sketch of the derivation and the analytical result can be found in Appendix A.

IV. FINITE-SIZE EFFECTS NEAR THE CRITICAL POINT

A. Rounding of the melting transition of DNA

It is well known that a finite-size system does not exhibit any phase transition. At the critical point its free energy is analytic and consequently all thermodynamical quantities are regular. Let L be the size of a system having a critical behaviour in the thermodynamical limit $L \rightarrow \infty$. For this system finite-size effects manifest themselves as $e^{-L/\xi}$, where ξ is the correlation length, by rounding the critical point singularity. In other words they become important over a region for which $\xi \sim L$. A simple example of this rounding phenomena concerning the Ising model can be found in [8]. For an infinite-size Ising system, as the magnetic field H varies, the order parameter jumps discontinuously from $-M_{cr}$ to $+M_{cr}$ at the critical point $H = 0$. On the other hand if the system's size is finite, this transition occurs on a finite region of order $\Delta H \simeq k_B T / (M_{cr} L^d)$ with a large but finite slope $\sim M_L^2 L^d / (k_B T)$ where M_L is the most probable value of the magnetization in the finite system.

For the two DNA models sketched in Sect. 2, the size L of the system is equal to the number N of base pairs in the sequence times the distance between two successive base pairs. This latter quantity playing no role in the dynamics of the investigated models, we will henceforth use indistinctly N or L to refer to the size of the sequence.

Given a sequence of length N , the first task consists in determining its critical temperature, which we denote by $T_c(N)$. Among the several methods listed for example in [25], we found it rather simple and convenient to search for the maximum of the specific heat C_V , which is more pronounced than that of the correlation length ξ , thus allowing for a more accurate localization of the temperature. Two observations confirm a posteriori that the critical temperatures thus obtained are correct. First, the shift in critical temperature is found to vary as a power of N , as predicted by finite-size scaling theory (see below). The top plot of Fig. 1 shows for example that $T_c - T_c(N)$, where T_c stands for $T_c(\infty)$, varies as $N^{-1.00}$ and $N^{-1.05}$ for the DPB and JB models, respectively. Note that this scaling is also in excellent agreement with the semi-empirical formula used by experimentalists to calculate the melting temperature of finite sequences. Moreover, as will be seen later (Figs. 3 and 4) the curves for the temperature evolution of C_V , ξ , $\langle y \rangle$, etc... for sequences with different lengths N all coincide sufficiently far from the critical temperature when plotted as a

function of the reduced temperature

$$t(N) = \frac{T - T_c(N)}{T_c(N)}.$$

In order to illustrate finite-size effects acting on DNA melting, we first computed the entropy per base, $s = S/N$, for an infinite chain and a short DNA sequence for both the DPB and the JB models. Results are shown in Fig. 2. At the thermodynamic limit, the entropy s is clearly discontinuous at the critical temperature, as is expected for first order phase transitions. In contrast, smooth curves are observed over the whole temperature range for the sequence with $N = 100$. We next computed the specific heat per base, $c_V = C_V/N$, for increasing sequence length and temperature. The top and bottom plots of Fig. 3 show the temperature evolution of c_V for seven values of N ranging from 100 to infinity for the DPB and JB models, respectively. It is seen in this figure that rounding manifests itself through a decrease in the maximum of c_V as N decreases, but also through the fact that the sharp rise of c_V takes place further and further from the critical temperature, that is, at increasingly larger values of $|t(N)|$. This is particularly clear for the DPB model, which at the thermodynamic limit undergoes a very sharp transition, i.e. a transition that is noticeable only at very small values of $|t| = |t(\infty)|$ [21]. Quite interestingly, examination of Fig. 3 also indicates that the two models consequently give very comparable results up to $N \simeq 1000$, while the narrower nature of the phase transition for the DPB model becomes apparent for longer sequences.

At this point, it should be emphasized that boundary effects may become important when the size of the system is small. In order to check whether such boundary effects play a role in the results presented above, we repeated these computations with periodic boundary conditions instead of open ones and found that this alters only very little the results for s and c_V down to $N = 100$. Conclusion is therefore that boundary effects play only a marginal role down to this size.

Finally, we computed, for the JB model, the temperature evolution of the correlation length ξ (Eq. (A2)) and the order parameter $\langle y \rangle$ (Eq. (17)) for increasing values of N . As the critical temperature is approached, the correlation length ξ of a finite-size system is expected to increase according to the power law $\xi \propto t^\nu(N)$ till it reaches the system's dimension L and freezes. This behaviour can be checked in the middle plot of Fig. 1, which shows the evolution of the maximum of the correlation length (in units of the separation

between successive base pairs) as a function of N : the maximum of ξ is indeed of the same order of magnitude as N and the curve scales as $N^{0.97}$. An exception however occurs for the last three points with $N \geq 3000$: we will come back later to this point. The bottom plot of Fig. 4 additionally shows the temperature evolution of ξ for seven values of N ranging from 100 to infinity. One observes just the same rounding effects as for the specific heat in Fig. 3. This is again the case for the temperature evolution of the order parameter $\langle y \rangle$, which is drawn in the top plot of Fig. 4. This latter plot however displays a remarkable feature, in the sense that all the curves converge to the same limit at $T_c(N)$. To understand why this is the case, it must be realized that $\langle y \rangle$, the average separation between paired bases, is the only quantity which diverges at the critical temperature whatever the size N of the sequence, while, for example, c_V and ξ diverge for infinitely long chains but remain finite for finite chains. The limit towards which all curves converge in the top plot of Fig. 4 is thus just the approximation of infinity imposed by the numerical procedure (size of the grid, etc...).

B. Finite-size scaling analysis

The basic idea of finite-size scaling is that the correlation length ξ is the only length that matters close to the critical temperature and that one just needs to compare the linear dimension L of the system to ξ : rounding and shifting indeed set in as soon as $L/\xi \sim 1$. Since, by definition of the critical exponent ν , ξ grows as $t^{-\nu}$, one has $(L/\xi) \propto (tL^{1/\nu})^\nu$ as $L \rightarrow \infty$ or, L being proportional to the number N of paired bases, $(L/\xi) \propto (tN^{1/\nu})^\nu$. In the absence of external field, it is therefore natural to write the singular part of the free energy of the finite-size system in the form

$$f_{sing} = N^{-d} Y(tN^{1/\nu}). \quad (18)$$

where Y is some homogeneous function. Differentiating Eq. (18) twice with respect to t , one obtains that c_V is equal to

$$c_V = N^\rho G(tN^\sigma), \quad (19)$$

where

$$\begin{aligned}\rho &= \frac{2}{\nu} - d \\ \sigma &= \frac{1}{\nu}\end{aligned}\tag{20}$$

and G is an homogeneous function which is proportional to the second derivative of Y . By using Josephson's identity, $2 - \alpha = \nu d$, where α is the critical exponent for c_V ($c_V \propto t^{-\alpha}$), coefficients ρ and σ can be recast in the form

$$\begin{aligned}\rho &= \frac{\alpha}{\nu} \\ \sigma &= \frac{1}{\nu}\end{aligned}\tag{21}$$

Conversely, if there occur several lengths that diverge at the critical point, as is the case for DNA melting, then it is no longer so clear to which of these lengths L should be compared. In order to tackle this more complex case, Binder et al [24] derived a method which is based on the use of an irrelevant variable u and an expression of the form

$$f_{sing} = N^{-d} F(tN^{y_t}, uN^{y_u})\tag{22}$$

After several approximations and a little bit of algebra, these authors obtain c_V in the form of Eq. (19) with, however,

$$\begin{aligned}\rho &= \frac{2d}{2\beta + \gamma} - d \\ \sigma &= \frac{2d}{2\beta + \gamma}.\end{aligned}\tag{23}$$

Finally, the lengths that diverge at DNA melting are ξ and $\langle y \rangle$, the average separation between paired bases. One might wonder whether L should not be compared to $\langle y \rangle$ instead of ξ . In order to check this hypothesis, let us denote by λ the characteristic exponent for $\langle y \rangle$ ($\langle y \rangle \propto t^\lambda$). Remember that if the external field is proportional to y then $\langle y \rangle$ is the order parameter m , so that λ is equal to β , the critical exponent for m ($m \propto t^\beta$). Let us next express the singular part of the free energy in the form

$$f_{sing} = N^{-d} F(tN^{-1/\lambda}).\tag{24}$$

Differentiating Eq. (24) twice with respect to t , one again obtains c_V in the form of Eq. (19) with, however

$$\begin{aligned}\rho &= -\frac{2}{\lambda} - d \\ \sigma &= -\frac{1}{\lambda}.\end{aligned}\tag{25}$$

Table I shows the values of ρ and σ calculated from the characteristic exponents reported in [21] and Eqs (20), (21), (23) and (25), as well as adjusted values. These latter ones were obtained by varying ρ and σ by hand in order that the plots of c_V/N^ρ as a function of tN^σ are superposed for an interval of values of N as large as possible. By setting $t = 0$ in Eq. (19), one sees that the maximum of c_V scales as N^ρ . ρ was therefore adjusted in the neighbourhood of the slope of the plot of the maximum of c_V as a function of N (bottom plot of Fig. 1). On the other hand, σ was adjusted in the neighbourhood of $1/\nu$. Examination of Table I indicates that the values of ρ and σ obtained from Eqs. (20) and (25) compare well with the adjusted ones, while this is certainly not the case for the values obtained from Eqs. (21) and (23). Figs. 5 and 6 further show plots of c_V/N^ρ as a function of tN^σ for, respectively, the DPB and JB models, and values of ρ and σ obtained from Eq. (21) (top plots) and adjusted ones (bottom plots). It is seen in the top plots that the curves with different values of N are far from being superposed for the values of ρ and σ obtained from Eq. (21), and the situation is still worse with Eq. (23). In contrast, the various curves are fairly well superposed for the adjusted values of ρ and σ (see bottom plots of Figs 5 and 6), as well as those obtained with Eqs. (20) and (25). An exception occurs for the curves corresponding to the largest values of N in the JB model (bottom plot of Fig. 6). Remember that the corresponding points also depart from the power law in the bottom plot of Fig. 1. The reason for this is that the TI method fails to give correct values of thermodynamical observables too close to the phase transition discontinuity. Examination of the top plots of Fig. 3 shows that sequences of length $N = 10000$ are still rather far from the thermodynamic limit for the DPB model, so that one needs not to worry about the effect of the discontinuity on TI calculations. In contrast, the bottom plot of Fig. 3 and the two plots of Fig. 4 indicate that sequences of length $N = 10000$ have reached the thermodynamic limit for the JB model, so that the perturbative effect of the discontinuity becomes noticeable in TI calculations.

The fact that Eq. (20) leads to a correct superposition of the curves for different values of the sequence length N is the proof that the basic hypothesis of finite-size scaling theory is satisfied. Since, however, Eq. (25) also leads to a correct superposition of the curves, it is, as expected, no longer clear to which diverging length (ξ or $\langle y \rangle$) L should be compared, both possibilities leading to a reasonable result. On the other hand, the fact that curves with different N are no longer superposed when Eq. (21) is used to calculate ρ and σ simply reflects the fact that Josephson's identity, $2 - \alpha = \nu d$, is not valid for these two models of DNA melting, a conclusion which was already arrived at in our preceding work [21]. Finally, the fact that curves also do not superpose when Eq. (23) is used indicates that one of the several hypotheses made by the authors of Ref. [24] to arrive to these expressions is not satisfied for the DNA models, although it is not an easy task to tell which one(s) is(are) invalidated. Alternatively, Eq. (23) can be straightforwardly derived from Eq. (20) by using Rushbrooke identity ($\alpha + 2\beta + \gamma = 2$) as well as Josephson's one. This latter identity being not valid, it comes as no surprise that Eq. (23) leads to as poor a result as Eq. (21).

V. CONCLUSION

To summarize, we modified the Transfer Integral method to adapt it to the calculation of thermodynamic quantities of finite sequences with open boundary conditions. Non-extensive quantities, like the average separation of paired bases $\langle y \rangle$ and the correlation length ξ , turned out to be the most tricky ones to evaluate. We then applied this modified procedure to the DPB and JB dynamical models, in order to clarify how the finiteness of the length of the sequence affects the phase transition that takes place at DNA melting temperature. We showed that the rounding of the transition that occurs when the size of the sequence decreases is clearly reflected in the temperature evolution of most quantities, including the specific heat c_V , the correlation length ξ , and the average separation of paired bases $\langle y \rangle$. We next performed a finite-size scaling analysis of the two systems and showed that the singular part of the free energy can indeed be expressed in terms of an homogeneous function. However, since both ξ and $\langle y \rangle$ diverge at the melting transition, it is no longer clear whether the argument of the homogeneous function should be (a power of) L/ξ or $L/\langle y \rangle$. Moreover, Josephson's identity is satisfied for none of the investigated systems, so that the derivation of the characteristic exponents ρ and σ , which appear in the asymptotic expression of the

specific heat c_V , requires some care.

The Transfer Integral (TI) method appears as the only efficient numerical tool to study the thermodynamics of DNA melting in detail. In the formulation used here, it however applies only to homogeneous chains, while it is well established that the heterogeneity of real DNA molecules may smear out the discontinuity of the melting transition, just like the finiteness of the sequence does. Our next goal is therefore to overcome the technical difficulty associated with the application of the TI method to inhomogeneous chains and investigate the effect of heterogeneities on the phase transition at DNA melting.

APPENDIX A: COMPUTATION OF THE CORRELATION LENGTH

The *static form factor* is defined as

$$S(q) = \left\langle \left| \sum_{n=1}^N (y_n - \langle y_n \rangle) e^{iqan} \right|^2 \right\rangle \quad (\text{A1})$$

and the correlation length is given by

$$\xi^2 = - \frac{1}{2S(q)} \frac{d^2 S(q)}{dq^2} \Big|_{q=0}. \quad (\text{A2})$$

We stress that the statistical weight of the bases at the extremities $n = 1, N$ is different from that of the other ones, so that they must be treated separately. We first write Eq. (A1) in the explicit form

$$S(q) = \sum_{n=1}^N \sum_{m=1}^N \langle \delta y_n \delta y_m \rangle e^{iqa(n-m)} \quad (\text{A3})$$

where $\delta y_n = y_n - \langle y_n \rangle$. By isolating averages concerning extremity values, we get

$$\begin{aligned} S(q) = & \langle \delta y_1^2 \rangle + \langle \delta y_N^2 \rangle + 2 \langle \delta y_1 \delta y_N \rangle \cos[qa(N-1)] \\ & + S_1 e^{iqa} + S_1^* e^{-iqa} + S_N e^{iNqa} + S_N^* e^{-iNqa} + S_{mid} \end{aligned} \quad (\text{A4})$$

where

$$\begin{aligned}
S_1 &= \sum_{m=2}^{N-1} \langle \delta y_1 \delta y_m \rangle e^{-iqam} \\
S_N &= \sum_{m=2}^{N-1} \langle \delta y_m \delta y_N \rangle e^{-iqam} \\
S_{mid} &= \sum_{m=2}^{N-1} \sum_{n=2}^{N-1} \langle \delta y_n \delta y_m \rangle e^{iqa(n-m)}.
\end{aligned} \tag{A5}$$

Defining

$$\begin{aligned}
c_i &= \int dy e^{-\beta V_M(y)/2} y^2 \Phi_i(y) \\
Y_{ij}^{(2)} &= \int dy \Phi_i(y) y^2 \Phi_j(y).
\end{aligned} \tag{A6}$$

we obtain the relations :

$$\begin{aligned}
\langle y_1^2 \rangle &= \langle y_N^2 \rangle = \frac{1}{Z} \sum_i a_i c_i \lambda_i^{N-1} \\
\langle y_1 y_N \rangle &= \frac{1}{Z} \sum_i b_i^2 \lambda_i^{N-1} \\
\langle y_1 y_m \rangle &= \frac{1}{Z} \sum_{ij} a_j b_i \lambda_i^{m-1} \lambda_j^{N-m} Y_{ij}^{(1)} \\
\langle y_m y_N \rangle &= \frac{1}{Z} \sum_{ij} a_i b_j \lambda_i^{m-1} \lambda_j^{N-m} Y_{ij}^{(1)} \\
\langle y_n^2 \rangle &= \frac{1}{Z} \sum_{ij} a_i a_j \lambda_i^{n-1} \lambda_j^{N-n} Y_{ij}^{(2)} \\
\langle y_n y_m \rangle &= \frac{1}{Z} \sum_{ijk} a_i a_k Y_{ij}^{(1)} Y_{jk}^{(1)} \cdot \begin{cases} \lambda_i^{n-1} \lambda_j^{m-n} \lambda_k^{N-m} & \text{if } m > n \\ \lambda_i^{m-1} \lambda_j^{n-m} \lambda_k^{N-n} & \text{if } m < n. \end{cases}
\end{aligned} \tag{A7}$$

According to the relations in Eqs. (16) and (A7), the summations in Eq. (A3) are just geometric series, which we evaluated formally in order to increase the speed of numerical calculations by a factor N^2 . After some tedious algebra, one obtains :

$$S(q) = \langle \delta y_1^2 \rangle + \langle \delta y_N^2 \rangle + 2 \langle \delta y_1 \delta y_N \rangle \cos[qa(N-1)] + \sum_{ij} H_{ij} \frac{f_{ij}(2) - f_{ij}(N)}{1 - f_{ij}(1)}$$

$$\begin{aligned}
& + \sum_{i,j} \frac{1}{\cosh(\alpha_{ij}) - \cos(qa)} \{ [D_{ij}f_{ij}(1) + C_{ij}f_{ij}(N)] \cos[(N-2)qa] - D_{ij}f_{ij}(N-1) - C_{ij}f_{ij}(2) \\
& \quad - [D_{ij}f_{ij}(2) + C_{ij}f_{ij}(N-1)] \cos[(N-1)qa] + [D_{ij}f_{ij}(N) + C_{ij}f_{ij}(1)] \cos[qa] \} \\
& + \sum_{ijk} \frac{2M_{ijk}}{(1-g(1,1))(1+f_{ij}(2)-2f_{ij}(1)\cos(qa))(1+f_{jk}(2)-2f_{jk}(1)\cos(qa))} \\
& \times \{ -g(3,3) - g(4,4) - g(2,4) + g(N,N) + g(N,N+2) + g(N+1,N+1) \\
& + [g(2,3) - g(N,N+1) + 2g(3,4) - g(N+1,N+2) + g(4,3) - g(N+1,N) - g(N,N+1)] \cos(qa) \\
& + [g(N+1,N+1) - g(3,3)] \cos(2qa) + [g(2,N+1) - g(3,N+2)] \cos[(N-3)qa] \\
& + [g(4,N+2) - g(2,N)] \cos[(N-2)qa] + [g(3,N) - g(4,N+1)] \cos[(N-1)qa] \} \\
& + \left| \sum_{ij} M_{ij} \frac{e^{2(iqa-\alpha_{ij})} - e^{N(iqa-\alpha_{ij})}}{1 - e^{iqa-\alpha_{ij}}} \right|^2,
\end{aligned} \tag{A8}$$

where

$$\begin{aligned}
\alpha_{ij} &= -\ln(r_{ij}) \\
M_{ij} &= \frac{1}{Z} \lambda_i^{-1} \lambda_j^N a_i a_j Y_{ij}^{(1)} \\
H_{ij} &= \frac{1}{Z} \lambda_i^{-1} \lambda_j^N a_i a_j Y_{ij}^{(2)} \\
T_{ij} &= \frac{1}{Z} \lambda_i^{-1} \lambda_j^N b_i a_j Y_{ij}^{(1)} \\
G_{ij} &= \frac{1}{Z} \lambda_i^{-1} \lambda_j^N a_i b_j Y_{ij}^{(1)} \\
C_{ij} &= T_{ij} - \langle y_1 \rangle M_{ij} \\
D_{ij} &= G_{ij} - \langle y_1 \rangle M_{ij} \\
f_{ij}(n) &= e^{-n\alpha_{ij}} \\
g(n, m) &= e^{-n\alpha_{ij} - m\alpha_{jk}}.
\end{aligned} \tag{A9}$$

Finally, $S(q)$ must be derivated twice with respect to the wave vector q in order to get the correlation length from Eq. (A2).

-
- [1] D.J. Wales and J.P.K. Doye, *J. Chem. Phys.* 103, 3061 (1995)
 - [2] P. Borrmann, Oliver Mulken and Jens Harting, *Phys. Rev. Lett.* 84 3511 (2000)
 - [3] O. Mulken, H. Stamerjohanns and P. Borrmann, *Phys. Rev. E* 64, 047105 (2001)
 - [4] D. J. Dean, *Int. J. Mod. Phys.* 17, 5093 (2001)
 - [5] N. A. Alves, J. P. N. Ferrite and U. H. E. Hansmann, *Phys. Rev. E* 65, 036110 (2002)
 - [6] M.E. Fisher and M.N. Barber, *Phys. Rev. Lett.* 28, 1516 (1972)
 - [7] K. Binder and D.P. Landau, *Phys. Rev. B* 30, 1477 (1984)
 - [8] M.S.S Challa, D.P. Landau and K. Binder, *Phys. Rev B* 34, 1841 (1984)
 - [9] K.K. Mon and K. Binder, *J. Chem. Phys.* 96, 6989 (1992)
 - [10] E. Luijten, K. Binder and H.W.J Blöte, *Eur. Phys. J. B* 9, 289 (1999)
 - [11] H.W.J. Blöte and E. Luijten, *Europhys. Lett.*, 38, 565 (1997)
 - [12] S. Clar, B. Drossel, K. Schenk and F. Schwabl, *Phys. Rev. E* 56, 2467 (1997)
 - [13] M. Masihi, P.R. King and P. Nurafza, *Phys. Rev. E* 74, 042102 (2006)
 - [14] B. Tadić, *Phys. Rev. E* 59, 1452 (1999)
 - [15] Z. Merdan and R. Erdem, *Phys. Letters A* 330, 403 (2004)
 - [16] S.S. Martinos, A. Malakis and I. Hadjiagapiou, *Physica A* 355, 393 (2005)
 - [17] C. Zhou, D.P. Landau and T.C. Schulthess, *Phys. Rev. B* 74, 064407 (2006)
 - [18] M. Peyrard and A.R. Bishop, *Phys. Rev. Lett.* 62, 2755 (1989)
 - [19] T. Dauxois, M. Peyrard and A.R. Bishop, *Phys. Rev. E* 47, R44 (1993)
 - [20] M. Joyeux and S. Buyukdagli, *Phys. Rev. E* 72, 051902 (2005)
 - [21] S. Buyukdagli and M. Joyeux, *Phys. Rev. E* 73, 051910 (2006)
 - [22] Y.L. Zhang, W.M. Zheng, J.X. Liu and Y.Z. Chen, *Phys. Rev. E* 56, 7100 (1997)
 - [23] V. Privman and M. E. Fisher, *J. Stat. Phys.* 33, 385(1983)
 - [24] K. Binder, M. Nauenberg, V. Privman, A. P. Young, *Phys. Rev. B* 31, 1498 (1985)
 - [25] K. Binder, *Ferroelectrics*, 73, 43 (1987)

TABLE CAPTION

Table I : Values of the coefficients ρ and σ of Eq. (19) for the DPB and JB models. First four lines show the values calculated from the characteristic exponents reported in Ref. [21] and Eqs (20), (21), (23) and (25). The last line shows the values adjusted by hand in order that the plots of c_V/N^ρ as a function of tN^σ are superposed on an interval of values of N as large as possible (see bottom plots of Figs. 5 and 6).

FIGURE CAPTIONS

Figure 1 : (color online) Log-log plots, as a function of the sequence length N , of the reduced critical temperature shift $1 - T_c(N)/T_c(\infty)$ (top plot), the maximum of the correlation length ξ (middle plot), and the maximum of c_V (bottom plot), according to the DPB (squares) and JB (circles) models. ξ is in units of the separation between two successive base pairs, and c_V in units of k_B . The solid and dash-dotted lines show the result of the adjustment of power laws against the calculated points.

Figure 2 : (color online) Plot of the entropy per site s as a function of the rescaled temperature $t(N)$ for an infinitely long chain (circles) and a sequence with $N = 100$ bp (squares), according to the DPB model (top plot) and the JB one (bottom plot). s is in units of k_B .

Figure 3 : (color online) Log-log plots of the specific heat per site c_V as a function of the opposite $-t(N)$ of the rescaled temperature for the DPB model (top plot) and the JB one (bottom plot) and seven values of the sequence length N ranging from 100 to ∞ . c_V is in units of k_B . Note that, at the thermodynamic limit of infinitely long chains, c_V becomes infinite at the critical temperature but numerical limitations of the TI method prevent the observation of such divergence.

Figure 4 : (color online) Log-log plots of the correlation length ξ (bottom plot) and the order parameter $\langle y \rangle$ (top plot) as a function of the opposite $-t(N)$ of the rescaled temperature for the JB model and seven values of the sequence length N ranging from 100 to ∞ . ξ is in units of the separation between two successive base pairs and $\langle y \rangle$ in units of the inverse $1/a$ of the Morse potential parameter. Although numerical limitations of the TI method prevent the observation of such divergences, ξ becomes infinite at the critical temperature at the thermodynamic limit of infinitely long chains, while $\langle y \rangle$ becomes infinite at the critical temperature whatever the length N of the sequence.

Figure 5 : (color online) Plots, for six values of the sequence length N ranging from 100 to 10000, of c_V/N^ρ as a function of tN^σ for the DPB model and values of ρ and σ obtained from Eq. (21) (top plot) or adjusted by hand (bottom plot).

Figure 6 : (color online) Plots, for six values of the sequence length N ranging from 100 to 10000, of c_V/N^ρ as a function of tN^σ for the JB model and values of ρ and σ obtained from Eq. (21) (top plot) or adjusted by hand (bottom plot).

	DPB model		JB model	
	ρ	σ	ρ	σ
Eq. (20)	0.79	0.89	0.63	0.81
Eq. (21)	1.29	0.89	0.92	0.81
Eq. (23)	1.78	1.39	1.82	1.41
Eq. (25)	0.87	0.93	0.53	0.76
adjusted	0.85	1.00	0.45	0.90

TABLE I:

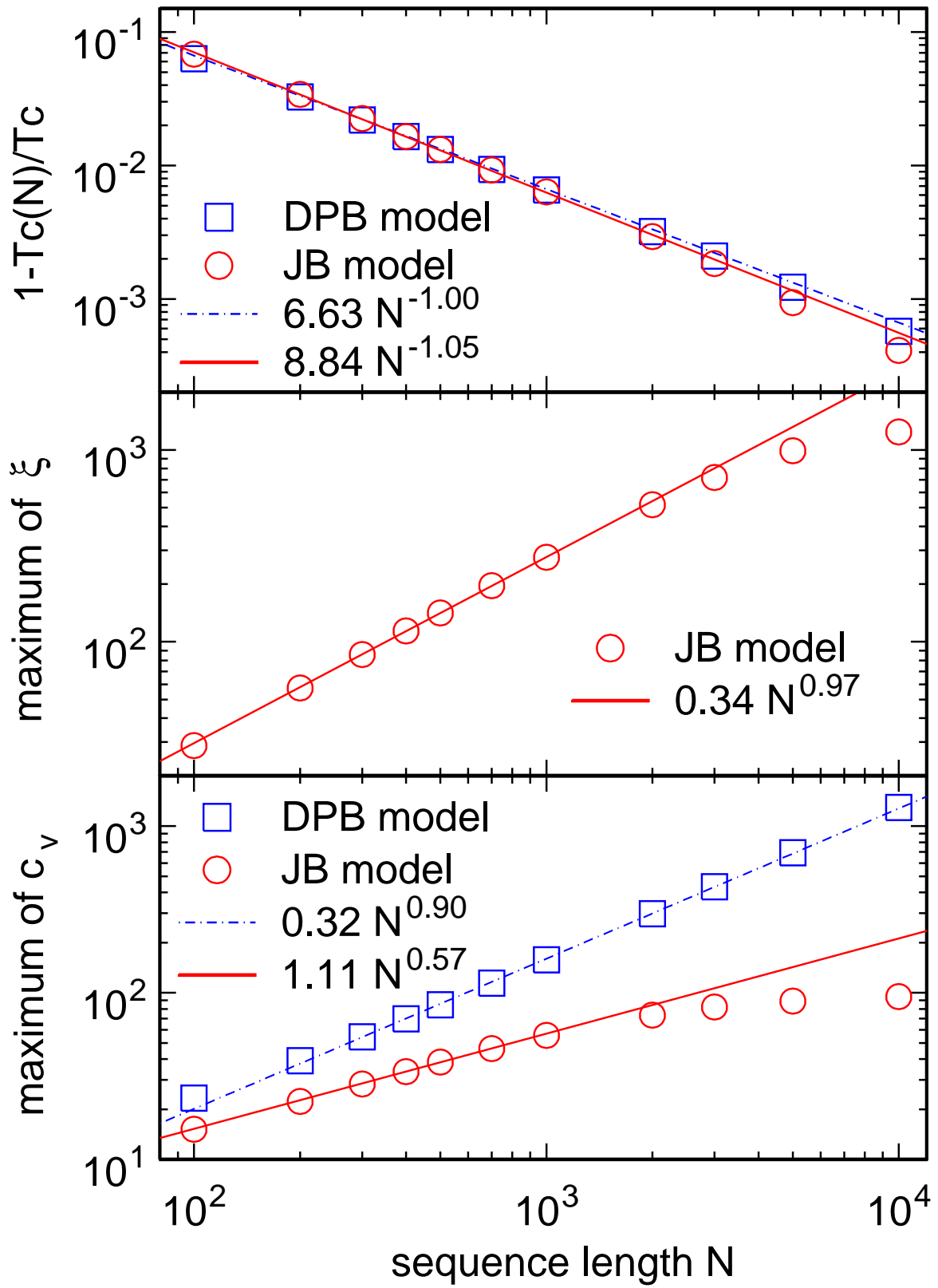


FIG. 1:

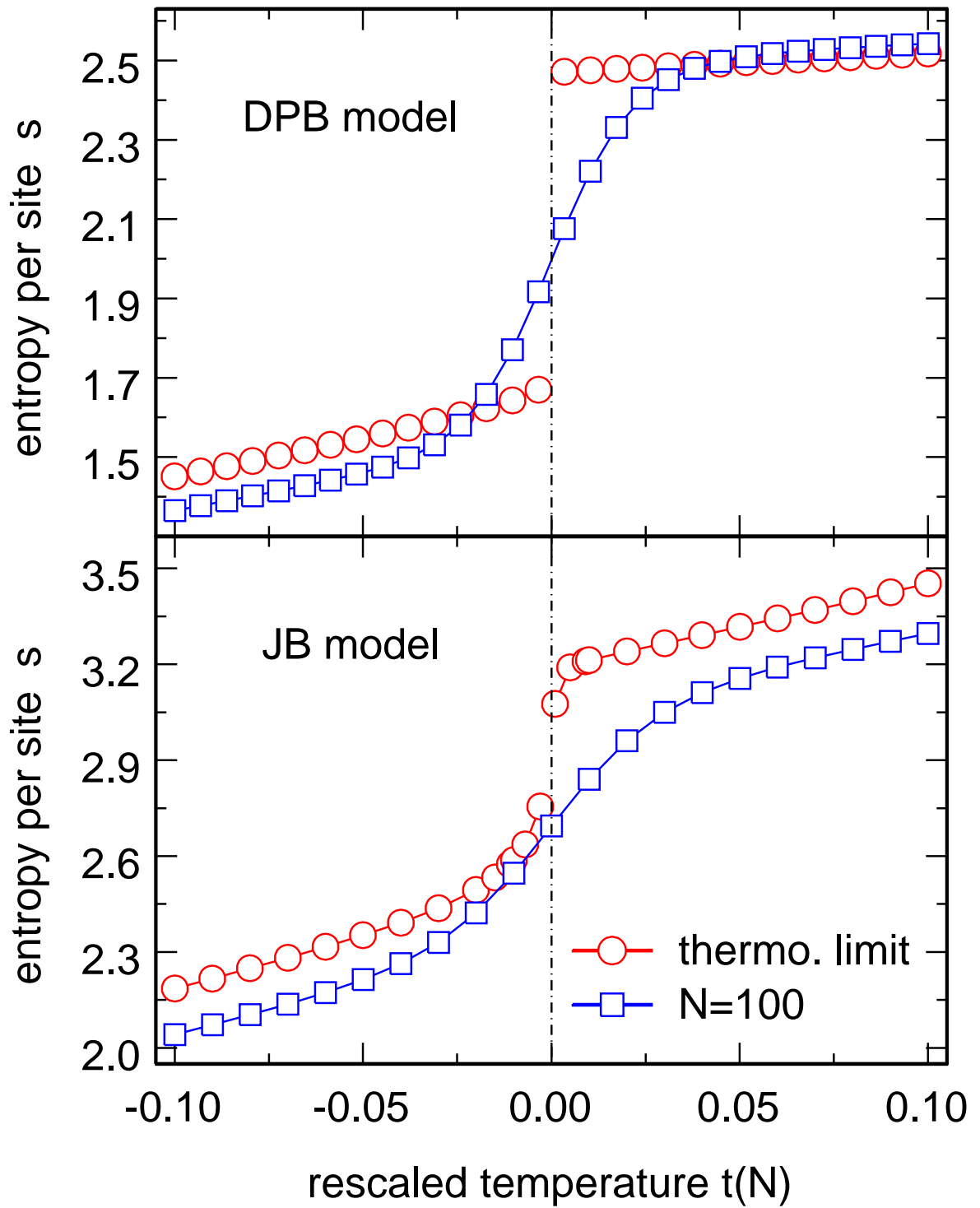


FIG. 2:

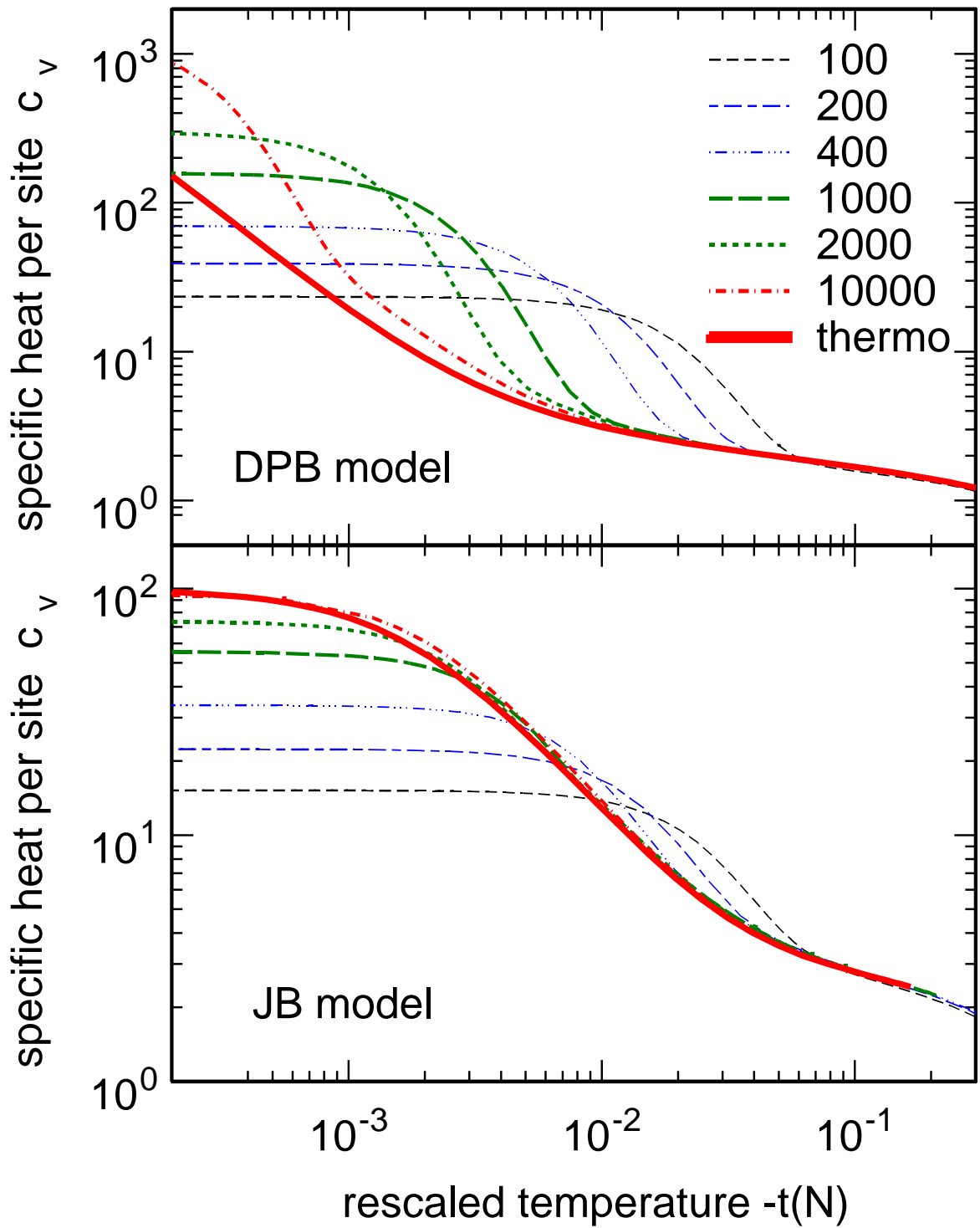


FIG. 3:

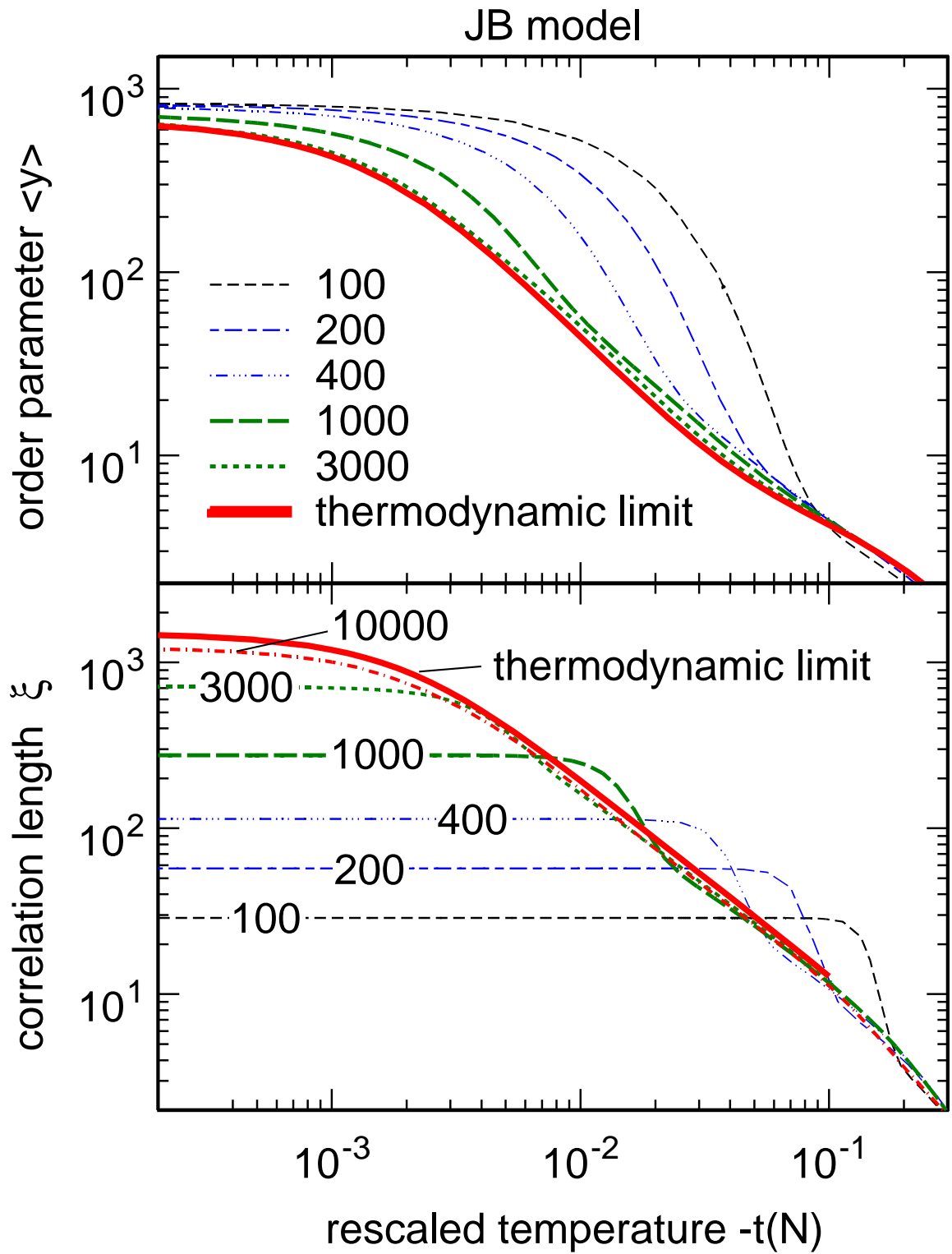


FIG. 4:

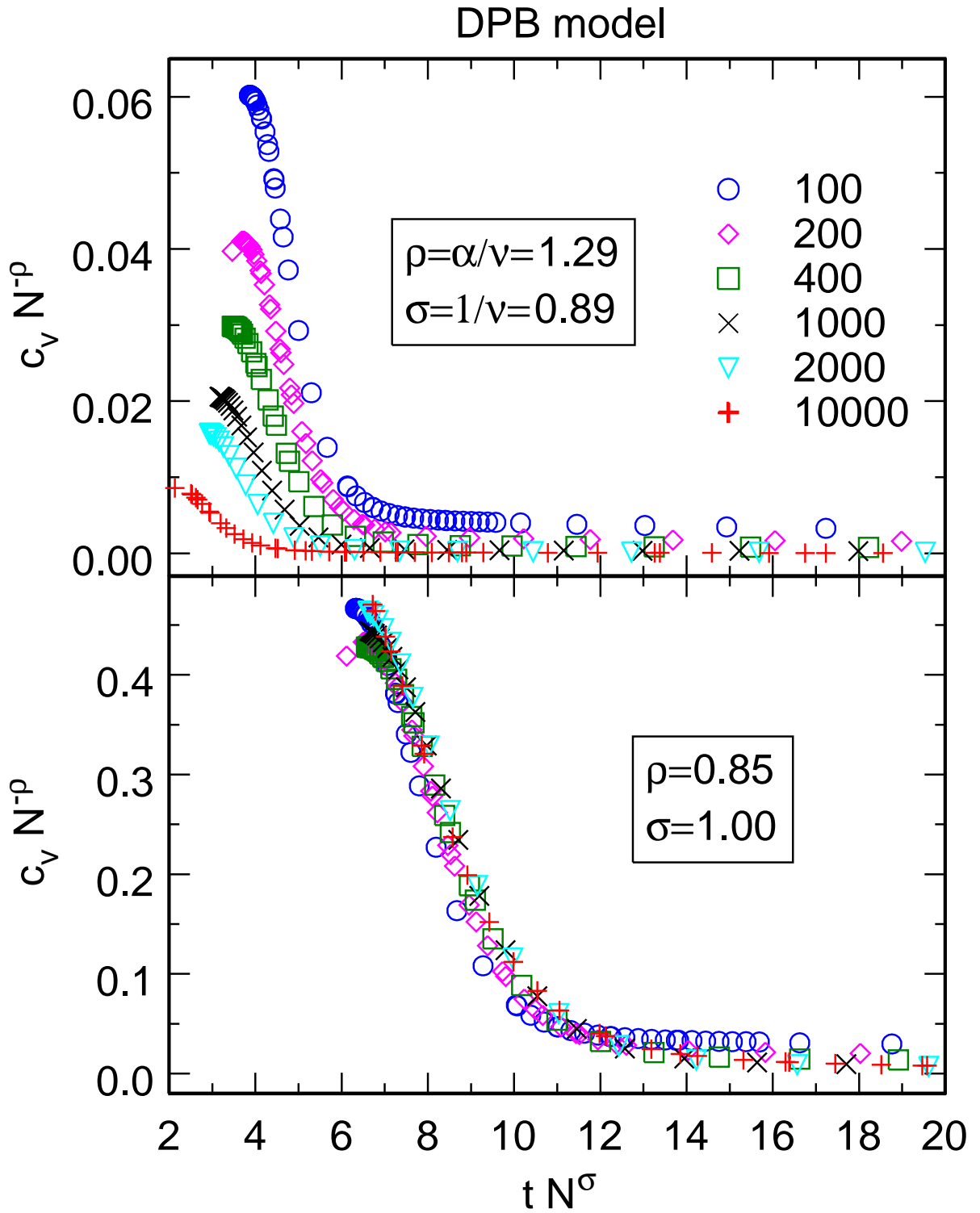


FIG. 5:

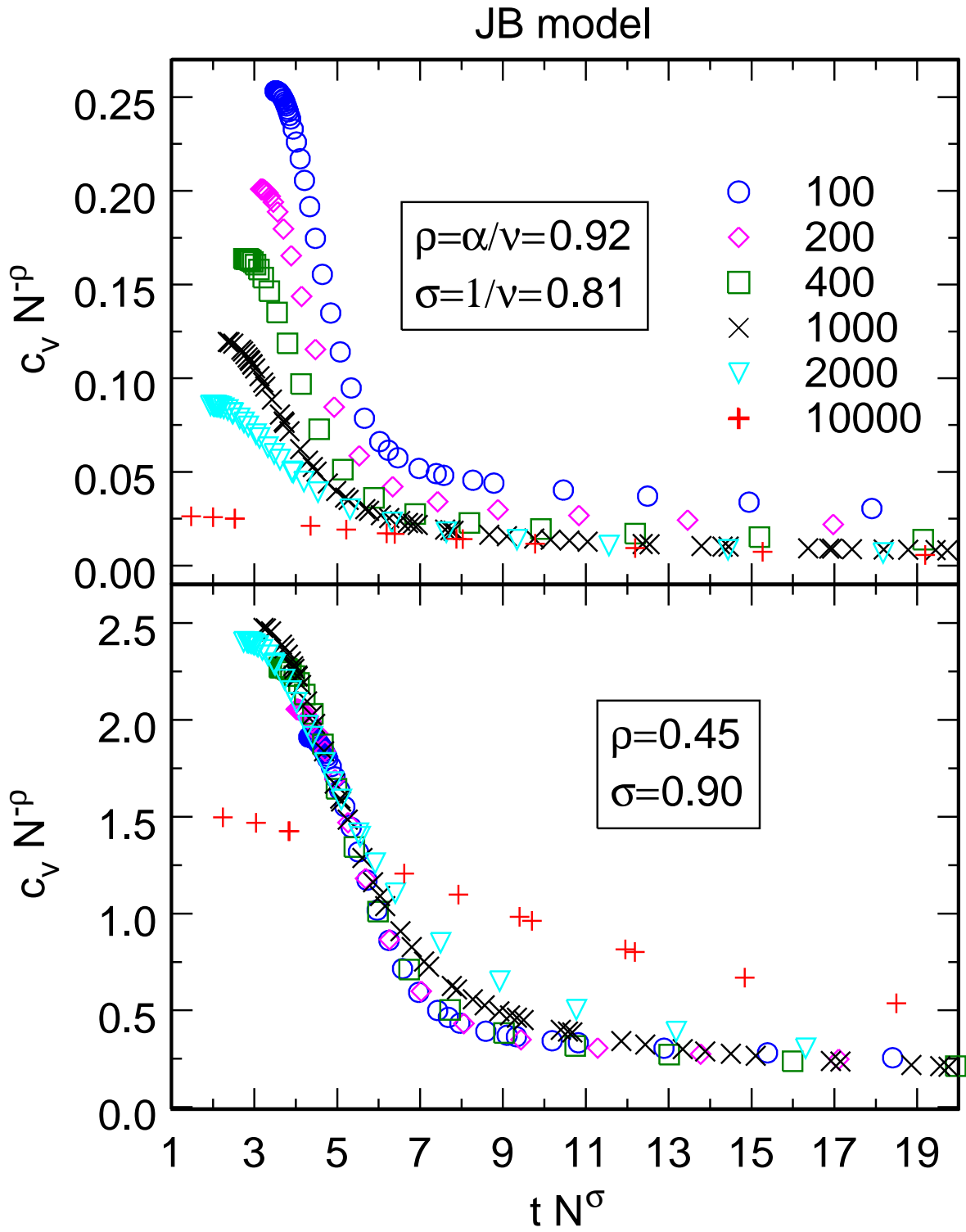


FIG. 6:

