PHYSICAL REVIEW E

STATISTICAL PHYSICS, PLASMAS, FLUIDS, AND RELATED INTERDISCIPLINARY TOPICS

THIRD SERIES, VOLUME 59, NUMBER 2 PART B FEBRUARY 1999

ARTICLES

Θ-point universality of random polyampholytes with screened interactions

Pietro Monari[1] and Attilio L. Stella[1,2]

[1]*INFM, Dipartimento di Fisica, Università di Padova, I-35131 Padova, Italy*

[2]*Sezione INFN, Università di Padova, I-35131 Padova, Italy*

(Received 9 July 1998)

By an efficient algorithm, we evaluate exactly the disorder-averaged statistics of globally neutral self-avoiding chains with quenched random charge $q_i = \pm 1$ in monomer i and nearest neighbor interactions $\propto q_i q_j$ on square (22 monomers) and cubic (16 monomers) lattices. At the Θ transition in two dimensions, the radius of gyration and entropic and crossover exponents are very compatible with the universality class of the corresponding transition of homopolymers. A further strong indication of such a class comes from direct comparison with the corresponding annealed problem. In three dimensions, classical exponents are recovered. The percentage of charge sequences leading to folding in a unique ground state approaches zero exponentially with the chain length. [S1063-651X(98)09212-5]

PACS number(s): 64.60.−i, 36.20.−r, 33.15.−e, 02.70.−c

During recent years random heteropolymers have been the object of intense study: they play a central role in our understanding of the properties of biologically active molecules [1], and represent a relatively handy model of a disordered system, of great interest in statistical mechanics [2–4]. Particular attention has recently been focused on randomly charged polymers (polyampholytes) [5] interacting via long- (Coulomb) and short-range (screened) potentials. Amino acids in proteins carry electric charges, and electrostatic interactions play an important role in determining their behavior [1]. Thus polyampholyte models are expected to be useful for the investigation of biologically relevant polymers.

Due to monomer-monomer and monomer-solvent interactions, even in the absence of charges, a linear polymer undergoes a collapse Θ transition as the temperature T is varied [6]. In particular, in the case of a homopolymer, for $T > T_\Theta$ the chain is swollen and its average radius of gyration grows as $R_g \propto N^\nu$ (N is the number of steps, $\nu = \frac{3}{4}$ in two dimensions [7], and $\nu = 0.588$ in three dimensions [8]), while for $T < T_\Theta$ the polymer is compact ($\nu = 1/d$). At $T = T_\Theta$ a distinct universality class ($\nu = \nu_\Theta = \frac{4}{7}$ in two dimensions [6] and $\frac{1}{2}$ in three dimensions [9]) characterizes scaling.

The existence of a collapse transition is by now well established for polyampholyte models with both short- and long-range interactions due to randomly distributed charges. In the long-range Coulomb case, the total charge distributed along the chain must not exceed a critical value ($\Delta Q \simeq \sqrt{N}$) in order for the Θ transition to occur [10]. In the short-range case there is a collapse transition as long as the charge unbalance $x = |N_+ - N_-|/(N_+ + N_-)$ is less than some cutoff value [11]. The collapse of randomly charged polymers owes much of its importance to the close connection with protein folding [12], and is similar to that of homopolymers, with a Θ-scaling regime separating compact and swollen phases. An important yet unsettled issue is that of establishing whether the Θ transition of polyampholytes falls into the same universality class as the collapse of homopolymers described above.

By a numerical study of randomly charged self-avoiding walks (SAW's) with nearest neighbor interactions on the square lattice, Kantor and Golding [13] obtained $\nu_\Theta = 0.60 \pm 0.02$ for globally neutral polyampholytes, slightly, but appreciably, different from that of homopolymers ($\nu_\Theta = \frac{4}{7}$). These results, based mostly on Monte Carlo (MC) enumerations (up to $N = 99$) and on a relatively limited sampling over disorder, suggest that the presence of quenched random interactions can modify the universality class of the Θ transition. This is quite an intriguing possibility, also in view of the constant focus on universality issues in the literature on Θ transitions [6]. A major limitation of this type of study is due to the need to perform averages over charge disorder, a task which, at an exact level, becomes impossible with standard algorithms as soon as $N > 15$ in two dimensions. It is not obvious whether, for relatively short chains, quenched averages over few disorder configurations (10–100 in Ref. [13]) should be sufficient for a satisfactory determination of

the radius of gyration. Especially in the low-temperature regime, fluctuations of thermodynamic quantities due to quenched disorder are indeed very large, and random sampling over a small number of disorder realizations can lead to inaccurate results for quenched averages. The investigation of the low-temperature phase of random polyampholytes is a formidable challenge for MC methods, because of the prohibitively large samplings it should require. On the other hand, even thermal averages become a problem at low T. Indeed, dynamic MC methods like the pivot algorithm are efficient in the high temperature regime but much less reliable as soon as the temperature is lowered below T_Θ and chains are compact [14]. Finally, Markov-chain-based MC methods do not allow a computation of entropic exponents, which would provide a more complete characterization of the universality of the transition.

The above considerations suggest that an interesting strategy in the study of random polyampholyte models can be that of extending as far as possible, by appropriate algorithms, the range of chain lengths within which we can still gather exact information. In the present work we perform exact enumerations up to chains with $N=21$ in two dimensions and $N=15$ in three dimensions for polyampholytes with charge disorder and nearest neighbor interactions, and carry out quenched averages over all disorder realizations. To this purpose we developed a powerful algorithm for short-range interacting SAW's with charge disorder, in such a way as to reduce by orders of magnitude the required computational effort, compared to more standard methods.

We represent each N-steps polymer chain configuration by a SAW ω ($|\omega|=N$) on a square or cubic lattice. In each of the $N+1$ visited sites sits a charged monomer. The Hamiltonian takes the form

$$H(\{q\},\omega)=-\sum_{\langle i,j\rangle} q_i q_j, \tag{1}$$

where $q_i=\pm 1$ is the charge carried by the ith monomer, and $\langle i,j\rangle$ indicates pairs of (nonconsecutive) nearest neighbor monomers. The sequence of the $N+1$ charges distributed along the chain is denoted by $\{q\}=\{q_0,q_1,\ldots,q_N\}$ and is assumed to be globally neutral ($N+1$ even). The quenched average squared radius of gyration is

$$R_g^2(N,T)=\frac{1}{N_q}\sum_{\{q\}} Z_{\{q\}}(N,T)^{-1}\left[\sum_{|\omega|=N} e^{-H(\{q\},\omega)/T} r_g^2(\omega)\right], \tag{2}$$

where N_q is the total number of charge sequences, $r_g(\omega)$ is the radius of gyration with respect to the center of mass of the configuration ω, and $Z_{\{q\}}(N,T)$ is the canonical partition function for a given realization $\{q\}$ of disorder:

$$Z_{\{q\}}(N,T)=\sum_{|\omega|=N} e^{-H(\{q\},\omega)/T}. \tag{3}$$

Near the Θ point, the radius of gyration is expected to obey the tricritical scaling form [15,16]

$$R_g^2(N,T)\simeq N^{2\nu_\Theta} f(N^{\phi_\Theta}\tau), \tag{4}$$

where $\tau=(T-T_\Theta)/T_\Theta$ is the temperature distance from the Θ transition, and ϕ_Θ is the crossover exponent. In the case of homopolymers in two dimensions, $\nu_\Theta=\frac{4}{7}$ and $\phi_\Theta=\frac{3}{7}$ [9].

The annealed partition function is defined as the average of Eq. (3) over all sequences:

$$Z_{(a)}(N,T)=\frac{1}{N_q}\sum_{\{q\}} Z_{\{q\}}(N,T), \tag{5}$$

while the quenched partition function is defined in terms of the quenched free energy:

$$Z_{(q)}(N,T)=\exp\left[\frac{1}{N_q}\sum_{\{q\}} \ln(Z_{\{q\}}(N,T))\right]. \tag{6}$$

Both annealed and quenched partition functions are expected to scale as

$$Z_{(a/q)}(N,T)\simeq \mu_{(a/q)}(T)^N N^{\gamma_{(a/q)}(T)-1}. \tag{7}$$

The connectivity μ is lattice and temperature dependent, while γ is the universal entropic exponent. As in the homopolymer case, for high T the exponent $\gamma(T)$ should identically take the value appropriate to SAW's in the swollen regime ($\gamma_{\rm SAW}=\frac{43}{32}$) [7]. In the presence of a Θ collapse at T_Θ, γ is expected to assume a different value γ_Θ, which is peculiar to the universality class of the transition. At lower temperatures, because of the globular shapes of the collapsed polymer, surface effects are present and can modify the above scaling form for Z with the appearance of an extra exponential factor besides μ^N [17].

It is not obvious *a priori* that μ and γ should take the same values for quenched and annealed problems. While this is plausible at relatively high T, where quenched disorder plays a minor role, discrepancies can be anticipated at low T. A main issue here is to establish whether T_Θ is still included in the high-T range.

Different entropic exponents have to be defined when polymers are subject to geometrical constraints: if the polymer chain is forced to live in half space by an impenetrable wall to which one of its ends is fixed, the critical entropic exponent assumes a value γ_1, different from the bulk γ [6]. The entropic behavior of SAW's at boundaries already played a major role in studies aimed at a precise characterization of the universality class of the Θ transition of homopolymers (in two dimensions, $\gamma_\Theta=\frac{8}{7}$, $\gamma_{1\Theta}=\frac{4}{7}$, and $\gamma_{11\Theta}=-\frac{4}{7}$ [18,19]).

The numerical study of entropic exponents is greatly facilitated by considering simultaneously data for bulk and boundary behavior [20]. Indeed, if the polymer is not adsorbed, the connectivity μ is insensitive to the presence of the boundary and remains the same for both bulk and boundary behavior of Z. Below we indicate by $Z^{\rm bulk}$ and $Z^{\rm half}$ the respective partition functions. Thus the ratio between bulk and boundary Z's scales as a power of the difference between the respective γ's, and does not depend on μ, whose estimation is then not necessary. Due to these circumstances, the determination of $\gamma-\gamma_1$ becomes easier and much more accurate.

TABLE I. Comparison between the number of different contact maps S_N and C_N in two dimensions for $N=7,8,\ldots,21$. Also, even values of N are reported for completeness.

N	S_N	C_N
7	40	2172
8	77	5916
9	211	16 268
10	423	44 100
11	1112	120 292
12	2308	324 932
13	5952	881 500
14	12 494	2 374 444
15	31 939	6 416 596
16	67 388	17 245 332
17	170 669	46 466 676
18	363 009	124 658 732
19	910 971	335 116 620
20	1 953 846	897 697 164
21	4 868 342	2 408 806 028

Here we call contact a pair of nonconsecutive monomers on nearest neighbor sites, i.e., two interacting monomers. The contact map of a given SAW configuration ω is the set of all contacts it contains:

$$X(\omega)=\{(i,j):|\{r\}\{\omega(i)\}-\{r\}\{\omega(j)\}|=1,|i-j|>1\}. \tag{8}$$

A contact map of an N-steps SAW can also be represented by an $(N+1)\times(N+1)$ matrix, whose (i,j) element is 1 or 0, according to whether the monomers i and j are interacting or not, respectively. For any given $\{q\}$, the energy of a configuration ω is fully determined by its contact map $X(\omega)$. Two configurations ω and ω' which are characterized by the same structure of contacts $[X(\omega)=X(\omega')]$ have the same energy for every $\{q\}$, and can be considered as equivalent. The set of all ω's of a given length can be partitioned into equivalence classes, each of them containing all the walks which are characterized by a given contact map. The number of equivalence classes is equal to the number, S_N, of distinguishable contact maps X_α, $\alpha=1,\ldots,S_N$. Each equivalence class $C_\alpha=\{\omega:X(\omega)=X_\alpha\}$ is characterized by its own degeneracy $g(\alpha)$ and cumulative squared radius of gyration $\rho_g^2(\alpha)$:

$$g(\alpha)=\sum_{\omega\in C_\alpha}1,\quad \rho_g^2(\alpha)=\sum_{\omega\in C_\alpha}r_g^2(\omega). \tag{9}$$

$g(\alpha)$ is expected to grow exponentially with the difference between the number of steps N and the number of contacts in X_α [21]. This means that S_N still grows exponentialy with N, but much more slowly than the total number of SAW's C_N (see Table I). In particular the ratio S_N/C_N is expected to approach zero exponentially.

The sum over configurations ω in Eqs. (2) and (3) can be replaced by the sum over equivalence classes, each of them taken with its own degeneracy and cumulative squared radius of gyration:

$$R_g^2(N,T)=\frac{1}{N_q}\sum_{\{q\}}Z_{\{q\}}(N,T)^{-1}\left[\sum_{\alpha=1}^{S_N}e^{-H(\{q\},X_\alpha)/T}\rho_g^2(\alpha)\right], \tag{10}$$

$$Z_{\{q\}}(N,T)=\sum_{\alpha=1}^{S_N}e^{-H(\{q\},X_\alpha)/T}g(\alpha). \tag{11}$$

In terms of computational cost, the last equations are considerably cheaper than Eqs. (2) and (3). The main improvement regards the thermal averages over configurations ω, which are made considerably faster, due to the fact that they involve summations over S_N rather than C_N terms. Detailed enumeration of all ω's for each given sequence $\{q\}$, would become unfeasible as soon as $N>15$, when computing exact averages over disorder.

In the present work, Eqs. (10) and (11) have been implemented by an efficient algorithm, in which SAW's of a given length are generated once and for all. The structure of contacts of each walk is registered on a binary map. Whenever a new walk is generated, its contact map is analyzed and sorted: if the contact configuration has already occurred, its degeneracy and cumulative gyration radius are updated; otherwise, a new contact map is added.

Once SAW's are fully enumerated, all contact maps X_α are stored together with their $g(\alpha)$ and $\rho_g^2(\alpha)$. Disorder averages of thermodynamic and geometric observables are then calculated over half the number of neutral sequences, the Hamiltonian invariant being under $\{q\}\to\{-q\}$.

On a DEC 600 DIGITAL work station, an exact enumeration of SAW's and a complete quenching over all sequences requires few minutes of CPU time for 16 monomer chains and about 130 hours for 22 monomer chains. The same algorithm was adapted later on in order to compute annealed averages over the same realizations of disorder. The computation of annealed averages is slightly faster than that of quenched averages. Thus, we could easily obtain exact results with annealed disorder for N up to 21 in two dimensions.

In order to study the Θ point we computed effective ν exponents

$$\nu(N,K,T)=\frac{1}{2}\ln\left[\frac{R_g^2(N,T)}{R_g^2(N-K,T)}\right]\ln\left[\frac{N}{N-K}\right]^{-1}. \tag{12}$$

In the $N\to\infty$ limit these curves should be step functions of T. However, at finite N, they show a rounded step. If the trends of approach of the $N\to\infty$ ν values in the high- and low-T phases are from opposite directions, curves (12) are expected to intersect among themselves in the neighborhood of the Θ point. They indeed show such behavior: effective exponents like $\nu(N,2,T)$ are monotonically increasing functions of N at high T and decreasing at not too low T. Linear extrapolation of these curves with respect to $1/N$, in the $1/N\to0$ limit, allows one to estimate an exponent $\nu_\infty(T)$ which is close to or even below the compact-polymer value $\nu=0.5$ for T just below the intersection region. On the other hand, $\nu_\infty(T)$ is almost equal to the swollen SAW value $\nu=0.75$ at high T (Fig. 1). Intersections of all the curves $\nu(N,K,T)$ occur in a small region of the (T,ν) plane, within which one can suppose the transition to be located. Follow-

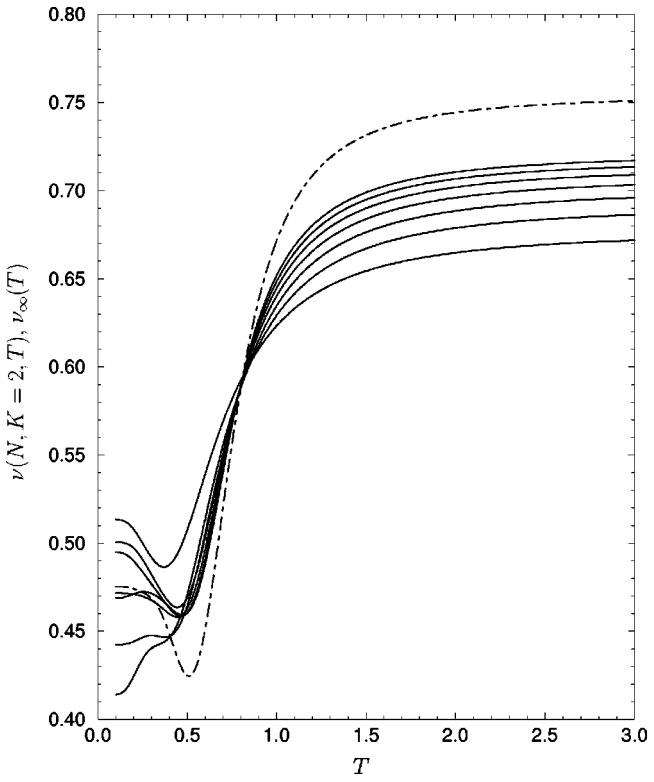

FIG. 1. Effective $\nu(N,K=2,T)$ exponents (solid lines) and their linear extrapolation $\nu_\infty(T)$ for $1/N\to 0$ (dot-dashed line). Temperature is normalized to monomer-monomer interaction. For sufficiently low T, the sequences cease to be monotonic. Of course, the relatively short length of the chains rounds off the expected steplike shape of ν_∞ at the Θ transition.

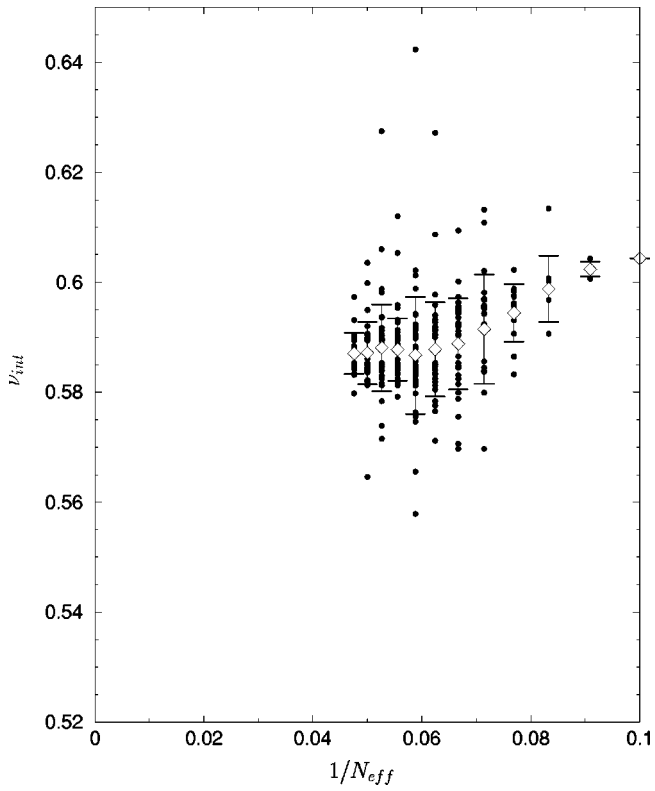

FIG. 2. Values of $\nu_{\rm int}$ as a function of $1/N_{\rm eff}$. Rhombs indicate the means of the exponent extimates at fixed $N_{\rm eff}$, while horizontal bars limit the variance of their distribution.

ing Privman [22], in order to pinpoint the Θ transition quantitatively, we calculated the coordinates $(T_{\rm int},\nu_{\rm int})$ of all intersections between every pair of curves $\nu(N,K,T)$, $\nu(N',K',T)$, and plotted these points against $1/N_{\rm eff}=2/(N+N')$. The definition of $N_{\rm eff}$ is of course subjective. In our choice no role is played by the integers K and K' because of the weak dependence of the intersection locations on these parameters. For each $1/N_{\rm eff}$, we computed the means, of $T_{\rm int}$ and $\nu_{\rm int}$ of the corresponding intersections, and extrapolated them linearly as a function of $1/N_{\rm eff}$ (Fig. 2), obtaining the estimates $\nu_\Theta=0.58\pm0.02$ and $T_\Theta=0.80\pm0.03$. Uncertainty estimates are also based on comparison between extrapolations from data in different ranges of $1/N_{\rm eff}$. The exponent is fully compatible with homopolymer Θ-point universality.

Another method can be applied in order to estimate ν_Θ and T_Θ. As illustrated above, the effective exponents $\nu_N(T)=\nu(N,2,T)$ are monotonic functions of $1/N$, decreasing for $T>T_\Theta$ and increasing for $T<T_\Theta$. Their linear correlation with respect to $1/N$ can be analyzed with the correlation coefficient defined by [23]

$$r(T)=\frac{\sum_N (1/N-\overline{1/N})(\nu_N(T)-\overline{\nu_N(T)})}{\sqrt{\sum_N (1/N-\overline{1/N})^2\sum_N (\nu_N(T)-\overline{\nu_N(T)})^2}}, \tag{13}$$

where bars indicate averages over N. The coefficient $r(T)$ is close to -1 for $T>T_\Theta$ and to 1 for $T<T_\Theta$ meaning that, in these regions, data are very well linear correlated and have opposite monotony. In the Θ region $r(T)$ undergoes a sudden jump between 1 and -1. Its derivative with respect to temperature shows a high and sharp peak whose mean value and width localize T_Θ and determine its uncertainty ΔT_Θ. The extremal values taken by $\nu_\infty(T)$ in the interval $[T_\Theta-\Delta T_\Theta, T_\Theta+\Delta T_\Theta]$ give an estimate of ν_Θ, and of the corresponding error $\Delta\nu_\Theta$. T_Θ and ν_Θ obtained with this method are almost identical to the values determined above by extrapolating the intersections $\nu_{\rm int}$ and $T_{\rm int}$, respectively.

In order to obtain the crossover exponent ϕ_Θ we analyzed the derivative of the squared radius of gyration with respect to temperature. Near the Θ point this quantity should scale as

$$\frac{d}{dT}R_g^2(N,T)\simeq N^{\phi_\Theta(T)+2\nu(T)}. \tag{14}$$

The effective exponent curves corresponding to $\phi+2\nu$ do not clearly intersect each other in a narrow region of the $(T,\phi+2\nu)$ plane. So the method used for determining ν_Θ cannot be applied in this case, because it would lead to ambiguous results. Following Ref. [22], we then calculated, for each intersection $(T_{\rm int},\nu_{\rm int})$ between $\nu(N,K,T)$ and $\nu(N',K',T)$, with $N>N'$, the quantity

$$\ln\left[\frac{dR_g^2(N,T_{\rm int})/dT}{dR_g^2(N-K,T_{\rm int})/dT}\right]\ln\left[\frac{N}{N-K}\right]^{-1}-2\nu_{\rm int}. \tag{15}$$

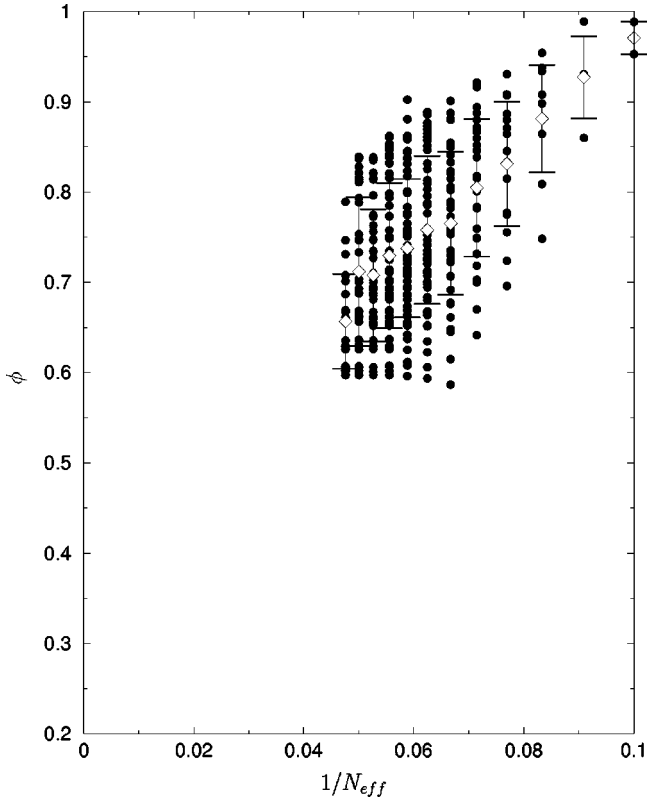

FIG. 3. Extrapolation of the crossover exponent ϕ_Θ.

Extrapolation of these data in $1/N_{\rm eff}$ leads to the estimate $\phi_\Theta = 0.40 \pm 0.08$ (Fig. 3). The computation of the crossover exponent for homopolymer Θ transitions is usually rather difficult and often leads to considerable overestimates [24,25]. Our result is very compatible with the exact homopolymer value $\phi_\Theta = 3/7 \simeq 0.42$ [9]. Attempts to determine ϕ on the basis of data collapse fits for R_g [Eq. (4)] were not very successful because the collapse quality does not depend sensibly enough on this exponent.

In the annealed system, frustration effects peculiar to quenched disorder are ruled out. The charges distributed along the chain are indeed free to rearrange among themselves in such a way to let nearest neighbor interactions be able to minimize the energy of each SAW configuration ω. It seems very plausible that such a rearrangement can produce a collapse in the same universality class as the Θ point of an ordered polymer with nearest neighbor attractive interactions for all monomers. Because of these reasons we expect annealed disorder to be irrelevant for the collapse transition. This conjecture is well confirmed by the analysis of our exact enumeration results for the annealed system (22 monomers). The analysis followed the lines of the quenched case. The transition exponents of the annealed model were estimated as $\nu_\Theta = 0.58 \pm 0.02$ and $\phi_\Theta = 0.41 \pm 0.08$.

A direct comparison between annealed and quenched partition functions then turns out to be a very significant test, in view of the fact that the annealed system represents a sort of substitute of the pure one. As explained above, to avoid difficulties due to the calculation of the nonuniversal constant μ, in the case of both quenched and annealed charges, we analyzed the ratio $\mathcal{Z}$ between the partitions of SAW's in the bulk and in the presence of the boundary, which is expected to scale as

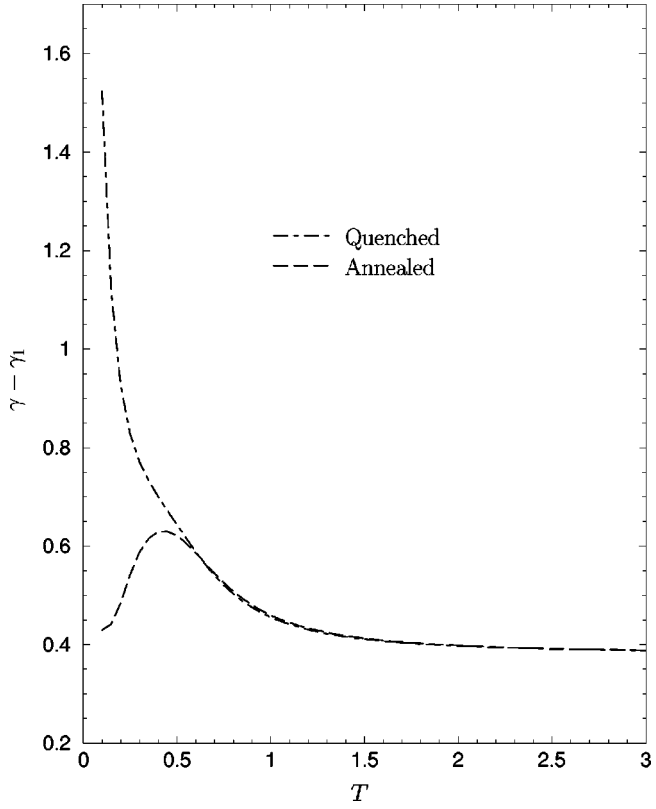

FIG. 4. Extrapolation of $\gamma - \gamma_1$ for annealed (dashed line) and quenched (dot-dashed line) disorder. The values are almost identical in a range of temperatures extending below $T_\Theta \sim 0.80$.

$$\mathcal{Z}_{(a/q)}(N,T) = \frac{Z^{\rm bulk}_{(a/q)}(N,T)}{Z^{\rm half}_{(a/q)}(N,T)} \simeq N^{(\gamma-\gamma_1)_{(a/q)}(T)}. \qquad (16)$$

Effective exponents can be obtained from

$$\left[\frac{\mathcal{Z}_{(a/q)}(N,T)}{\mathcal{Z}_{(a/q)}(N-2,T)}\right]^{1/2} = 1 + \frac{1}{N}(\gamma-\gamma_1)_{(a/q)}(N,T) + o\left(\frac{1}{N^2}\right). \qquad (17)$$

The sequences $(\gamma-\gamma_1)_{(a/q)}(N,T)$, plotted against $1/N$, show remarkably good linear correlation. Their extrapolation for $1/N \to 0$ gives a reasonable estimate of the expected $(\gamma - \gamma_1)_{(a/q)}$ in the high-T range and close to T_Θ. Even more precise is the comparison between the annealed and quenched cases based on these $\gamma - \gamma_1$ estimates. It turns out that the difference $\gamma - \gamma_1$ is almost identical for annealed and quenched systems on a range of temperatures which clearly extends below the Θ temperature (Fig. 4). We estimated $(\gamma - \gamma_1) \sim 0.50$ and $(\gamma - \gamma_1) \sim 0.39$ at the Θ point and in the high-T region, respectively. The Θ-point determination is slightly below the homopolymer value $[(\gamma - \gamma_1)_\Theta = \frac{4}{7}$ [18]], while the high-T one almost coincides with the SAW one: $[(\gamma - \gamma_1)_{\rm SAW} = 25/64$ [6]].

In three dimensions, for a homopolymer, ν_Θ is expected to be equal to $\frac{1}{2}$ with logarithmic corrections [9]. Indeed $d = 3$ is the upper critical dimension for the transition. We applied our methods to our model of random polyampholytes in three dimensions, and computed exact averages for chains up to 15 monomers. A simple analysis of the radius of gyra-

tion, not including logarithmic corrections, gives $\nu_\Theta = 0.51 \pm 0.04$, again consistent with the homopolymer universality class.

Altogether the above results give very strong evidence that the collapse transition of the globally neutral random polyampholyte model falls in the same universality class as the Θ point of homopolymers. Support for such a conclusion comes from the exponent determinations we were able to perform. Further evidence comes from our comparative analysis of entropic properties in the case of quenched and annealed disorders. Our study of $\gamma - \gamma_1$ shows that annealed and quenched partition functions start to deviate appreciably at some temperature falling definitively below the estimated T_Θ. In order to obtain a collapse with exponents different from those of homopolymer models, one should have conditions such that the effect of quenched disorder become important above or, at least, at the collapse transition temperature. The identification and investigation of models where such conditions could possibly be realized remains an important open issue in the field, whose solution would sensibly increment our understanding of the possible role played by chain disorder in polymer statistics.

With the exact enumeration methods developed for the study of the Θ transition, we could also perform an analysis of how the actual partition function at fixed $\{q\}$, $Z_{\{q\}}$, deviates from its (annealed) average at low temperature. Histograms of quantities like $Z_{\{q\}}(N,T)/Z_{(a)}(N,T)$ show very clearly a lack of self-averaging at T sufficiently lower than T_Θ. While in a range of high T including T_Θ they are narrow peaked around the value 1, for lower temperature they are quite broad. At very low T such histograms acquire a sparse structure, and allow one to investigate folding properties of the model.

While for T approaching zero homopolymers collapse to many compact conformations with the same ground state energy, most heteropolymer sequences usually collapse to very few lowest-energy conformations (see, e.g., Ref. [26]). In general, in order to better represent properties of real proteins, a heteropolymer model is expected to admit a unique compact conformation with lowest energy, i.e., a nondegenerate ground state, at least for some sequences. The percentage of sequences admitting a unique ground state for the *H-P* heteropolymer model is believed from numerical analysis to remain almost constant as N increases [27]. The *H-P* model is a SAW in which each monomer can have either a hydrophobic or a polar character, with short-range interactions to the nearest neighbor solvent molecules. This model has been applied often to protein folding studies (see, e.g., Ref. [26,28]). Here we investigated the number f_N of sequences having a unique ''native state'' in our two dimensional model. This analysis was performed by applying the exact method described above to the investigation of ground states of Hamiltonian walks [29] on the square lattice, for chain lengths up to 25. It turns out that f_N grows with N at a reduced exponential rate with respect to the total number of sequences N_q. In particular we found $f_N \simeq 1.85^N$, while $N_q \simeq N^{-1/2} 2^N$. Thus the percentage of sequences which possess a unique ground state tends asymptotically to zero as $N \to \infty$. This behavior is in sharp contrast with that found in the *H-P* model [27].

We thank F. Seno for useful discussions and for help with the analysis. We are also indebted to S. G. Whittington for discussions, and to C. Vanderzande for a critical reading of the manuscript. Partial support from the European Network under Contract No. ERBFMRXCT 980183 is also acknowledged.

[1] T. E. Creighton, *Proteins: Structure and Molecular Properties*, 2nd ed. (Freeman, New York, 1993).
[2] D. L. Stein, Proc. Natl. Acad. Sci. USA **82**, 3670 (1985).
[3] H. Frauenfelder and P. G. Wolynes, Phys. Today **47**(2), 58 (1994).
[4] T. Garel and H. Orland, Europhys. Lett. **6**, 307 (1988).
[5] Y. Kantor, M. Kardar, and H. Lin, Phys. Rev. E **49**, 1383 (1994).
[6] C. Vanderzande, *Lattice Models of Polymers* (Cambridge University Press, Cambridge, 1998).
[7] B. Nienhuis, Phys. Rev. Lett. **49**, 1062 (1982); J. Stat. Phys. **34**, 731 (1983).
[8] J. C. Le Guillou and J. Zinn-Justin, J. Phys. (France) **50**, 1365 (1989).
[9] B. Duplantier and H. Saleur, Phys. Rev. Lett. **59**, 539 (1987); B. Duplantier, J. Phys. (France) Lett. **41**, L-409 (1980).
[10] Y. Kantor and M. Kardar, Phys. Rev. E **51**, 1299 (1995); Europhys. Lett. **27**, 643 (1994); Phys. Rev. E **52**, 835 (1995).
[11] Y. Kantor and M. Kardar, Europhys. Lett. **28**, 169 (1994).
[12] *Protein Folding*, edited by T. E. Creighton (Freeman, New York , 1992).
[13] Y. Kantor and I. Golding, Phys. Rev. E **56**, R1318 (1997).
[14] M. C. Tesi, E. J. Janse van Rensburg, E. Orlandini, and S. G. Whittington, J. Stat. Phys. **82**, 155 (1996).
[15] M. Daoud and G. Jannink, J. Phys. (France) **37**, 973 (1976).
[16] P. G. De Gennes, J. Phys. (France) Lett. **39**, L-299 (1978).
[17] A. L. Owczareck, T. Prellberg, and R. Brak, Phys. Rev. Lett. **70**, 951 (1993).
[18] C. Vanderzande, A. L. Stella, and F. Seno, Phys. Rev. Lett. **20**, 1757 (1991).
[19] A. L. Stella, F. Seno, and C. Vanderzande, J. Stat. Phys. **73**, 21 (1993).
[20] F. Seno and A. Stella, Europhys. Lett. **7**, 605 (1988).
[21] J. Douglas, C. M. Guttman, A. Mah, and T. Ishinabe, Phys. Rev. E **55**, 738 (1997).
[22] V. Privman, J. Phys. A **19**, 3287 (1986).
[23] W. Feller, *An Introduction To Probability Theory And Its Applications* (Wiley, New York, 1978).
[24] I. Chang and H. Meirovitch, Phys. Rev. E **48**, 3656 (1993).
[25] P. Grassberger and R. Hegger, J. Phys. I **5**, 597 (1995).
[26] C. Camacho and D. Thirumalai, Phys. Rev. Lett. **71**, 2505 (1993).
[27] K. A. Dill and H. S. Chan, J. Chem. Phys. **95**, 3775 (1991).
[28] K. A. Dill *et al.*, Protein Sci. **4**, 561 (1995).
[29] E. Shakhnovich and A. Gutin, J. Chem. Phys. **93**, 5967 (1990).