

# Dealing with long range interactions in the determination of polyelectrolyte ionization properties. Extension of the transfer matrix formalism to the full range of ionic strengths

Josep L. Garcés<sup>(a),\*</sup>, Sergio Madurga<sup>(b)</sup>, Carlos Rey-Castro<sup>(a)</sup> and Francesc Mas<sup>(b), \*</sup>

22nd November 2016

(a) Chemistry Department. Technical School of Agricultural Engineering & AGROTECNIO of Lleida University (UdL), Lleida (Catalonia, Spain)

(b) Physical Chemistry Unit, Materials Science and Physical Chemistry Department & Research Institute of Theoretical and Computational Chemistry (IQTUCUB) of Barcelona University (UB), Barcelona (Catalonia, Spain)

(\*) corresponding authors; e-mails: jlgarcés@quimica.udl.cat; fmas@ub.edu

[www.ub.edu/biophyschem](http://www.ub.edu/biophyschem)

**Key words:** polyelectrolytes; ionization properties; long-range interactions; Debye-Hückel interactions; Ising model; Potts model; transfer matrix; Monte Carlo simulations

## Abstract

The ionization state of charged macromolecules in solution is usually determined by the extent of the binding processes. These processes are very sensitive to the ionic strength of the medium, which are of long-range nature. The ionization properties of weak polyelectrolytes can be described by means of Ising-type models, which is only feasible when long-range interactions are neglected. Here, this formalism is extended to include long-range interactions by introducing a modified free energy involving only effective short-range interaction parameters. These parameters can be systematically calculated by using the Gibbs-Bogoliubov variational principle. The technique is illustrated with the calculation of titration curves of homogeneous and heterogeneous polyelectrolytes in

a wide range of ionic strengths. The correction of the site protonation free energy (first order correction) is enough to obtain an excellent agreement between theory and Monte Carlo simulations. Corrections to other cluster parameters (higher order corrections) are also implemented. In general, the correction to a particular parameter represents the average change in the long-range energy when a new interaction is created in the polyelectrolyte. The method presented here represents an improvement in the description of the ionization state of polyelectrolytes that can be relevant in a wide range of areas.

## 1 Introduction

Binding of simple ions present in the background medium (protons, metal cations, etc.) controls the ionization state of weak polyelectrolytes in solution. This process is of paramount importance to understand the behaviour of charged macromolecules in a wide range of situations, from receptor-ligand interactions in biochemical systems [1, 2], to the design of advanced coatings in materials science, wastewater treatment, pharmaceutical industry, etc. [3, 4, 5], and the role of natural organic matter in the geochemical cycling of trace metals [6], to mention just a few examples. Acid-base equilibria in weak polyelectrolytes represent the paradigmatic case of charge regulation in weak polyelectrolytes (due to the ubiquitous nature of proton ions in aqueous solution) and its study has been a classical topic in the past [7, 8, 9]. Among the different theoretical approximations used to describe the ionization properties in these systems, the site binding (SB) model has proven to be one of the most productive frameworks [10, 11]. In this model, the ionization state of the molecule is defined as a set of protonating sites, which can adopt two possible states, *i.e.*, protonated and deprotonated. For a linear polyelectrolyte, the resulting treatment is equivalent to the classical Ising model of ferromagnetism in the presence of an external field [12, 13]. Potts models, a generalization of the Ising models which considers more than two possible states for the spins, have been applied to the calculation of conformational properties of linear polymers [14, 15], coupling of ionization and conformational properties [16] or metal binding to polyelectrolytes [17].

The protonated sites experience both short and long range interactions. The latter are usually mediated by the solvent and can be described by simple potentials, such as the Debye-Hückel potential, or by suitable mean field approximations [18, 19, 20, 21]. Short range interactions, however, can produce important correlations between the binding to neighbouring sites, so that the continuous mean field approach is no longer valid [22, 23]. Moreover, they are mediated by the molecular skeleton rather than by the solvent, so that they cannot in general be modelled by simple interaction potentials. As a consequence, parameters accounting for chemically specific interactions must be included in the model [19, 21, 24, 25]. Within the SB model, the problem can be tackled by expanding the free energy in a set of parameters (also known as “cluster” parameters) which involve interactions among an increasing number of sites: site protonation free energies, pair interaction energies, triplet (or three body) interaction energies,

etc. [16, 19, 26, 27, 28, 29, 30, 31]. In many cases, the resulting “cluster” expansion converges very fast to the exact free energy. If the number of sites is small ( $N \leq 20$ ), the necessary thermal averages can then be performed by direct enumeration. For systems with a large number of sites, direct enumeration becomes impracticable and other techniques, such as Monte Carlo (MC) simulation must be used [20, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42].

For linear polyelectrolytes, the transfer matrix formalism is probably the most elegant and powerful method to compute thermal averages of Ising and Potts models. The key point is to relate the partition function of a system with  $N+1$  units (spins, bonds or binding sites) to the partition of the system with  $N$  units. The resulting recurrence relationship can be expressed through the transfer matrix, whose elements are proportional to the conditional probability of finding the new unit in a particular state for a given state of the preceding one [12, 43]. When applied to the binding of ions to polyelectrolytes, the method can be adapted to include a wide range of phenomena such as triplet interactions between sites [13], chelate complexation of metal ions [17], proton binding to polyampholytes [44, 45], protein-DNA binding [46], super-capacitor charging [47], or coupling between ionization and conformational degrees of freedom [16, 21, 48]. This treatment, however, is only practicable if long range interactions are neglected, since the size of the transfer matrices grows exponentially with the range of the interactions [18]. In practice, this means that the study must be restricted to high ionic strengths, an important limitation specially for macromolecules which become insoluble under such conditions [49, 50, 51, 52].

In the present work, the transfer matrix method is modified in order to include long range interactions in an approximate but systematic way. With this aim, some necessary results on transfer matrices are recalled in section 2. The basic idea of the approach is presented in section 3. The original free energy is replaced by a new one involving only short range cluster parameters, which account for the long range interactions in an effective way. Equations for the new parameters are derived by using the Gibbs-Bogoliubov variational principle [43]. In the resulting formalism, both short and long range interactions are treated simultaneously. Although the results are valid for ionizable small molecules, surfaces or macromolecules of any architecture, they are particularly easy to implement for linear polyelectrolytes, since the transfer matrix method can be applied for the effective free energy in these systems as usual. The technique is illustrated with the calculation of titration curves of homogeneous and heterogeneous polyelectrolytes and polyampholytes in a wide range of ionic strengths (virtually, the method can be applied to the full range of ionic strengths achievable in practice). In the simplest form (which we will refer to as “first order correction”), the method computes effective site protonation free energies. The resulting isotherms reproduce with great accuracy the results obtained by MC simulations using the exact free energy. Corrections to other cluster parameters (higher order corrections), such as nearest neighbour interactions, are provided in section 4, although little significant improvement of the isotherms is obtained. The calculations are fast enough to perform straightforward fitting of binding parameters to experimental data.

## 2 Site binding (SB) model and transfer matrices

A particular protonation state (or microstate) of a system with  $N$  ionizable sites is characterized by a set of state variables  $s = \{s_i\}$  where  $s_i = 1$  if the site is protonated and  $s_i = 0$  if the site is deprotonated. The microstate free energy can be expressed as the so called ‘‘cluster expansion’’ in the form

$$\frac{F(s)}{\ln 10} = - \sum_i \text{p}K_i s_i + \sum_{i>j} \epsilon_{ij} s_i s_j + \sum_{i>j>k} \tau_{ijk} s_i s_j s_k + \dots \quad (1)$$

$\text{p}K_i$  is the common logarithm of the protonation constant of the site  $i$  given all other sites are deprotonated,  $\epsilon_{ij}$  is the interaction energy of sites  $i$  and  $j$ ,  $\tau_{ijk}$  represents the triplet interaction energy among sites  $i$ ,  $j$  and  $k$ , and so on. The energy is expressed in thermal units, i.e.,  $\beta = 1/k_B T = 1$ , and the interaction parameters are divided by a factor  $\ln 10$  in order to be comparable in the pH scale. In most cases, the cluster expansion converges very fast to the exact free energy. Note that the choice of the protonation variables is not unique. For instance, for polyanions such as polycarboxylic acids, it could seem more natural to work with the variables  $q_i = 1 - s_i$ , where  $q_i = 1$  where the site is charged (i.e. deprotonated) and  $q_i = 0$  is the site is uncharged (i.e. protonated). However, it has been shown that the choice of the protonation variables does not affect the resulting thermal averages if the cluster parameters are properly redefined [19]. The foregoing arguments and calculations are thus applicable regardless of the charge sign of the protonated sites. It is also worth mentioning that although the flexibility of the chain is not explicitly included in  $F(s)$ , it can be shown that the cluster expansion (1) is still exact if the cluster parameters are regarded as proper averages over the conformational degrees of freedom [16].

The probability of a particular microstate in the semi-grandcanonical ensemble is given by

$$p(s) = \frac{e^{-F(s)} a_{\text{H}}^n}{\Xi} ; n = \sum_i s_i \quad (2)$$

where  $n$  is the total bound protons,  $a_{\text{H}}$  is the proton activity, and  $\Xi = \sum_s e^{-F(s)} a_{\text{H}}^n$  is the semi-grandcanonical partition function. For the purposes of this work, it is convenient to work in terms of the reduced free energy

$$\frac{H(s)}{\ln 10} = \sum_i \mu_i s_i + \sum_{i>j} \epsilon_{ij} s_i s_j + \sum_{i>j>k} \tau_{ijk} s_i s_j s_k + \dots ; p(s) = \frac{e^{-H(s)}}{\Xi} \quad (3)$$

where  $\mu_i = (\text{pH} - \text{p}K_i) = \log(K_i a_{\text{H}})$  is the reduced chemical potential in the pH scale. The relevant physical properties can be obtained by performing the proper thermal averages, which in many cases can be expressed in terms of derivatives of  $\Xi$ , related to the reduced free energy as

$$\Xi = \sum_s e^{-H(s)} \quad (4)$$

For instance, the average degree of protonation of a particular site  $i$  is obtained by

$$\theta_i = \langle s_i \rangle = -\frac{\partial \log \Xi}{\partial \mu_i} = \frac{1}{\ln 10} \frac{\partial \Omega}{\partial \mu_i} \quad (5)$$

where  $\Omega = -\ln \Xi$  is the thermodynamic potential or free energy associated to the semi-grandcanonical ensemble. The average number of bound protons is given by

$$\nu = \left\langle \sum_i s_i \right\rangle = \frac{\partial \Omega}{\partial \ln a_{\text{H}}} \quad (6)$$

As we will see later, it is also important to determine the correlation of the protonation degrees of two sites  $i$  and  $j$ , which can be expressed as

$$h_{ij} = \langle s_i s_j \rangle = \frac{1}{\ln 10} \frac{\partial \Omega}{\partial \epsilon_{ij}} = -\frac{1}{(\ln 10)^2} \frac{\partial \Omega}{\partial \mu_i \partial \mu_j} \quad (7)$$

The computation of the partition function can be performed by direct enumeration of the microstates only for  $N \leq 20$ . Otherwise, methods borrowed from statistical mechanics must be used. For 2D systems, only a few cases can be exactly solved [12]. For 1D systems, however, transfer matrix techniques can be applied to a wide range of models, although only if long range interactions are neglected. The transfer matrix method consists of relating the partition function for a system with  $N+1$  sites with that of a system with  $N$  sites in a recursive way. The link between both partition functions can be shown to be the transfer matrices. As a result, the partition function can be expressed as [14]

$$\Xi = \mathbf{q} \mathbf{T}_1 \mathbf{T}_2 \dots \mathbf{T}_N \mathbf{p}^{\text{T}} \quad (8)$$

where  $\mathbf{T}_i$  is the transfer matrix of site  $i$ , whose elements  $T_{i;\sigma,\rho}$  are the Boltzmann factors corresponding to the increase in the reduced free energy. They are proportional to the conditional probability of adding a site  $i$  in a state  $\rho$  provided that the previous site  $i-1$  was in the state  $\sigma$ .  $\mathbf{q}$  and  $\mathbf{p}$  are proper initiating and terminating vectors which depend on the details of the end of the chain.

For identical sites eqn. (8) reduces to

$$\Xi = \mathbf{q} \mathbf{T}^N \mathbf{p}^{\text{T}} \quad (9)$$

where  $\mathbf{T} = \mathbf{T}_1 = \mathbf{T}_2 = \dots = \mathbf{T}_N$ . In the limit  $N \rightarrow \infty$ ,  $\Xi$  becomes proportional to  $\lambda^N$ , the maximum eigenvalue of  $\mathbf{T}$ , so that  $\Omega \sim -N \ln \lambda$  and the average degree of protonation  $\theta = \theta_i = \nu/N$  reads

$$\theta = \frac{\partial \ln \lambda}{\partial \ln a_{\text{H}}} \quad (10)$$

A useful alternative expression to (10), which involves derivatives of the transfer matrix instead of  $\lambda$ , is derived in the appendix. The result is

$$\theta = \mathbf{b} \tilde{\mathbf{T}} \mathbf{a}^{\text{T}} ; \quad \tilde{\mathbf{T}} = -\frac{1}{\lambda \ln 10} \frac{\partial \mathbf{T}}{\partial \mu} \quad (11)$$

If  $\mathbf{Q}$  is the matrix whose columns are the eigenvectors of  $\mathbf{T}$ ,  $\mathbf{a} = (Q_{1r}, Q_{2r}, Q_{3r}, \dots)$  represents the eigenvector corresponding to  $\lambda$ , and  $\mathbf{b}$  is the  $r$ th row of  $\mathbf{Q}^{-1}$ ,  $\mathbf{b} = (Q_{r1}^{-1}, Q_{r2}^{-1}, Q_{r3}^{-1}, \dots)$ .

An equivalent expression for the correlation function  $h_{ij}$  is given by

$$h_{i,i+k-1} = h_{1k} = \mathbf{b} \tilde{\mathbf{T}} (\mathbf{T}/\lambda)^{k-2} \tilde{\mathbf{T}} \mathbf{a}^T \quad (12)$$

In the simplest case, when only nearest neighbour interactions are considered, the transfer matrix adopts the form [19]

$$\mathbf{T} = \begin{pmatrix} 1 & z \\ 1 & zu \end{pmatrix} \quad (13)$$

where  $z = \hat{K}a_H$  is the reduced activity and  $\epsilon = -\log u$  is the interaction free energy between consecutive sites. The corresponding initial and final vectors are  $\mathbf{p} = (1, 1)$  and  $\mathbf{q} = (1, 0)$ , respectively. If next-nearest neighbour and triplet interactions are included, the transfer matrix can be shown to be [45]

$$\mathbf{T} = \begin{pmatrix} 1 & z & 0 & 0 \\ 0 & 0 & 1 & zu \\ 1 & zv & 0 & 0 \\ 0 & 0 & 1 & zuvw \end{pmatrix} \quad (14)$$

where  $\zeta \equiv \epsilon_{i,i+2} = -\log v$  represents the next-nearest neighbour interaction free energy and  $\tau = -\log w$  the triplet interaction free energy, only present when three adjacent sites are protonated. In this case,  $\mathbf{p} = (1, 1, 1, 1)$  and  $\mathbf{q} = (1, 0, 0, 0)$ .

### 3 Effective site protonation free energies: first order correction

If long range interactions are included in the free energy, approximate methods are necessary. When the range of the interactions is increased in one site, the size of the transfer matrix is multiplied by a factor of 2. For instance, in order to include next-nearest neighbour approximation we need the  $4 \times 4$  transfer matrix (14);  $8 \times 8$  matrices are needed to account for next-next nearest neighbour interactions, and so forth [18]. Therefore, the size of the matrices grows exponentially with the range of the interactions and the method becomes impractical. The approach here proposed to overcome this difficulty is to replace the original free energy by an approximate one which only includes short range interactions, but which accounts for the long range interactions in an effective way. With this aim, the Gibbs-Bogoliubov variational principle [43] will be used. Let us assume that  $H(s)$  can be expressed as a sum of two terms

$$H(s) = H_0(s, \alpha) + \Delta H(s, \alpha) \quad (15)$$

so that  $H_0(s, \alpha)$  can be exactly solved, and  $\alpha = \{\alpha_k; k = 1 \dots P\}$  is a set of  $P$  parameters, not necessarily present in  $H(s)$ .  $\alpha_k$  are chosen by means of

some physical criterion such that  $\Delta H$  is “small” compared to  $H_0$ . Then, it can be shown that if  $H(s)$  is replaced by the new microstate reduced free energy  $\tilde{H} = H_0 + \langle \Delta H \rangle_0$ , the Gibbs-Bogoliubov inequality holds

$$\Omega \leq \tilde{\Omega} \equiv \Omega_0(\alpha) + \langle \Delta H(s, \alpha) \rangle_0 \quad (16)$$

where  $\Omega_0(\alpha) = -\ln \Xi_0$  and  $\langle \dots \rangle_0$  represent the free energy and the thermal average corresponding to  $H_0$ , respectively.  $\tilde{\Omega}$  is thus an upper bound of  $\Omega$ , so that the best approximation to  $\Omega$  is obtained by choosing  $\alpha_k$  such that  $\tilde{\Omega}$  is minimum

$$\left[ \frac{\partial \tilde{\Omega}}{\partial \alpha_k} \right]_{\alpha=\alpha_{\text{opt}}} = 0 ; k = 1, 2, \dots, P \quad (17)$$

The thermal averages are then approximated by using  $\tilde{H}(s, \alpha_{\text{opt}})$  instead of  $H(s)$ , which in turn coincide with the thermal averages calculated using  $H_0(s, \alpha_{\text{opt}})$ , which is exactly solvable

$$\langle \dots \rangle \simeq \sum_s \frac{e^{-\tilde{H}}}{\Xi} (\dots) = \sum_s \frac{e^{-(H_0 + \langle \Delta H \rangle_0)}}{\Xi_0 e^{-\langle \Delta H \rangle_0}} (\dots) = \langle \dots \rangle_0 \quad (18)$$

### 3.1 Identical sites

Let us consider a linear chain composed by  $N$  identical but interacting sites. The protonated sites experience short range interactions with  $m$  neighbouring sites and long range interactions with the rest of sites by a potential  $\phi_{ij}$ . Let us also assume that there is no interaction between different chains. The reduced free energy is expressed as the sum of two terms

$$\frac{H}{\ln 10} = \left( \mu \sum_i s_i + \epsilon \sum_i s_i s_{i+1} + \zeta \sum_i s_i s_{i+2} + \tau \sum_i s_i s_{i+1} s_{i+2} + \dots \right) + \left( \sum_{j>i+m} \phi_{ij} s_i s_j \right) \quad (19)$$

The first brackets contain the short range interactions and the second ones the long range interactions. Let us split  $H$  into two parts  $H = H_0 + \Delta H$  defined as

$$\frac{H_0}{\ln 10} = (\mu + \mu') \sum_i s_i + \epsilon \sum_i s_i s_{i+1} + \zeta \sum_i s_i s_{i+2} + \tau \sum_i s_i s_{i+1} s_{i+2} + \dots$$

$$\frac{\Delta H}{\ln 10} = \sum_{j>i+m} \phi_{ij} s_i s_j - \mu' \sum_i s_i \quad (20)$$

In doing so we try to find a correction  $\mu'$  to the site protonation reduced free energy, so that the main effects of the long range interactions are embedded in  $H_0$ .  $\mu'$  can also be understood as a correction to the site pK-value, so that  $\text{pK}^{\text{eff}} = \text{pK} - \mu'$  is the effective pK-value. The exact free energy  $\Omega$  is then replaced by

$$\varphi = \left\langle \sum_{j>i+m} \phi_{ij} s_i s_j \right\rangle_0 = \sum_{j>i+m} \phi_{ij} h_{ij}^0 \quad (21)$$

is the long range energy averaged over the unperturbed system, whose correlation function (7) is  $h_{ij}^0$ . Taking  $\partial\tilde{\Omega}/\partial\mu' = 0$ , one obtains for  $\mu'$

$$\mu' \ln 10 = \frac{d\varphi/d\mu'}{d\nu_0/d\mu'} = \frac{d\varphi/d\mu}{d\nu_0/d\mu} = \frac{d\varphi}{d\nu_0} \quad (22)$$

where  $\nu_0 = \langle \sum_i s_i \rangle_0 = (\partial\Omega_0/\partial\mu') / \ln 10 = (\partial\Omega_0/\partial\mu) / \ln 10$  is, according to (6), the average number of protons for the unperturbed model. An alternative to eqn. (22) is to determine  $\mu'$  by directly minimizing (??) using a suitable optimization routine. Eqn. (22), however, provides a transparent interpretation for  $\mu'$ , which results to be the average change of long range interaction energy involved in adding a new proton to the chain (calculated using  $H_0$ ). Note that no consideration has been made about the shape or dimensionality of the system, so that eqn. (22) is valid for small molecules, surfaces or polyelectrolytes. For surfaces, not much is known about  $\Omega_0$  for arbitrary  $\mu$ -values [12, 53]. However, in problems involving linear chains, short range interactions can always be treated by means the suitable transfer matrix, with the reduced activity  $z$  replaced by

$$z^{\text{eff}} = z 10^{-\mu'} \quad (23)$$

where  $\mu'$  is calculated by using (21-22). For  $N \rightarrow \infty$ ,  $\varphi$  becomes

$$\varphi = N \sum_{j>m} \phi_{1j} h_{1j}^0 \quad (24)$$

which can be evaluated by means of eqn. (12). If only nearest neighbour interactions are considered, the maximum eigenvalue  $\lambda_0$  of the transfer matrix (13) can be analytically evaluated and, using (10) or (11), the average degree of protonation  $\theta_0 = \nu_0/N$  remains [30]

$$\begin{aligned} \theta_0 &= [2 + (\lambda_0/z^{\text{eff}}) (1 - z^{\text{eff}}u) / (1 - u + \lambda_0u)]^{-1} \\ \lambda_0 &= (1 + z^{\text{eff}}u) / 2 + \sqrt{z^{\text{eff}} + (1 - z^{\text{eff}}u)^2 / 4} \end{aligned} \quad (25)$$

If the correlation between the long range interacting binding sites is neglected, we can assume that for  $j > i + m$

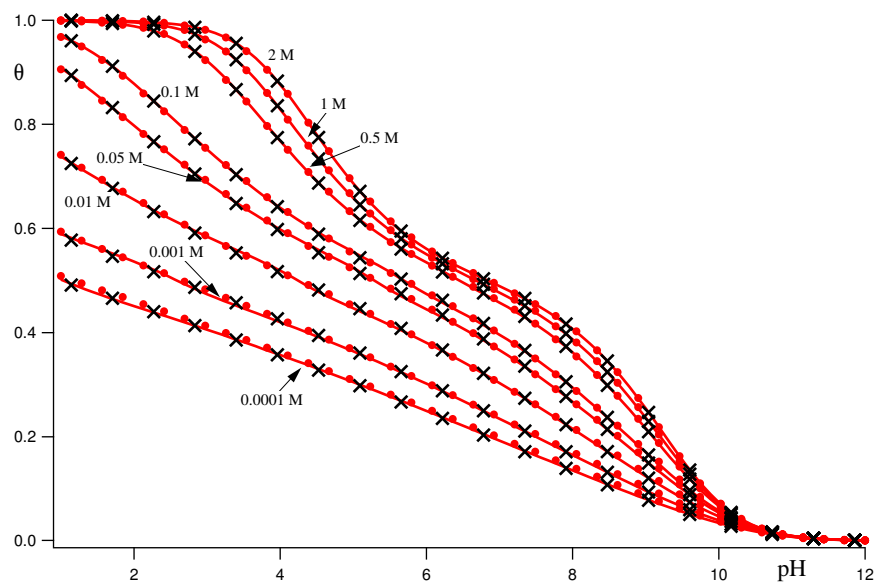
$$h_{ij}^0 = \langle s_i s_j \rangle_0 \simeq \langle s_i \rangle_0 \langle s_j \rangle_0 = \theta_0^2 \quad (26)$$

Introducing (26) in (22) we obtain

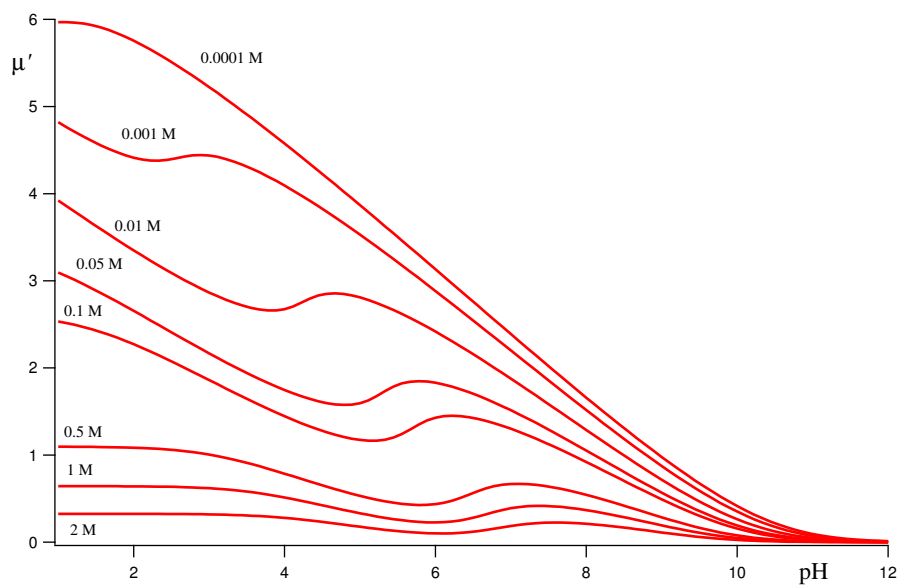
$$\mu' = 2\rho\theta_0 ; \rho = \sum_{i=m+1, \dots, \infty} \phi_{1i} \quad (27)$$

which together with (23-25) generalizes the Frumkin isotherm, since it includes long range interactions within the mean field approximation, while short range interactions are treated in an explicit way. However, due to the drastic approximation (26), eqns. (25-27), unlike eqns. (22-24), are not accurate for the whole range of ionic strengths.





(a) Fig. 1a



(b) Fig. 1b

Figure 1: (a) Titration curves corresponding to a polyelectrolyte with identical sites experiencing nearest neighbour interactions. Long range interactions are described by the Debye-Hückel potential. The ionic strengths range from 2M to  $10^{-4}$ M. Dots: MC simulations using the exact free energy; continuous lines: first order correction; crossed markers: second order correction. (b) First order corrections to the site protonation free energy,  $\mu'$ . The chosen parameters are  $pK = 9$ ,  $\epsilon = 2$  and  $b=0.2$  nm, which are typical values for polyethylene(imine) [16, 54].

Fig. 1a shows the titration curves corresponding to a linear chain with identical sites and nearest neighbour interactions at the short range level. As discussed above, short range interactions are mediated by the molecular skeleton rather than the solvent [19], so that specific values for them are necessary. Long range interactions are described by the Debye-Hückel potential for an extended linear chain

$$\phi_{ij} = \frac{1}{\ln 10} \frac{\ell_B e^{-\kappa b|j-i|}}{b|j-i|} ; \quad j > i + 1 \quad (28)$$

where  $\ell_B \simeq 0.7$  nm is the Bjerrum length in water at 298 K,  $b$  is the separation between consecutive protonating sites and  $\kappa^{-1}$  (nm) =  $0.304/\sqrt{I(\text{M})}$  is the Debye length. Note that the choice of the DH potential implies that the distance of two interacting sites is constant so that the chain is assumed to be rigid. Consequences of the chain flexibility in the ionization properties are analyzed, for instance, in [16, 21, 55] and in [56, 57, 58, 59]. The chosen parameters are  $pK = 9$ ,  $b=0.2$  nm and  $\epsilon = 2$ . They are typical values for polyethylene(imine) [16, 54]. Dots represent the values obtained from MC simulations, while continuous lines correspond to those obtained by using eqns. (21-25). The agreement between theory and MC simulations is excellent. Surprisingly they can be regarded as exact from the practical point of view. It is worth noting that the variation of the profiles reveals the importance of including the long range interactions even for relatively high ionic strengths (0.1 M), being completely necessary at low ionic strengths. In Fig. 1b the correction to the protonation free energy,  $\mu'$  is depicted. As expected, it increases in lowering the pH (increase in the charge), and in decreasing the ionic strength (lower electrostatic screening). Note that the wavy behaviour at pH values for which  $\theta \simeq 0.5$  corresponds to the maximum correlation between protonated and deprotonated sites, which tend to appear alternated in the chain [30]. As will be shown in section 4, this behaviour disappears when higher order corrections are included, since part of the effect is embedded in the effective nearest neighbour interactions.

Fig. 2 shows the titration curves corresponding for the same model, but now triplet interactions and next-nearest interactions have been added at the short range level. The chosen parameters are  $\tau = 0.43$  while  $\zeta$  has been calculated for each curve using the Debye-Hückel potential (28) taking  $|j-i| = 2$ , and introduced in the transfer matrix (14), so that correlations between next-nearest neighbour sites are treated in an explicit way. Triplet interactions are necessary to explain the titration curves of some small molecules and polyelectrolytes [13, 19, 28, 30, 54]. Recent papers have reported that they can be understood as the result of the coupling of conformational and ionization degrees of freedom [16, 21]. Dots represent MC simulations, while continuous lines come from eqns. (21-22), (23) and (10), together with the transfer matrix (14). A very good agreement is again obtained. Finally, we would like to comment that proper modifications of the free energy (19) could account for the interaction between two or several chains, by including the suitable interacting terms. Interaction between polyelectrolytes could drive interesting phenomena, such as charge regulation or phase separation, previously reported in colloidal suspen-

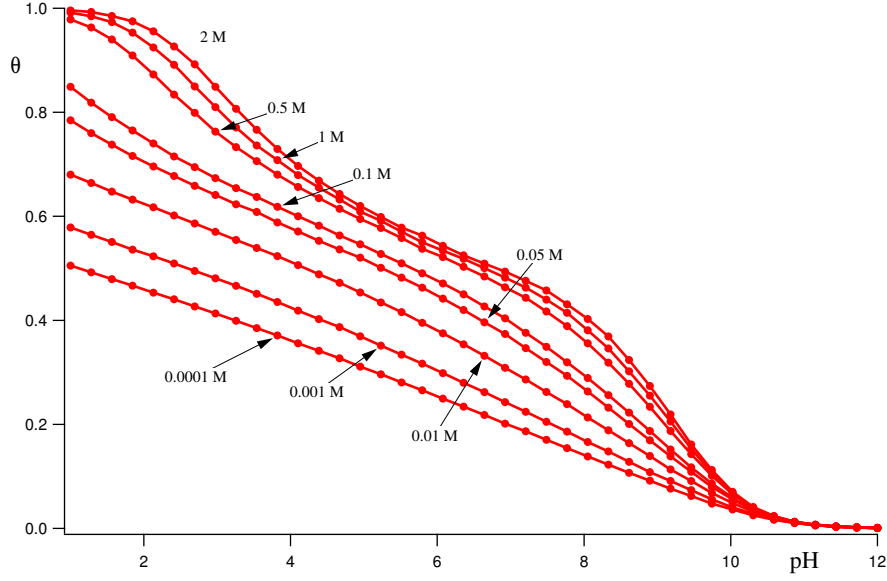


Figure 2: Titration curves corresponding to a polyelectrolyte with identical sites experiencing nearest neighbour, next-nearest neighbour and triplet interactions. The ionic strength range from 2M to  $10^{-4}$ M. Long range and next-nearest neighbour interactions are calculated by using the Debye-Hückel potential. The chosen parameters are  $pK = 9$ ,  $\epsilon = 2$ ,  $\tau = 0.43$  and  $b=0.2$  nm. Dots: values obtained with MC simulations; continuous lines: results obtained using eqns. (21-22), (23), (10) and transfer matrix (14).

sions [23, 60, 61, 62, 63]. This important topic is, however, out of the scope of this work.

### 3.2 Heterogeneous polyelectrolytes

The treatment presented in the previous section can be straightforwardly adapted to molecules with arbitrary shape and chemical composition, by correcting each site protonation free energy  $pK_i$ ,  $i = 1, \dots, N$ , so that  $pK_i^{\text{eff}} = pK_i - \mu'_i$ . As an illustrative example, let us consider a linear polyelectrolyte with two different kind of sites, A and B, such that they form the alternating structure ABABAB... The corresponding reduced free energy is given by

$$\frac{H}{\ln 10} = \mu_A \sum_{i=1,3,5,\dots} s_i + \mu_B \sum_{i=2,4,6,\dots} s_i + \epsilon \sum_i s_i s_{i+1} + \sum_{j>i+1} \phi_{ij} s_i s_j \quad (29)$$

We need corrections to both site protonation free energies,  $\mu'_A$  and  $\mu'_B$ , and

$H = H_0 + \Delta H$  is splitted as

$$\begin{aligned} \frac{H_0}{\ln 10} &= (\mu_A + \mu'_A) \sum_{i=1,3,5,\dots} s_i + (\mu_B + \mu'_B) \sum_{i=2,4,6,\dots} s_i + \epsilon \sum_i s_i s_{i+1} \\ \frac{\Delta H}{\ln 10} &= \sum_{j>i+1} \phi_{ij} s_i s_j - \mu'_A \sum_{i=1,3,5,\dots} s_i - \mu'_B \sum_{i=2,4,6,\dots} s_i \end{aligned} \quad (30)$$

Introducing (30) in (16) the trial free energy  $\tilde{\Omega}$  is given by

$$\frac{\tilde{\Omega}}{\ln 10} = \frac{\Omega_0(\mu'_A, \mu'_B)}{\ln 10} + \varphi(\mu'_A, \mu'_B) - \mu'_A \nu_A^0 - \mu'_B \nu_B^0 \quad (31)$$

where  $\nu_A^0 = \langle \sum_{i=1,3,5,\dots} s_i \rangle_0$  and  $\nu_B^0 = \langle \sum_{i=2,4,6,\dots} s_i \rangle_0$  are the average number of protonated sites of type A and B calculated using  $H_0$ . Taking derivatives in (31) with respect  $\mu'_A$  and  $\mu'_B$  and setting to zero, we obtain

$$\mathbf{J}_1 \begin{pmatrix} \mu'_A \\ \mu'_B \end{pmatrix} = \begin{pmatrix} \left( \frac{\partial \varphi}{\partial \mu'_A} \right)_{\mu'_B} \\ \left( \frac{\partial \varphi}{\partial \mu'_B} \right)_{\mu'_A} \end{pmatrix} ; \mathbf{J}_1 = \begin{pmatrix} \left( \frac{\partial \nu_A^0}{\partial \mu'_A} \right)_{\mu'_B} & \left( \frac{\partial \nu_B^0}{\partial \mu'_A} \right)_{\mu'_B} \\ \left( \frac{\partial \nu_A^0}{\partial \mu'_B} \right)_{\mu'_A} & \left( \frac{\partial \nu_B^0}{\partial \mu'_B} \right)_{\mu'_A} \end{pmatrix} \quad (32)$$

where

$$\begin{aligned} \nu_A^0 &= (\partial \Omega_0 / \partial \mu'_A) / \ln 10 = \nu_A^0(\mu'_A, \mu'_B) \\ \nu_B^0 &= (\partial \Omega_0 / \partial \mu'_B) / \ln 10 = \nu_B^0(\mu'_A, \mu'_B) \end{aligned} \quad (33)$$

have been used. Again, we can numerically solve the system of equations (32) or, alternatively, find  $\mu'_A$  and  $\mu'_B$  by direct optimization of  $\tilde{\Omega}$  using a suitable routine.

Eqn. (32) can be modified such that it provides a direct physical interpretation for  $\mu'_A$  and  $\mu'_B$ . Eqn. (33) states that, for given values of  $\mu'_A$  and  $\mu'_B$ , there corresponds a unique pair of values of  $\nu_A^0$  and  $\nu_B^0$ . Let us now suppose that we work in terms of the new variables  $\nu_A^0$  and  $\nu_B^0$  instead of  $\mu'_A$  and  $\mu'_B$ . Since the matrix  $\mathbf{J}_1$  is the jacobian of the transformation (33), eqns. (32) for  $\mu'_A$  and  $\mu'_B$  can be inverted and rewritten in terms of  $\nu_A^0$  and  $\nu_B^0$  in the simpler form

$$\begin{pmatrix} \mu'_A \\ \mu'_B \end{pmatrix} = \mathbf{J}_1^{-1} \begin{pmatrix} \left( \frac{\partial \varphi}{\partial \mu'_A} \right)_{\mu'_B} \\ \left( \frac{\partial \varphi}{\partial \mu'_B} \right)_{\mu'_A} \end{pmatrix} = \begin{pmatrix} \left( \frac{\partial \varphi}{\partial \nu_A^0} \right)_{\nu_B^0} \\ \left( \frac{\partial \varphi}{\partial \nu_B^0} \right)_{\nu_A^0} \end{pmatrix} \quad (34)$$

which states that  $\mu'_A$  ( $\mu'_B$ ) can be interpreted as the change in the long range free energy  $\varphi$  when a new A (B) site is protonated, keeping constant the number of protonated B (A) sites. For linear polyelectrolytes  $H_0$  can be solved using suitable transfer matrices. For the alternating sites model  $\mathbf{T}$  is given, as in (13), by

$$\mathbf{T} = \mathbf{T}_A \mathbf{T}_B ; \mathbf{T}_A = \begin{pmatrix} 1 & z_A^{\text{eff}} \\ 1 & z_A^{\text{eff}} u \end{pmatrix} ; \mathbf{T}_B = \begin{pmatrix} 1 & z_B^{\text{eff}} \\ 1 & z_B^{\text{eff}} u \end{pmatrix} \quad (35)$$

where  $z_i^{\text{eff}} = z_i 10^{-\mu'_i}$  are the effective reduced activities, corrected using eqns. (32-33), and  $\epsilon = -\log u$  represents the nearest neighbour interaction energy between A and B sites. The resulting titration curves are compared to MC

simulations in Fig. 3a. with parameters  $pK_A = 7$ ,  $pK_B = 9$ ,  $\epsilon = 1.5$ , and  $b=0.2$  nm, which represent alternating copolymers with two different basic functional groups. The agreement is very good for the full range of ionic strengths. In Fig. 3b the corrections to the site protonation energies are plotted as a function of pH. As expected, they increase in decreasing the ionic strength due to the lowering of the screening of coulombic forces. At ionic strengths lower than 0.1M, the correction in the protonation free energy at  $\theta \simeq 0.5$  becomes larger than 1 pK-unit and the effect of the long range interactions cannot be neglected. Note that the correction is larger for B sites than for A sites, since the latter present a lower affinity for the proton.

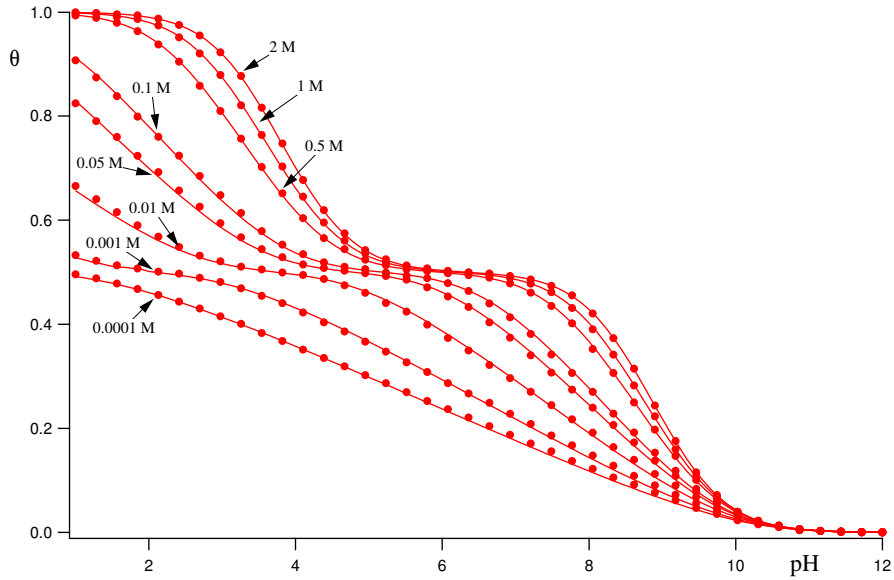
It is also interesting to test the method if, not only repulsive, but also attractive interactions are present. Attractive interactions are expected to produce high correlations between groups with opposite charge so that their presence is a good test to the present treatment. Let us consider a polyampholyte composed by two kind of alternating groups, forming a structure ABABAB... A groups are acidic (for instance, carboxylic groups) and negatively charged at high pH-values, while they are not charged at low pH-values. B groups have basic character (for instance, amino groups) and they are positively charged at low pH-values, while they are uncharged at high pH-values. The corresponding transfer matrix is

$$\mathbf{T} = \mathbf{T}_A \mathbf{T}_B ; \mathbf{T}_A = \begin{pmatrix} 1 & z_A^{\text{eff}} \\ u & z_A^{\text{eff}} \end{pmatrix} ; \mathbf{T}_B = \begin{pmatrix} 1 & z_B^{\text{eff}} u \\ 1 & z_B^{\text{eff}} \end{pmatrix} \quad (36)$$

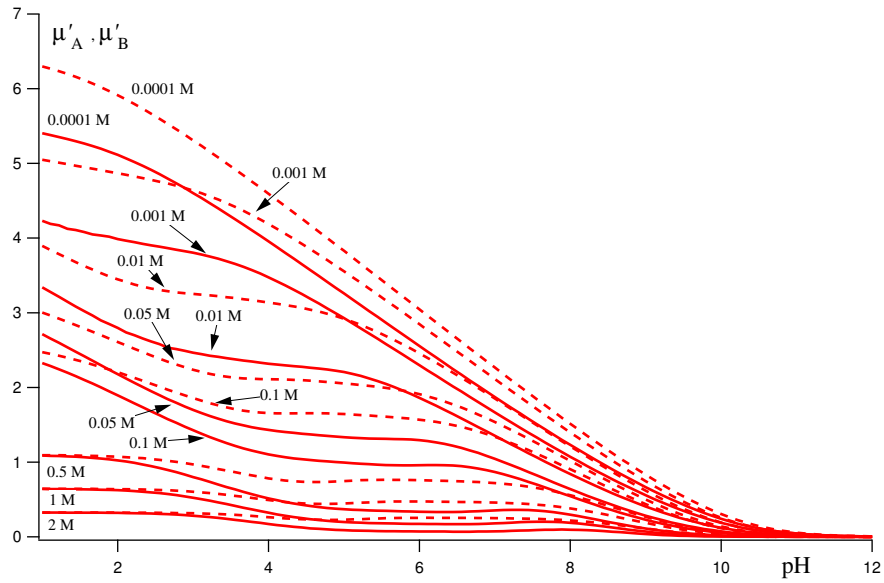
where  $\epsilon = -\log(u)$  represents the interaction between a negatively charged (*i.e.* deprotonated) A group and a positively charged (*i.e.* protonated) B group, so that  $\epsilon < 0$ . Fig 4a compares the theoretical and simulated titration curves, for a polyampholyte with parameters  $pK_A = 4.5$ ,  $pK_B = 6$ ,  $b= 0.2$  nm and  $\epsilon = -1.5$ . Again a very good agreement is obtained for the full range of ionic strengths, a particularly important point when working with polyampholytes since they can become insoluble at high ionic strengths [49]. It is worth to note that the isoelectric point, which for this particular model coincides with  $\theta = 1/2$ , is independent of the ionic strength. Fig 4b presents the average protonation degree of sites A and B,  $\theta_A$  and  $\theta_B$ , together with the average charge per monomer ( $q$ ) and the number of zwitterions ( $D = \langle \sum_i s_i s_{i+1} \rangle$ ) per monomer. The isoelectric point appears around  $\text{pH} \simeq 5$ , which, as expected, coincides with the maximum number of zwitterions. The effective parameters capture both the strong binding correlations due to the short range interactions and the more delocalized effect of long range interactions.

## 4 Higher order corrections

The treatment here presented can be extended to calculate higher order corrections to the titration curves by defining other local effective parameters, such as effective nearest neighbour interactions, effective next-nearest neighbour interactions, effective triplet interactions, etc. The resulting free energies are

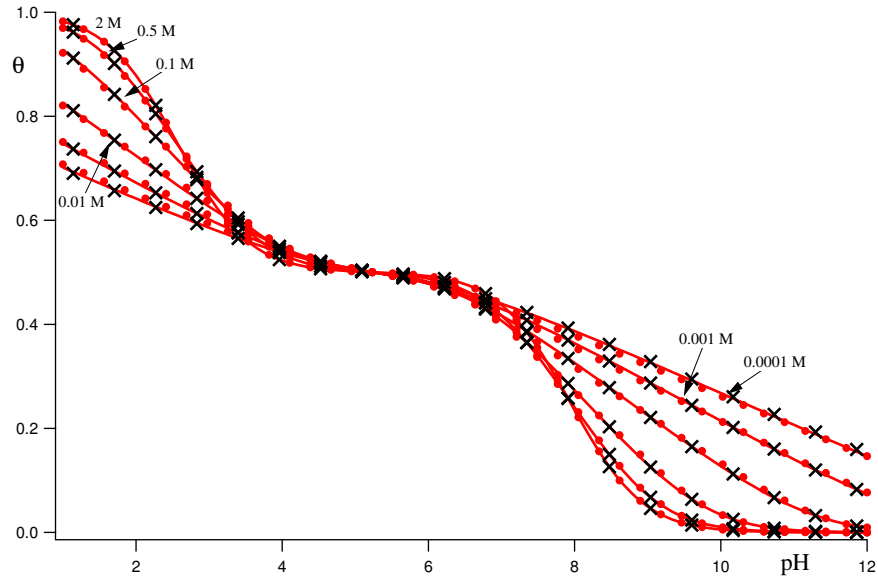


(a) Fig. 3a

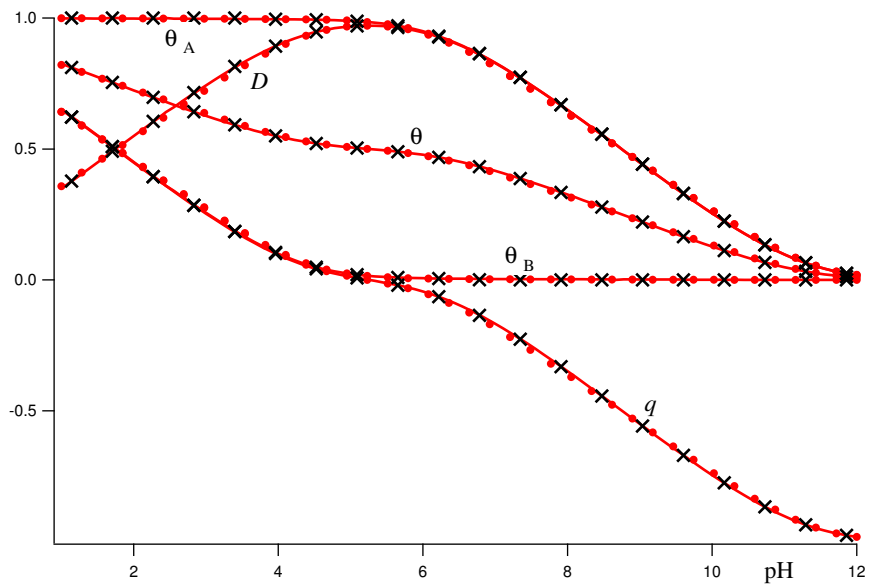


(b) Fig. 3b

Figure 3: (a) Titration curves corresponding to a heterogeneous polyelectrolyte with alternating sites A and B experiencing nearest neighbour interactions. Long range interactions are described by the Debye-Hückel potential. The ionic strengths range from 2M to  $10^{-4}$ M. The chosen parameters are  $pK_A = 7$ ,  $pK_B = 9$ ,  $\epsilon = 1.5$ ,  $b=0.2$  nm, which represent alternating copolymers with two different basic functional groups. Dots: values obtained from MC simulations; continuous lines: obtained by using the first order correction (b) First order correction to the site protonation free energy of sites A and B: continuous lines:  $\mu'_A$ ; dashed lines:  $\mu'_B$ .



(a) Fig. 4a



(b) Fig. 4b

Figure 4: (a) Titration curves corresponding to a polyampholyte with alternating acidic (A) and basic (B) groups with  $b = 0.2$  nm,  $pK_A = 4.5$ ,  $pK_B = 6$  and  $\epsilon = -1.5$  for ionic strengths ranging from 2M to  $10^{-4}$ M. (b) Average degree of protonation ( $\theta$ ), average degree of protonation of the acidic ( $\theta_A$ ) and basic ( $\theta_B$ ) groups, average number of zwitterions per monomer ( $D$ ) and average charge per monomer ( $q$ ), at ionic strength 0.01 M. Continuous lines: first order correction; crossed markers: second order correction; dots: MC simulations.

expected to converge very fast to the exact free energy. For simplicity, let us firstly treat the situation in which the short range interactions are described by the nearest neighbour model. Besides the effective reduced chemical potential, we also search an effective interaction energy. The free energy can now be splitted into

$$\begin{aligned} \frac{H_0}{\ln 10} &= (\mu + \mu'_2) \sum_i s_i + (\epsilon + \epsilon') \sum_i s_i s_{i+1} \\ \frac{\Delta H}{\ln 10} &= \sum_{j>i+m} \phi_{ij} s_i s_j - \mu'_2 \sum_i s_i - \epsilon' \sum_i s_i s_{i+1} \end{aligned} \quad (37)$$

The trial free energy adopts the form

$$\frac{\tilde{\Omega}}{\ln 10} = \frac{\Omega_0(\mu'_2, \epsilon')}{\ln 10} + \varphi(\mu'_2, \epsilon') - \mu'_2 \nu_0 - \epsilon' D_0 \quad (38)$$

where  $D_0 = \langle \sum_i s_i s_{i+1} \rangle_0$  is the number of neighbouring protonated sites over the unperturbed system. Minimizing (38) with respect to  $\mu'$  and  $\epsilon'$  we obtain

$$\mathbf{J}_2 \begin{pmatrix} \mu'_2 \\ \epsilon' \end{pmatrix} = \begin{pmatrix} \left( \frac{\partial \varphi}{\partial \mu'_2} \right)_{\epsilon'} \\ \left( \frac{\partial \varphi}{\partial \epsilon'} \right)_{\mu'_2} \end{pmatrix}; \quad \mathbf{J}_2 = \begin{pmatrix} \left( \frac{\partial \nu_0}{\partial \mu'_2} \right)_{\epsilon'} & \left( \frac{\partial D_0}{\partial \mu'_2} \right)_{\epsilon'} \\ \left( \frac{\partial \nu_0}{\partial \epsilon'} \right)_{\mu'_2} & \left( \frac{\partial D_0}{\partial \epsilon'} \right)_{\mu'_2} \end{pmatrix} \quad (39)$$

where

$$\begin{aligned} \nu_0 &= \partial \Omega_0 / \partial \mu'_2 = \nu_0(\mu'_2, \epsilon') \\ D_0 &= \partial \Omega_0 / \partial \epsilon' = D_0(\mu'_2, \epsilon') \end{aligned} \quad (40)$$

has been used. Note that, as pointed out in subsection 3.2,  $\tilde{\Omega}$  is the Legendre transformation of  $\Omega_0$  corresponding to the change of variables (40).

A physical interpretation of  $\mu'_2$  and  $\epsilon'$  arises using an argument similar to the one used in section 3.2. Eqn. (40) can be seen as the transformation between the two pair of variables  $(\mu'_2, \epsilon')$  and  $(\nu_0, D_0)$  and  $\mathbf{J}_2$  is the jacobian of such a transformation. Inverting (39),  $\mu'_2$  and  $\epsilon'$  remain

$$\begin{pmatrix} \mu'_2 \\ \epsilon' \end{pmatrix} = \mathbf{J}_2^{-1} \begin{pmatrix} \left( \frac{\partial \varphi}{\partial \mu'_2} \right)_{\epsilon'} \\ \left( \frac{\partial \varphi}{\partial \epsilon'} \right)_{\mu'_2} \end{pmatrix} = \begin{pmatrix} \left( \frac{\partial \varphi}{\partial \nu_0} \right)_{D_0} \\ \left( \frac{\partial \varphi}{\partial D_0} \right)_{\nu_0} \end{pmatrix} \quad (41)$$

so that  $\mu'_2$  represents the average increase in the long range energy when a new proton is bound (keeping constant the number of neighbouring interactions), while  $\epsilon'$  accounts for the change in the long range energy in creating a new neighbouring interaction (at a constant number of bound protons). This result explains why the first order correction works so well. Intuitively, one can expect that the change in the long range energy in creating a new neighbouring interaction is very small if no new proton is bound to the molecule. This is confirmed by the calculations, which show that the second order correction does not add significant improvement to the titration curves. In Fig. 1a (homogeneous polyelectrolyte) and Fig. 4a (polyampholyte) the second order correction (39) has been plotted (crossed markers). No significant improvement is obtained compared to the ones calculated by using the first order correction. In fig. 5 the



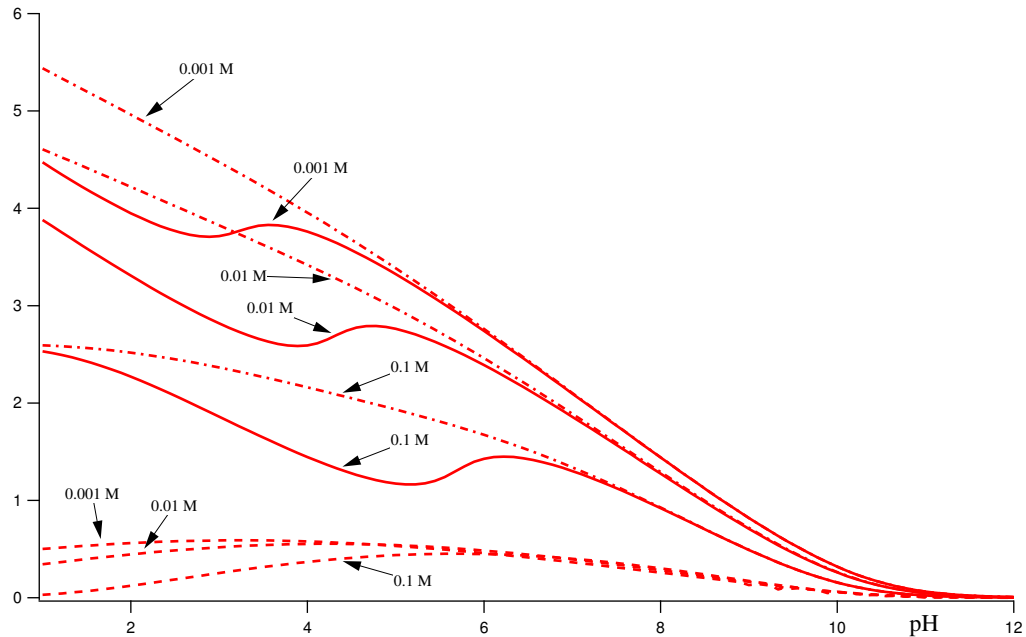


Figure 5: Continuous lines: first order correction to the protonation free energy  $\mu'$ ; dash-dot line: second order correction to the protonation free energy  $\mu'_2$ ; dashed line: second order correction to the nearest neighbour interaction energy  $\epsilon'$ . Ionic strengths: 0.1M,  $10^{-2}$ M,  $10^{-3}$ M. The parameters are same as in Fig. 1

second order corrections  $\mu'_2$  and  $\epsilon'$  are plotted *versus* pH, for the same homogeneous polyelectrolyte depicted in Fig. 1, and compared to  $\mu'$ , the first order correction to the site protonation free energy. As expected,  $\epsilon'$  is small compared to  $\mu'_2$ . We also observe that the wavy behaviour of  $\mu'$  is no longer present in  $\mu'_2$ , but it has been replaced by the contribution to the nearest neighbour interaction,  $\epsilon'$ , which is maximum at pH-values for which  $\theta \simeq 0.5$ . This could be related to the increment in free energy produced when the structure of alternating protonated and deprotonated sites is modified when two protonated sites become neighbours. Using the same procedure, higher order corrections to the triplet, next-nearest neighbour interactions can be calculated, and expressions of the type (41) derived. Similarly, the correction to a particular cluster parameter can be interpreted as the average change in the long range energy when a new interaction (of the type described by the cluster parameter), is created in the polymer. However, as discussed above, no gain in accuracy is obtained. Therefore, at least for the models here studied, the first order correction to the protonation free energy is enough to reproduce almost exactly the MC simulations.

## 5 Conclusions

An approach to Ising-type models in polyelectrolytes which treats short and long range interactions simultaneously, is presented. The effect of long range interactions is included by means of effective short range cluster parameters, which can be systematically calculated by means of the Gibbs-Bogoliubov variational principle. The first order correction to the site protonation free energy correspond to the change of the average long range interaction energy when a new proton is added, a result which is valid for small molecules, polyelectrolytes and surfaces. However, it is specially straightforward to apply to linear molecules, where the transfer matrix method can be implemented for a wide variety of situations. In general this formalism has been used when long range interactions are neglected, since the size of the matrices grows exponentially with the range of the interactions. This fact often restricted the study of the titration curves to high ionic strengths, sometimes a hard limitation since many molecules of interest become insoluble under such conditions. This difficulty is overcome with the treatment here presented.

Different models have been used to test the method. In all the cases investigated, where long range interactions have been included by means of the Debye-Hückel potential, the first order correction provides results which are almost indistinguishable from those obtained from MC simulations, allowing fast parameter fitting to experimental data for the full range of ionic strengths. As examples of heterogeneous systems, polyelectrolytes and polyampholytes with two kind of alternating sites have been chosen to test the method. In the case of polyampholytes, both repulsive and attractive interactions are present. Effective protonation free energies have been calculated for each chemically different site. It is important to emphasize that long range interactions imply important corrections to the titration curves even for not too low ionic strengths, which could lead to a significant bias in the fitted parameters if this type of interactions is not take into account.

Higher order corrections have also been implemented. In the second order correction, both the site protonation free energy and the nearest neighbour interaction are recalculated to include long range effects. It is shown that the correction to the nearest neighbour interaction energy corresponds to the average change in the long range energy when a new neighbouring proton-proton pair is created (keeping constant the number of bound protons). This justifies why the second order correction is so small and add little improvement to the resulting titration curves. As a general statement, we can say that the correction to a particular cluster parameter represents the average change in the long range energy when a new interaction (of the type described by the cluster parameter), is created in the polymer.

The basic ideas here presented could possibly be extended to a wider group of situations where Ising and Potts models are useful, such as ionization of surfaces and small molecules, conformational properties of macromolecules, polymer-polymer and polymer-surface interactions, molecular stretching, among others.

## Acknowledgments

This work was supported by grants from Spanish Ministerio de Economía y Competitividad (CTM2012-39183) and Generalitat de Catalunya (grants 2014SGR1132, 2014SGR1017 and XrQTC).

## APPENDIX: degree of protonation and binding correlation in terms of derivatives of the transfer matrix for $N \rightarrow \infty$

For a linear chain of  $N+1$  identical sites the partition function is given by eqn. 9. In the limit  $N \rightarrow \infty$  the average degree of protonation  $\theta$  is the same for all the binding sites, so that we can choose, for instance, the site located in the middle of the chain.  $\theta$  is thus given by

$$\theta = -\frac{\partial \log \Xi}{\partial \mu} = \frac{\mathbf{q}\mathbf{T}^{N/2}\mathbf{T}'\mathbf{T}^{N/2}\mathbf{p}^{\mathbf{T}}}{\mathbf{q}\mathbf{T}^{N+1}\mathbf{p}^{\mathbf{T}}} \quad (\text{A.1})$$

where  $N$  is even and

$$\mathbf{T}' = -\frac{\mathbf{1}}{\ln 10} \frac{\partial \mathbf{T}}{\partial \mu} \quad (\text{A.2})$$

$\mathbf{T}$  can be expressed in the diagonal form  $\mathbf{T} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}$ , where  $\mathbf{Q}$  is a matrix whose columns are the eigenvectors of  $\mathbf{T}$ , and  $\mathbf{\Lambda}$  is a diagonal matrix. Let us suppose that the maximum eigenvalue is the  $(r, r)$  element of  $\mathbf{\Lambda}$  and the corresponding eigenvector is the  $r$ th column of  $\mathbf{Q}$  and  $\mathbf{\Lambda}$ . Eqn. (A.1) reads

$$\theta = \frac{\partial \ln \Xi}{\partial \mu} = \frac{1}{\lambda} \frac{\mathbf{q}\mathbf{Q}\mathbf{\Lambda}^{N/2}\mathbf{Q}^{-1}\mathbf{T}'\mathbf{q}\mathbf{Q}\mathbf{\Lambda}^{N/2}\mathbf{Q}^{-1}\mathbf{p}^{\mathbf{T}}}{\mathbf{q}\mathbf{Q}\mathbf{\Lambda}^{N+1}\mathbf{Q}^{-1}\mathbf{p}^{\mathbf{T}}} \quad (\text{A.3})$$

Dividing numerator and denominator by  $\lambda^{N+1}$ , where  $\lambda$  is the maximum eigenvalue of  $\mathbf{T}$ , and noting that for  $M \rightarrow \infty$

$$\left(\frac{\mathbf{\Lambda}}{\lambda}\right)^M \rightarrow \mathbf{E} \quad (\text{A.4})$$

where  $\mathbf{E}$  is a matrix whose elements are zero excepting  $E_{rr} = 1$ . Therefore, eqn. (A.3) reads

$$\theta = \frac{\mathbf{q}\mathbf{B}\tilde{\mathbf{T}}\mathbf{B}\mathbf{p}^{\mathbf{T}}}{\mathbf{q}\mathbf{B}\mathbf{p}^{\mathbf{T}}} \quad (\text{A.5})$$

where  $\mathbf{T} = \mathbf{T}'/\lambda$  and  $\mathbf{B} = \mathbf{Q}\mathbf{E}\mathbf{Q}^{-1}$ . After some algebra  $\mathbf{B}$  can be expressed as

$$\mathbf{B} = \mathbf{a}^{\mathbf{T}}\mathbf{b} \quad (\text{A.6})$$

where  $\mathbf{a} = (Q_{1r}, Q_{2r}, Q_{3r}, \dots)$  contains the eigenvector corresponding to  $\lambda$  and  $\mathbf{b}$  is the  $r$ th row of  $\mathbf{Q}^{-1}$ ,  $\mathbf{b} = (Q_{r1}^{-1}, Q_{r2}^{-1}, Q_{r3}^{-1}, \dots)$ . Introducing (A.6) in (A.5)

and taking into account that  $\mathbf{qa}^T$  and  $\mathbf{bp}^T$  are scalar quantities, the sought result (11) is obtained

$$\theta = \frac{1}{\lambda} \frac{(\mathbf{qa}^T) \mathbf{b} \tilde{\mathbf{T}} \mathbf{a}^T (\mathbf{bp}^T)}{(\mathbf{qa}^T) (\mathbf{bp}^T)} = \mathbf{b} \tilde{\mathbf{T}} \mathbf{a}^T \quad (\text{A.7})$$

An expression for the correlation function  $h_{ij} = \langle s_i s_j \rangle$  can be derived in a similar way. For  $N \rightarrow \infty$ , the correlation function  $h_{ij}$  becomes independent of  $i$ , so that  $h_{ij}$  can be expressed

$$h_{ii+k-1} = h_{1k} = \frac{1}{(\ln 10)^2} \frac{\partial \ln \Xi}{\partial \mu_i \partial \mu_{i+k-1}} = \frac{\mathbf{q} \mathbf{T}^{N/2} \mathbf{T}' \mathbf{T}^{k-2} \mathbf{T}' \mathbf{T}^{N/2-k+1} \mathbf{p}^T}{\mathbf{q} \mathbf{T}^{N+1} \mathbf{p}^T} \quad (\text{A.8})$$

Proceeding as in the steps (A.3-A.7), eqn. (A.8) leads to eqn. (12).

## References

- [1] Warshel, A.; Sharma, P. K.; Kato, M.; Parson, W. W. *BBA-Proteins Proteomics* **2006**, *1764*, 1647-1676
- [2] Lipfert, J.; Doniach, S.; Das, R.; Herschlag, D. *Annu. Rev. Biochem.* **2014**, *83*, 813-841
- [3] Rivas, B. L.; Pereira, E. D.; Moreno-Villoslada, I. *Prog. Polym. Sci.* **2003**, *28*, 173-208
- [4] Hartig, S. M.; Greene, R. R.; Dikov, M. M.; Prokop, A.; Davidson, J. M. *Pharm. Res.* **2007**, *24*, 2353-2369
- [5] Chapel, J. P.; Berret, J. F. *Curr. Opin. Colloid Interface Sci.* **2012**, *17*, 97-105
- [6] Stumm, W.; Morgan, J. J. *Aquatic chemistry : chemical equilibria and rates in natural waters*; 3rd ed.; Wiley: New York, 1996.
- [7] Marcus, R. A. *J. Phys. Chem.* **1954**, *58*, 621-623
- [8] Steiner, R. F. *J. Chem. Phys.* **1954**, *22*, 1458-1460
- [9] Katchalsky, A.; Mazur J.; Spitnik, P. *J. Polym. Sci.* **1957**, *23*, 513-532
- [10] Borkovec, M.; Koper, G. J. M.; Piguet, C. *Curr. Opin. Colloid Interface Sci.* **2006**, *11*, 280-289
- [11] Borkovec, M.; Koper, G. J. M.; Spiess, B. *New J. Chem.* **2014**, *38*, 5679-5685
- [12] Baxter, R. J. *Exactly solved models in statistical mechanics*; Academic Press: London, 1982

- [13] Borkovec, M.; Koper, G. J. M. *J. Phys. Chem.* **1994**, *98*, 6038-6045
- [14] Flory, P. L. *Statistical Mechanics of Chain Molecules*; John Wiley: New York, **1969**
- [15] Flory, P. L. *Macromolecules* **1974**, *7*, 381-392
- [16] Garcés, J. L. ; Koper G. J. M.; Borkovec M. *J. Phys. Chem. B* **2006**, *110*, 10937-10950
- [17] Koper, G.J.M.; Borkovec, M, *J Phys. Chem. B* **2001**, *105*, 6666-6674
- [18] Reed, C. E.; Reed, W. F. *J. Chem. Phys.* **1992**, *96*, 1609-1620
- [19] Borkovec, M.; Jönsson, B.; Koper, G. J. M. J. In *Surface and Colloid Science*; Matijevic, E., Ed.; Plenum Press: New York, **2001**; Vol. 16
- [20] Aleksandrov, A.; Polydorides S.; Archontis G.; Simonson T. *J. Phys. Chem. B* **2012**, *114*, 10634-10648
- [21] Garcés J. L. ; Madurga S.; Borkovec, M. *Phys. Chem. Chem. Phys.*, **2014**, *16*, 4626-4638
- [22] Borkovec, M.; Daicic, J.; Koper, G. J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 3499-3503
- [23] Burak, Y.; Netz R. R. *J. Phys. Chem.* **2004**, *108*, 4840-4849
- [24] Liem, R.; Poland, D.; Scheraga, H. *J. Am. Chem. Soc.* **1970**, *92*, 5717-5724
- [25] Vila, J. *J. Chem Phys.* **1986**, *84*, 6421-6425
- [26] Borkovec, M.; Koper, G. J. M. *Anal. Chem.* **2000**, *72*, 3272-3279
- [27] Felemez, M.; Bernard, P.; Schlewer, G.; Spiess, B. *J. Am. Chem. Soc.* **2000**, *122*, 3156-3165
- [28] Koper, G. J. M.; van Duijvenbode, R. C.; Stam, D. D. P. W.; Steuerle, U.; Borkovec, M. *Macromolecules* **2003**, *36*, 2500-2507
- [29] Cakara, D. ; Borkovec M. *Croat. Chem. Acta* **2007**, *80*, 421-428
- [30] Koper, G. J. M.; Borkovec, M. *Polymer* **2010**, *51*, 5649-5662
- [31] Borkovec M. ; Cakara D. ; Koper G. J. M. *J. Phys. Chem. B* ; **2012**, *116*, 4300-4309
- [32] Cannas, S. A. *Phys. Rev. B*, **1995**, *52*, 3034-3037
- [33] Ullner, M.; Jönsson, B. *Macromolecules* **1996**, *29*, 6645-6655
- [34] Nishio, T. *Biophys. Chem.* **1998**, *71*, 173-184

- [35] Zarzycki, P. *Langmuir*, **2006**, *22*, 11234-11240
- [36] Zarzycki, P. *J. Coll. Interf. Sci.* **2007**, *311*, 622-627
- [37] Ulrich, S.; Seijo, M.; Stoll, S. *J. Phys. Chem. B*, **2007**, *111*, 8459-8467
- [38] Companys, E.; Garcés, J.L.; Salvador, J.; Galceran, J.; Puy, J.; Mas, F. *Coll. Surf. A: Physchem. Eng.*, **2007**, *306*, 2-13
- [39] Madurga, S.; Garcés, J.L.; Companys, E.; Rey-Castro, C.; Salvador, J.; Galceran, J.; Puy, J.; Mas, F. *Theor. Chem. Acc.*, **2009**, *123*, 127-135
- [40] Garcés, J. L.; Rey-Castro, C.; David, C.; Madurga, S.; Mas, F.; Pastor, I.; Puy, J. *J. Phys. Chem. B* **2009**, *113*, 15145-15155
- [41] Carnal, F.; Ulrich S.; Stoll, S. *Macromolecules*, **2010**, *43*, 2544-2553
- [42] Carnal, F.; Stoll, S. *J. Chem. Phys.*, **2011**, *134*, 044904:1-11
- [43] Chandler, D. *Introduction to Modern Statistical Mechanics*: Oxford University Press: New York, **1987**
- [44] Koper, G. J. M.; Borkovec, M. *Langmuir* **1994**, *10*, 2863-2865
- [45] Koper, G. J. M.; Borkovec, M. *J. Chem. Phys.* **1996**, *104*, 4204
- [46] Teif, V.B. *Nucleic Acids Res.* **2007**, *35*, 3519-3867
- [47] Lee, A. A.; Kondrat, S.; Kornyshev, A. A. *Phys. Rev. Lett.* **2014**, *113*, 048701
- [48] Poland, D.; Scheraga, H. A. *Theory of Helix-Coil Transitions in Biopolymers*; Academic Press: New York, **1970**
- [49] Alfrey Jr., T. ; Morawetz, H. *J. Am. Chem. Soc.* **1952**, *74*, 436-438
- [50] Mazur J. ; Silberberg A. ; Katchalsky, A. *J. Polym. Sci.* **1959**, *35*, 43-70
- [51] David, C.; Companys, E.; Galceran, J.; Garcés, J. L.; Mas, F.; Rey-Castro, C.; Salvador, J.; Puy, J. *J. Phys. Chem. B* **2007**, *111*, 10421-10430
- [52] David, C.; Companys, E.; Galceran, J.; Garcés, J. L.; Mas, F.; Rey-Castro, C.; Salvador, J.; Puy, J. *J. Phys. Chem. B* **2008**, *112*, 10092-10100
- [53] Stosic, B. ; Milosevic, S. ; *Phys. Rev. B*, **1990**, *41*, 11466-11478
- [54] Smits, R. G.; Koper, G. J. M.; Mandel, M. *J. Phys. Chem.* **1993**, *97*, 5745-5751

- [55] Ziebarth, J. D.; Wang, Y. *Biomacromolecules*, **2010**, *11*, 29-38
- [56] Brilliantov, N. V.; Kuznetsov, D. V.; Klein, R. *Phys. Rev. Lett.*, **1998**, *81*, 1433-1436
- [57] Schiessel, H.; Pincus, P. *Macromolecules* **1998**, *31*, 7953-7959
- [58] Winkler, R. G.; Gold, M.; Reinecker, P. *Phys. Rev. Lett.* **1998**, *17*, 3731-3734
- [59] Tom, A. M.; Vemparala, S.; Rajesh, R.; Brilliantov, N. V. *Phys. Rev. Lett.*, **2016**, *117*, 147801(5)
- [60] Brilliantov, N. *Physical Rev. E*, **1993**, *48*, 4536-4548
- [61] Linse, O.; Lobaskin, V. *Phys. Rev. Lett.* **1999**, *83*, 4208(4)
- [62] Diehl, A.; Barbosa, M.C.; Levin, Y. *Europhys. Lett.* **2001**, *53*, 86-92
- [63] Zoetekouw B.; van Roij, R. *Phys. Rev. Lett.*, **2006**, *97*, 258302(4)