ARTICLE IN PRESS

Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications ▮ (▮▮▮▮) ▮▮▮–▮▮▮

stochastic processes and their applications

www.elsevier.com/locate/spa

Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes

Giambattista Giacomin*, Maha Khatib

Université Paris Diderot, Sorbonne Paris Cité, Laboratoire de Probabilités et Modèles Aléatoires, UMR 7599, F-75205 Paris, France

Received 22 November 2015; received in revised form 20 June 2016; accepted 20 June 2016

Abstract

The Poland–Scheraga model describes the denaturation transition of two complementary – in particular, equally long – strands of DNA, and it has enjoyed a remarkable success both for quantitative modeling purposes and at a more theoretical level. The solvable character of the homogeneous version of the model is one of features to which its success is due. In the bio-physical literature a generalization of the model, allowing different length and non complementarity of the strands, has been considered and the solvable character extends to this substantial generalization. We present a mathematical analysis of the homogeneous generalized Poland–Scheraga model. Our approach is based on the fact that such a model is a homogeneous pinning model based on a bivariate renewal process, much like the basic Poland–Scheraga model is a pinning model based on a univariate, i.e. standard, renewal. We present a complete analysis of the free energy singularities, which include the localization–delocalization critical point and (in general) other critical points that have been only partially captured in the physical literature. We obtain also precise estimates on the path properties of the model.

MSC: 60K35; 82D60; 92C05; 60K05; 60F10

Keywords: DNA denaturation; Polymer pinning model; Two-dimensional renewal processes; Critical behavior; Sharp and large deviation estimates; Path properties

* Corresponding author.
E-mail address: giambattista.giacomin@univ-paris-diderot.fr (G. Giacomin).

http://dx.doi.org/10.1016/j.spa.2016.06.017
0304-4149/

1. Introduction and main results

1.1. General overview

The localization–delocalization phenomenon in various polymer models has been the object of much attention in the physics, biophysics and mathematics literature [13,16,10,21,26,27]. One of the main biological and physical phenomenon that motivates this work is DNA denaturation, that is the separation of the two DNA strands at high temperature and, more generally, the fluctuation phenomena observed at lower temperatures, when the two strands are tied together. The most basic and studied model in this field is the Poland–Scheraga (PS) model [26] which is limited to the case of sharp complementarity of two equal length strands: only bases with the same index can form pairs. From the theoretical physics and mathematical viewpoint what is most remarkable in the homogeneous version of the model is its solvable character and the fact that at the delocalization (or denaturation) transition the behavior – i.e. the critical behavior – can be fully captured. A mathematical viewpoint on this solvable character and on the solution itself is that the Poland–Scheraga model is a Gibbs measure with only one body potentials and built on a one-dimensional process [16].

In [15,14,24] (see also [23,28] for a primitive version), a generalization of the Poland–Scheraga (gPS) model has been introduced and the novelties are:

- The possibility of formation of non-symmetrical loops in the two strands (i.e., the contribution to a loop, in terms of number of bases, from the two strands is not necessarily the same).
- The two strands may be of different lengths.

These novelties are very substantial (we invite the reader to compare Fig. 1 with [13, Fig. 6] or [18, Fig. 2.5]). Nevertheless, as already pointed out in [14,24,28], the solvable character is preserved. However, that the novelties are really substantial is witnessed by a richer phenomenology (partially captured and understood in [12,24]): in addition to the expected denaturation transition, the gPS model undergoes other transitions.

Here we develop a mathematical analysis of the gPS model based on the observation that it is a pinning model based on a two-dimensional renewal process. Much like for the original PS model, tools from Renewal Theory allow going far toward a complete understanding of the model. Nevertheless, as we will explain, some important questions are still open and they correspond to open problems in the theory of two and higher dimensional renewal processes.

1.2. The gPS model: biophysics version

This subsection, as well as Section 1.4, can be skipped if one is not focusing on the biophysics set-up. The model we consider has been introduced in [14]. The two DNA strands, of lengths M and $N \geq 1$ – the length of course corresponds to the number of bases – interact by forming some base pairs. We talk of N-strand, M-strand and of base i of the N- or M-strand with the obvious meaning. An allowed configuration of our system is a collection of base pairs

$$((i_1, j_1), (i_2, j_2), \ldots, (i_n, j_n)) \in \mathbb{N}^{2n}, \quad \text{with } \mathbb{N} = \{1, 2, \ldots\}, \tag{1.1}$$

where $n \in \mathbb{N} := \{1, 2, \ldots, N\}$ and

(1) $(i_1, j_1) = (1, 1)$;
(2) $i_k < i_{k'}$, as well as $j_k < j_{k'}$, for $1 \leq k < k' \leq n$.

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

The first condition is simply saying that the first two bases form a pair and the second condition is imposed by the geometric constraint (see Fig. 1). The weight of every configuration is assigned by the following rules:

(1) Each base pair is energetically favored and carries an energy $-E_b < 0$;
(2) A base which is not in a pair is either in a loop or in the free ends:
- It is in a loop if it is in $L_k := ((i_k, i_{k+1}) \cup (j_k, j_{k+1})) \cap \mathbb{N}$ for some $k \in \{1, \ldots, n-1\}$: the loop L_k has length $\ell_k := (i_{k+1} - i_k) + (j_{k+1} - j_k) - 2$ and we associate to L_k an entropy factor $B(\ell_k)$ with
$$B(\ell) := s^\ell \ell^{-c}, \tag{1.2}$$
where $s \geq 1$ and $c > 2$. There is also an energetic penalty $E_l > 0$ penalty associated to a loop.
- The free ends have length $N - i_n$ and $M - j_n$ and to each free end we associate the entropy term $A(\ell) := s^\ell (\ell + 1)^{-\bar{c}}$ where $\bar{c}$ is another positive constant.

As we will see the value of s is irrelevant. The value of $\bar{c}$, chosen equal to 0.1 in [24], is somewhat more relevant, but what is very relevant is the value of c: in [24] it is chosen equal to 2.15.

These rules easily lead to a formula for the partition function, i.e. the sum of the weights over all allowed configurations, of our system
$$Z_N^M := \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} A(i) A(j) W_{N-i}^{M-j}, \tag{1.3}$$
where W_l^r obeys the recursion relation ($\beta \geq 0$ is proportional to the inverse of the temperature)
$$W_{m+1}^{r+1} = \exp(\beta E_b) W_m^r + \exp(\beta(E_b - E_l)) \sum_{\substack{i,i' : i+i' > 0 \\ i < m,\, i' < r}} B(i + i') W_{m-i}^{r-i'}, \tag{1.4}$$
with $W_1^1 = 1$ and $W_1^i = W_i^1 = 0$ for $i > 1$.

1.3. The gPS model: renewal process viewpoint

From a mathematical perspective we take a more general viewpoint and we introduce a two-dimensional renewal pinning model (see Fig. 2). A discrete two-dimensional renewal issued from the origin is a random walk $\tau = \{\tau_n\}_{n=0,1,\ldots} = (\tau^{(1)}, \tau^{(2)}) = \{(\tau_n^{(1)}, \tau_n^{(2)})\}_{n=0,1,\ldots}$ where $\tau_0 = (0, 0)$ and, for $n \in \mathbb{N} := \{1, 2, \ldots\}$, τ_n is a sum of n independent identically distributed random variables taking values in $\mathbb{N}^2$. So if we set $\mathrm{K}(n, m) := \mathbf{P}(\tau_1 = (n, m))$ then given $\{(i_n, j_n)\}_{n=0,1,2,\ldots}$, with $(i_0, j_0) = (0, 0)$, for every $k \in \mathbb{N}$
$$\mathbf{P}\left(\tau_n = (i_n, j_n) \text{ for } n = 1, 2, \ldots, k\right) = \prod_{n=1}^{k} \mathrm{K}\left(i_n - i_{n-1}, j_n - j_{n-1}\right), \tag{1.5}$$
and, by construction, such a probability is zero unless the i's and j's form strictly increasing sequences.

We can then introduce for given N and $M \in \mathbb{N}$ a pinning model of length (N, M) by forcing, i.e. conditioning, τ to visit (N, M) and by penalizing ($h \leq 0$) or rewarding ($h \geq 0$) the number of

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

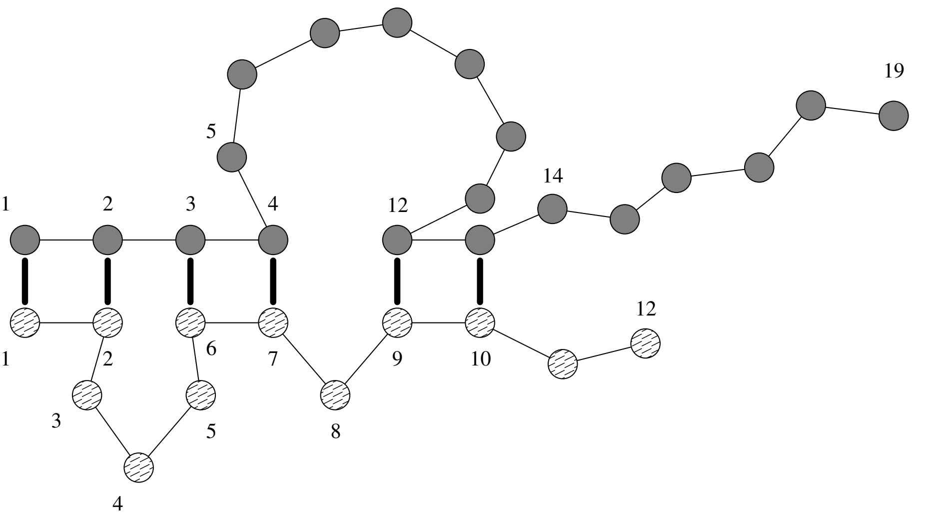

Fig. 1. A representation of a trajectory of the gPS model in the biophysics representation. The first strand contains 12 bases, the second strand 19. The six base pairs determining the configuration are (1, 1), (2, 2), (6, 3), (7, 4), (9, 12) and (10, 13).

renewals up to (N, M). More formally, we introduce the probability measure $\mathbf{P}_{N,M,h}$ by setting for every $k \in \mathbb{N}$ such that $k \leq \min(N, M) =: N \wedge M$ and for every $\{(i_n, j_n)\}_{n=0,1,\ldots,k}$ with $(i_0, j_0) = (0, 0)$, $i_n - i_{n-1} > 0$ as well as $j_n - j_{n-1} > 0$ for $n = 1, \ldots, k$ and $(i_k, j_k) = (N, M)$

$$\frac{\mathbf{P}_{N,M,h}\left(\tau_n = (i_n, j_n) \text{ for } n = 1, 2, \ldots, k\right)}{\mathbf{P}\left(\tau_n = (i_n, j_n) \text{ for } n = 1, 2, \ldots, k\right)} := \frac{1}{\mathcal{Z}_{N,M,h}} \exp(hk), \tag{1.6}$$

where $\mathcal{Z}_{N,M,h}$ is the partition function (or normalization constant):

$$\mathcal{Z}_{N,M,h} := \mathbf{E}\left[\exp\left(h|\tau \cap ([1, N] \times [1, M])|\right) \mathbf{1}_{(N,M)\in\tau}\right], \tag{1.7}$$

in which we are interpreting τ as a random subset of $\mathbb{N}^2$ and $|\cdot|$ denotes the cardinality. Note that $\mathcal{Z}_{0,0,h} = 1$, as well as $\mathcal{Z}_{0,M,h} = \mathcal{Z}_{N,0,h} = 0$. Of course $\mathbf{P}_{N,M,h}$ requires $\mathcal{Z}_{N,M,h} \neq 0$ and whether this is the case or not depends on the inter-arrival distribution $\mathrm{K}(n, m)$ and, possibly, on N and M.

We have used the atypical notation $\mathcal{Z}$ instead of Z because the latter is going to be employed for the model on which we really focus: we consider in fact a very special choice of the *inter-arrival distribution* $\mathrm{K}(\cdot, \cdot)$, namely $\mathrm{K}(n, m) = K(n + m)$ where $K : \{2, 3, \ldots\} \to (0, \infty)$ and

$$K(n) := \frac{L(n)}{n^{2+\alpha}}, \tag{1.8}$$

where $L(\cdot)$ is a slowly varying function and $\alpha \geq 0$ (see Appendix for the properties of slowly varying functions). With this definition, we have that $K(n) > 0$ for every n: all statements generalize to the case in which $K(n) = 0$ for finitely many n, but we make this choice aiming at conciseness of some proofs. We require $\sum_{n,m\in\mathbb{N}} K(n + m) = 1$ and of course

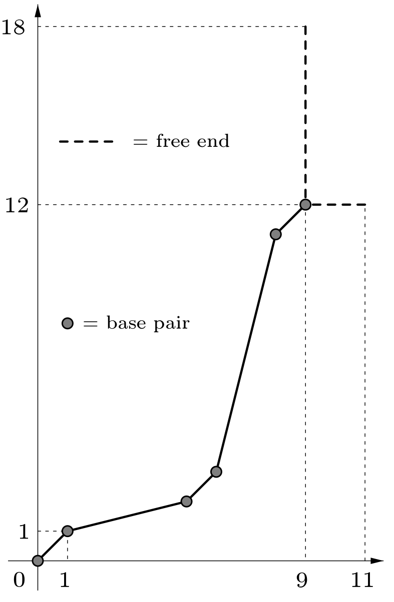

Fig. 2. A representation of a trajectory of the gPS model in the renewal process representation, corresponding to Fig. 1. The base pairs are now renewal points of a two-dimensional discrete renewal process: these points correspond to the (six) base pairs of Fig. 1, except that the first base pair is now (0, 0) and also all the other ones are translated down of (1, 1) with respect to Fig. 1. The renewal trajectory is drawn up to the renewal point (9, 12) and the trajectory up to this point corresponds to one of the possible trajectories of $Z^c_{9,12,h}$. The free ends, of lengths 2 and 6, are then represented as straight lines that go till the boundary of the rectangle with opposite vertices (0, 0) and (11, 18). $Z^f_{11,18,h}$ is obtained by summing up with respect to the position of the last renewal point – (9, 12) in this example – the contribution of the constrained partition function times the contribution due to the two free ends.

$\sum_{n,m\in\mathbb{N}} K(n+m) = \sum_{m=1}^{\infty} mK(m+1)$. We introduce then the *constrained partition function* $Z^c_{N,M,h}$ which coincides with $\mathcal{Z}_{N,M,h}$ once the specific choice of the inter-arrival is made. In an alternative explicit fashion

$$Z^c_{N,M,h} := \sum_{n=1}^{N\wedge M} \sum_{\substack{\underline{l}\in\mathbb{N}^n:\\ |\underline{l}|=N}} \sum_{\substack{\underline{t}\in\mathbb{N}^n:\\ |\underline{t}|=M}} \prod_{i=1}^{n} \exp(h) K\left(l_i + t_i\right), \tag{1.9}$$

where $|\underline{l}| = \sum_{i=1}^{n} l_i$. The *free partition function* is then defined by

$$Z^f_{N,M,h} := \sum_{i=0}^{N} \sum_{j=0}^{M} K_f(i) K_f(j) Z^c_{N-i,M-j,h}, \tag{1.10}$$

where $K_f : \{0\} \cup \mathbb{N} \to (0, \infty)$ is defined as $K_f(n) := \overline{L}(n)/n^{\overline{\alpha}}$ for every $n \geq 1$ and $K_f(0) = 1$ (an arbitrary choice: there is no loss of generality with respect to requiring just $K_f(0) > 0$ and, once again, one can even allow $K_f(n) = 0$ for finitely many n, but we choose positivity for conciseness) with $\overline{\alpha} \in \mathbb{R}$. The free partition function is the normalization associated to the probability $\mathbf{P}^f_{N,M,h}$ defined by setting for every $k \in \mathbb{N}$ such that $k \leq \min(N, M)$ and for every $\{(i_n, j_n)\}_{n=0,1,\dots,k}$ with $(i_0, j_0) = (0, 0)$, $i_n - i_{n-1} > 0$ as well as $j_n - j_{n-1} > 0$ for $n = 1, \dots, k$

and i_k (respectively, j_k) that does not exceed N (respectively, M)

$$\begin{aligned}&\mathbf{P}^f_{N,M,h}\left(\tau \cap [1,N]\times[1,M] = \{(i_1,j_1),(i_2,j_2),\ldots,(i_k,j_k)\}\right)\\&:= \frac{\exp(hk)}{Z^f_{N,M,h}} K_f(N-i_k)K_f(M-j_k)\mathbf{P}\left(\tau_n=(i_n,j_n)\quad \text{for } n=1,2,\ldots,k\right).\end{aligned} \tag{1.11}$$

Here we introduce also the free energy density (the existence of the limit is proven in Section 4)

$$\mathrm{F}_\gamma(h) := \lim_{\substack{N,M\to\infty:\\ \frac{M}{N}\to\gamma}} \frac{1}{N}\log Z^f_{N,M,h}, \tag{1.12}$$

for $\gamma > 0$ or, as we will often do without loss of generality, for $\gamma \geq 1$. The limit in (1.12) means: for every $\{(N_n, M_n)\}_{n=1,2,\ldots}$ with $\lim_n M_n/N_n = \gamma$ and $\lim_n N_n = \infty$.

1.4. Matching biophysics and renewal process viewpoints

The first remark to make on the biophysics model is that the dependence on s of Z^M_N (cf. (1.3)) is trivial: Z^M_N for a given value of s coincides with $s^{N+M}\widetilde{Z}^M_N$, where $\widetilde{Z}^M_N$ is Z^M_N with $s = 1$ and both E_b and E_l are decreased by $\frac{2}{\beta}\log s$. We can therefore set $s = 1$ without true loss of generality. We then remark that we can match $\exp(-\beta E_b)Z^{M+1}_{N+1}$ and $Z^f_{N,M,h}$ by an appropriate choice of h, $K(\cdot)$ and $K_f(\cdot)$: the fact that we consider the biophysics model with lengths augmented by 1 and renormalized by the factor $\exp(-\beta E_b)$ is just because in parallel with one-dimensional renewal pinning works we have chosen to start the renewal from *time* (or *renewal epoch*) zero, but without giving an energetic reward to the base pair 0. Matching is then made by observing that:

(1) The match for the free end terms $K_f(\cdot)$ and $A(\cdot)$ is easily made.
(2) A *minimal* inter-arrival step, that is $(1, 1)$, in the renewal model corresponds to a base pair and contributes $K(2)\exp(h)$. It would then be matched to $\exp(\beta E_b)$ in the biophysics model.
(3) All other inter-arrival steps (i, i'), i.e. $i + i' = 3, 4, \ldots$, which give a contribution $K(i + i')\exp(h)$ correspond to loops with $i + i' - 2$ unpaired bases and the contribution in the biophysics model is $\exp(\beta(E_b - E_l))B(i + i' - 2)$.

Then of course $c = 2 + \alpha$ and $\bar{c} = \bar{\alpha}$ and, at last, the matching between the two models is done thanks to the ample freedom that we have in choosing $K(\cdot)$, except that we have required that $\sum_{n,m} K(n + m) = 1$. This *probability constraint* in reality just corresponds to a shift in the parameter h. In a nutshell, renewal models include the biophysics ones: from a qualitative view point the matching is immediate, form the quantitative one it requires some bookkeeping care. Explicit matchings are presented in Section 3.4.

1.5. The free energy and the localization transition

We start from:

Proposition 1.1. *Recall (1.12). We have*

$$\lim_{\substack{N,M\to\infty:\\ \frac{M}{N}\to\gamma}} \frac{1}{N}\log Z^c_{N,M,h} = \lim_{\substack{N,M\to\infty:\\ \frac{M}{N}\to\gamma}} \frac{1}{N}\log Z^f_{N,M,h}. \tag{1.13}$$

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

In practice, $Z^c_{N,M,h}$ is a more fundamental quantity for our computations and we will first identify the free energy density by looking at the exponential growth of this quantity and only after, in Section 4, we will match it with the exponential growth rate in the free case.

Note that we are just speaking of exponential *growth* and not of *decrease*. In fact $\mathrm{F}_\gamma(h) \geq 0$ simply because $Z^f_{N,M,h} \geq K_f(0)^2 Z^c_{N,M,h} = Z^c_{N,M,h} \geq \exp(h)K(N+M)$. This is a very important issue because it is natural to set

$$h_c := \inf\{h : \mathrm{F}_\gamma(h) > 0\} = \max\{h : \mathrm{F}_\gamma(h) = 0\}, \tag{1.14}$$

where the equality on the right comes from the fact that $\mathrm{F}_\gamma(\cdot)$ is locally bounded, convex (hence continuous) and non-decreasing. These facts are evident from the definitions, like the following two preliminary observations:

(1) $Z^c_{N,M,h} \leq 1$ for $h \leq 0$: hence $h_c \geq 0$;
(2) we will see just below that $h_c = 0$, but it is worth pointing out that $h_c < \infty$ by elementary arguments. For example: $h_c \leq -\log K(2)$ because $Z^f_{N,M,h} \geq K_f(0)K_f(M-N)Z^c_{N,N,h}$ and $Z^c_{N,N,h} \geq (\exp(h)K(2))^N$.

From (1.14) we readily see that h_c is a non analyticity point of $\mathrm{F}_\gamma(\cdot)$ and there is a phase transition of the system. By standard arguments based on convexity (see for example [16, Section 1.2] for the univariate pinning model) one realizes that this transition is the denaturation, or localization/delocalization, transition: $\partial_h \mathrm{F}_\gamma(h)$ – in case to be interpreted as, say, left derivative, but we will soon see that $\partial_h \mathrm{F}_\gamma(h)$ exists except, in some cases, at $h = h_c$ – is the density of base pairs (or contact fraction), which is therefore positive, respectively zero, for $h > h_c$, respectively $h < h_c$.

The next result is much more quantitative about this transition: since $Z^f_{N,M,h} \geq K_f(0)K_f(M-N)Z^c_{N,N,h}$, we get that $\mathrm{F}_\gamma(h) \geq \mathrm{F}_1(h)$ for every $\gamma \geq 1$. All asymptotic statements in the next theorem are for $h \searrow 0$:

Theorem 1.2. *For every $\alpha \geq 0$ and $\gamma \geq 1$ we have $h_c = 0$ and there exists $h_{c,\gamma} \in (0, \infty]$ such that $\mathrm{F}_\gamma(\cdot)$ is real analytic in $(-\infty, 0) \cup (0, h_{c,\gamma})$ and $h_{c,\gamma}$ is a non analyticity point (when $h_{c,\gamma} < \infty$). Moreover if $\sum_n n^2 K(n) < \infty$, a condition implied by $\alpha > 1$, we have as $h \searrow 0$*

$$\mathrm{F}_\gamma(h) \sim \mathrm{F}_1(h) \sim c\, h, \tag{1.15}$$

with $c^{-1} := \frac{1}{2}\sum_{n=2}^\infty n(n-1)K(n)$. If instead $\sum_n n^2 K(n) = \infty$, implied by $\alpha \in [0, 1)$, there exists $c_{\alpha,\gamma} \geq 1$ such that

$$\mathrm{F}_\gamma(h) \sim c_{\alpha,\gamma}\mathrm{F}_1(h) \quad \textit{and} \quad \mathrm{F}_1(h) \sim L_\alpha(h)h^{1/\alpha} \tag{1.16}$$

where $L_\alpha(\cdot)$ is slowly varying at 0. In the case $\alpha = 0$, (1.16) should be interpreted as $\mathrm{F}_\gamma(h) = O(h^{1/\varepsilon})$ for every $\varepsilon > 0$.

In Section 3.3 $c_{\alpha,\gamma}$ and $L_\alpha(\cdot)$ are determined. The expression of $c_{\alpha,\gamma}$ implicitly contains nontrivial information on the system, see Proposition 3.7.

We will see that it may be that $h_{c,\gamma} = \infty$, for example $h_{c,1} = \infty$ in full generality, but when $h_{c,\gamma} < \infty$, $h_{c,\gamma}$ may not be the only critical point inside the localized regime (Theorem 1.5). In any case, $h_{c,\gamma} < \infty$ means that there is more than one localized phase in the system: this is what we treat next, but we need to introduce more concepts and definitions. By doing so we will start outlining the proof of Theorem 1.2.

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

1.6. Transitions in the localized regime

A crucial elementary observation is that (1.9) can be written as

$$Z^c_{N,M,h} = \exp\left((N+M)\mathrm{G}\right) \sum_{n=1}^{N\wedge M} \sum_{\substack{l\in\mathbb{N}^n:\\ |l|=N}} \sum_{\substack{t\in\mathbb{N}^n:\\ |t|=M}} \prod_{i=1}^{n} \widetilde{K}_h\left(l_i, t_i\right)$$

$$= \exp\left((N+M)\mathrm{G}\right) \mathbf{P}\left((N,M)\in\widetilde{\tau}_h\right), \tag{1.17}$$

where

- $\widetilde{K}_h(n,m) = \exp(h-(n+m)\mathrm{G})K(n+m)$ – note that $\widetilde{K}_h(n,m)$ is just function of $n+m$ – and $\mathrm{G} = \mathrm{G}(h)$ is the only solution to
$$\sum_{n,m} K(n+m)\exp(h-(n+m)\mathrm{G}) = 1, \tag{1.18}$$
when such a solution exists (that is, when $h \geq 0$), and $\mathrm{G} = 0$ otherwise. We have therefore defined a function $\mathrm{G}: \mathbb{R} \to [0,\infty)$.
- $\widetilde{\tau}_h$ is the two-dimensional renewal issued from $(0,0)$ with inter-arrival distribution $\widetilde{K}_h$: if $h < 0$ then $\widetilde{K}_h : \mathbb{N}^2 \to [0,1)$ is a sub-probability that we make a probability by defining $\widetilde{K}_h(\infty) := 1 - \sum_{(n,m)\in\mathbb{N}^2} \widetilde{K}_h(n,m)$ and, in this case, $\widetilde{\tau}_h$ contains a.s. a finite number of points (and $\{\infty\}$) and we refer to it as a *terminating* renewal.

$\mathrm{G}(h)$ accounts for part of the free energy of the system and some basic features are straightforward:

Lemma 1.3. *The function* $\mathrm{G}(\cdot)$ *is convex and real analytic except at* $h = 0$.

Proof. First of all the function is well defined for $h \geq 0$ because the function $\mathrm{G} \mapsto \sum_{n,m} K(n+m)\exp(-(n+m)\mathrm{G})$ is real analytic on the positive semi-axis and decreases from 1 to 0 as G goes from 0 to ∞. In particular, $\mathrm{G}(0) = 0$. Analyticity of $\mathrm{G}(\cdot)$ on the positive semi-axis follows directly by the Inverse Function Theorem (for analytic functions: see for example [22, Sec. 2.5]). Convexity can be proven directly by differentiating the implicit expression (1.18). A less tedious proof of convexity can be achieved by recognizing that $\mathrm{G}(h)$ is a free energy (see (3.6)). □

In view of (1.17) we see that (limits are with $M \sim \gamma N$)

$$\lim \frac{1}{N} \log Z^c_{N,M,h} = (1+\gamma)\mathrm{G}(h) + \lim \frac{1}{N} \log \mathbf{P}\left((N,M)\in\widetilde{\tau}_h\right). \tag{1.19}$$

Let us focus on $h > 0$. The inter-arrivals of $\widetilde{\tau}_h$ are exponentially integrable, in particular they have finite mean:

$$\mu_h = \left(\mu_h^{(1)}, \mu_h^{(2)}\right) := \mathbf{E}[(\widetilde{\tau}_h)_1] = \left(\sum_{n,m} n\widetilde{K}_h(n,m), \sum_{n,m} m\widetilde{K}_h(n,m)\right). \tag{1.20}$$

Evidently $\mu_h^{(1)} = \mu_h^{(2)}$ and the Law of Large Numbers directly implies that $\widetilde{\tau}_h$ stays close to the diagonal of the first quadrant. But then the event $\{(N,M)\in\widetilde{\tau}_h\}$ is a Large Deviation event for $\gamma \neq 1$ and it contributes to the free energy: the singularities of the free energy for $h > 0$ come from this extra Large Deviation contribution and key word to understand these new transitions is *Cramér regime*. The point in fact is whether or not the Large Deviation event can be made typical by an exponential change of measure (a *tilt*): the larger γ is the more possible it is that a tilt does not suffice and the typical Large Deviation trajectories will not correspond to a tilt of the measure (in this case we say that we are outside of the Cramér regime). On the other hand, the interaction strength directly impacts whether or not the process is in the Cramér regime. The formulas that follow precisely characterize the switching between Cramér and non Cramér regimes.

We introduce the convex function $q_h : \mathbb{R}^2 \to (0, \infty]^2$

$$q_h(\lambda) = q_h(\lambda_1, \lambda_2) := \sum_{n,m} e^h K(n+m) \exp\left(-(\mathrm{G}(h) - \lambda_1)n - (\mathrm{G}(h) - \lambda_2)m\right), \tag{1.21}$$

which is bounded in $(-\infty, \mathrm{G}(h)]^2$ and it is analytic in the interior of this domain. We set for $h > 0$

$$\overline{\lambda}_1(h) := \sup\{\lambda_1 < 0 : q_h(\lambda_1, \mathrm{G}(h)) \le 1\}, \tag{1.22}$$

and, since $q_h(\lambda_1, \mathrm{G}(h))$ increases continuously in λ_1 from $q_h(-\infty, \mathrm{G}(h)) = 0$ to $q_h(0, \mathrm{G}(h)) > 1$, $\overline{\lambda}_1(h)$ is negative and it is characterized by $q_h(\overline{\lambda}_1, \mathrm{G}(h)) = 1$. Finally, we set, always for $h > 0$

$$\gamma_c(h) := \frac{\sum\limits_{n,m} m K(n+m) \exp\left(-n\left(\mathrm{G}(h) - \overline{\lambda}_1(h)\right)\right)}{\sum\limits_{n,m} n K(n+m) \exp\left(-n\left(\mathrm{G}(h) - \overline{\lambda}_1(h)\right)\right)}, \tag{1.23}$$

which is the ratio of averaged strand sizes. Both denominator and numerator are bounded because for $c > 0$ and for every $\alpha \ge 0$

$$\sum_{n,m} (n+m) K(n+m) e^{-cn} = \sum_{t=2}^{\infty} t K(t) \sum_{n=1}^{t-1} \exp(-cn) \le \frac{1}{e^c - 1} \sum_{t=2}^{\infty} t K(t) < \infty. \tag{1.24}$$

Here are some properties (see Section 3.3 for the proof):

Lemma 1.4. *Choose $\alpha > 0$. The function $\gamma_c : (0, \infty) \longrightarrow (1, \infty)$ is real analytic and*

$$\gamma_c(0) := \lim_{h \searrow 0} \gamma_c(h) = \frac{1}{\alpha} \vee 1 \quad \text{and} \quad \gamma_c(\infty) = \frac{\sum\limits_m m K(1+m)}{\sum\limits_m K(1+m)}. \tag{1.25}$$

The examples worked out in Section 3.4 show that $\gamma_c(\cdot)$ can have various behaviors: in particular, in general it is not monotonic.

Theorem 1.5. *Fix $\gamma \ge 1$. The function $\mathrm{F}_\gamma(\cdot)$ is analytic on $\{h : h > 0 \text{such that} \gamma_c(h) - \gamma \ne 0\}$ and $\mathrm{F}_\gamma(\cdot)$ is not analytic for the values $h > 0$ at which $\gamma_c(h) - \gamma$ changes sign. However, $\mathrm{F}_\gamma(\cdot)$ is C^1 on the positive semi-axis.*

Theorem 1.5 is just a sample of the results we have and that can be gotten on these transitions that are transitions between localized regimes, because, by convexity of the free energy, the expected number of contacts does not decrease in h. In particular the *tangential case* – when $h \mapsto \gamma_c(h) - \gamma$ touches zero without changing sign – is treated in detail and while we can deal with most of the cases we are unable to produce a concise statement that says which zeros of $\gamma_c(h) - \gamma$ are critical points (some are not!) and, in general, what is the precise order of

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

the transition. This is due to the fact that these transitions, unlike the denaturation transition, do depend on the details of $K(\cdot)$ and, to a certain extent, one needs to do a case by case study. Examples and more considerations on all these issues are developed at the end of the introduction and in Section 3.4.

1.7. Outline of the approach, sharp estimates and limit path properties

As we already mentioned, the cornerstone is (1.17). In fact (1.17) reduces sharp, respectively Laplace, estimates on $Z^c_{N,M,h}$ to sharp, respectively Laplace, estimates on the renewal function $\mathbf{P}((N, M) \in \widetilde{\tau}_h)$. A quick overview of the behavior of $\mathbf{P}((N, M) \in \widetilde{\tau}_h)$ is

(1) If $h < 0$, so $\widetilde{\tau}_h$ is terminating, we will show that there exists $C_h > 0$ such that

$$\mathbf{P}((N, M) \in \widetilde{\tau}_h) \overset{N,M\to\infty}{\sim} C_h \widetilde{K}_h(N, M). \tag{1.26}$$

(2) For $h > 0$, recall (1.20) and the discussion right after that formula, one can show that

$$\lim_{\varepsilon \searrow 0} \lim_{t\to\infty} \frac{1}{t} \log \mathbf{P}\left(t\{x \in \mathbb{R}^2 : |x - v| \le \varepsilon\} \cap \widetilde{\tau}_h \neq \emptyset\right) = -D_h(v), \tag{1.27}$$

where $D_h(\cdot)$ is a non-negative function defined in $\mathbb{R}^2$, but equal to $+\infty$ outside of the first quadrant. We shall see that $D_h(\cdot)$ is linear along rays, that is $D_h(sv) = sD_h(v)$, and $D_h(v) = 0$ if and only if $v \propto \mu_h$ (for us $v \propto (1, 1)$). By the symmetry $\mathrm{K}(n, m) = \mathrm{K}(m, n)$, we get that $D_h(v_1, v_2) = D_h(v_2, v_1)$ so there is no loss of generality in sticking to $D(1, \gamma)$, $\gamma \ge 1$. Moreover we will then see that $\gamma \mapsto D(1, \gamma)$ is affine for γ larger than a critical value $\gamma_c(h) > 1$ or smaller than another critical value that coincides with $1/\gamma_c(h)$. On the other hand $\gamma \mapsto D(1, \gamma)$ is strictly convex in the interval $(1/\gamma_c(h), \gamma_c(h))$ – the *Cramér region* – and, when γ is in this interval, sharp asymptotic estimates are known. Namely, we have that

$$\mathbf{P}((N, M) \in \widetilde{\tau}_h) \sim \frac{c_v}{\sqrt{t}} \exp(-tD_h(v)), \tag{1.28}$$

where $t = \sqrt{N^2 + M^2}$, $v = (N, M)/t$, c_v is a positive constant (which depends of course also on h and $K(\cdot)$) and the asymptotic statement is for $t \to \infty$ and it is uniform provided that the unit vector v is in a compact arc of circle subset contained in an open arc of the unit circle that goes from $(\gamma_c(h), 1)/\sqrt{1 + (\gamma_c(h))^2}$ to $(1, \gamma_c(h))/\sqrt{1 + (\gamma_c(h))^2}$.

We remark that by putting (1.17) and (1.28) together we readily see that we can make a substantial step ahead with respect to (1.19):

$$\lim_{\substack{N,M\to\infty:\\ M\sim\gamma N}} \frac{1}{N} \log Z^c_{N,M,h} = (1 + \gamma)\mathrm{G} - D_h(1, \gamma). \tag{1.29}$$

From this formula, since G is (implicitly) determined by (1.18) and since $D_h(\cdot)$ has a variational formulation, we will be able to use it to establish Theorems 1.2 and 1.5.

But with what we just outlined we can go beyond Laplace type estimates: (1.26) and (1.28) yield sharp estimates on $Z^c_{N,M,h}$ as $N \to \infty$ for $h < 0$, with $M \sim \gamma N$, any $\gamma > 0$. The same holds for $h > 0$, but only for γ in the Cramér region. We state here the result for the free case, which is less immediate than the constrained one:

Theorem 1.6. *We have the following sharp estimates for $M \sim \gamma N$ and $\alpha > 0$:*
(1) *For $h > 0$ and $\gamma \in (1/\gamma_c(h), \gamma_c(h))$ there exists $c_{\gamma,h} > 0$ such that*

$$Z^f_{N,M,h} \overset{N\to\infty}{\sim} \frac{c_{\gamma,h}}{\sqrt{N}} \exp\left(N \mathrm{F}_{\frac{M}{N}}(h)\right). \tag{1.30}$$

(2) *For $h < 0$, if $\overline{\alpha} < 1 + \alpha/2$ we have*

$$Z^f_{N,M,h} \overset{N\to\infty}{\sim} \frac{K_f(N)K_f(M)}{1-\exp(h)}. \tag{1.31}$$

Moreover, if $\overline{\alpha} > 1 + \alpha/2$

$$Z^f_{N,M,h} \overset{N\to\infty}{\sim} \frac{\exp(h)\left(\sum_{n\geq 0} K_f(n)\right)^2}{(1-\exp(h))^2} \mathrm{K}(N, M). \tag{1.32}$$

Sharp estimates on the partition function lead to sharp control on path properties:

Theorem 1.7. *Choose $\gamma > 0$ and consider the case $M \sim \gamma N$ and $\alpha > 0$.*
(1) *Let $(\mathcal{F}_1, \mathcal{F}_2) := \max\{\tau \cap [0, N] \times [0, M]\}$ be the last renewal epoch in $[0, N] \times [0, M]$. For $h < 0$ and $\overline{\alpha} < 1 + \alpha/2$, the law of $(\mathcal{F}_1, \mathcal{F}_2)$ under $\mathbf{P}^f_{N,M,h}$– a probability measure on $(\{0\} \cup \mathbb{N})^2$– converges for $N \to \infty$ to the probability distribution that assigns to (i, j) probability*

$$(1-\exp(h))\mathbf{P}\left((i, j) \in \widetilde{\tau}_h\right). \tag{1.33}$$

Set $\mathcal{L}_1 := N - \mathcal{F}_1$ and $\mathcal{L}_2 := M - \mathcal{F}_2$. For $h < 0$ and $\overline{\alpha} > 1 + \alpha/2$, the law of $(\mathcal{L}_1, \mathcal{L}_2)$ under $\mathbf{P}^f_{N,M,h}$– a probability measure on $(\{0\}\cup\mathbb{N})^2$– converges for $N \to \infty$ to the probability distribution that assigns to (i, j) probability

$$\frac{1}{\left(\sum_{n\geq 0} K_f(n)\right)^2} K_f(i)K_f(j). \tag{1.34}$$

Moreover, for $h < 0$ and $\overline{\alpha} > 1 + \alpha/2$, we have

$$\lim_{L\to\infty} \lim_{N\to\infty} \mathbf{P}^f_{N,M,h}(\tau \cap [L, N-L] \times [L, M-L] = \emptyset) = 1. \tag{1.35}$$

(2) *For $h > 0$ and $\gamma \in (1/\gamma_c(h), \gamma_c(h))$ we have that both $\mathrm{F}_\gamma(h) - \gamma\partial_\gamma \mathrm{F}_\gamma(h)$ and $\partial_\gamma \mathrm{F}_\gamma(h)$ are positive and the law of $(\mathcal{L}_1, \mathcal{L}_2)$ under $\mathbf{P}^f_{N,M,h}$– a probability measure on $(\{0\} \cup \mathbb{N})^2$– converges for $N \to \infty$ to the probability distribution that assigns to (i, j) probability*

$$\frac{1}{C_{\gamma,h}} K_f(i) \exp\left(-i\left(\mathrm{F}_\gamma(h) - \gamma\partial_\gamma \mathrm{F}_\gamma(h)\right)\right) K_f(j) \exp\left(-j\partial_\gamma \mathrm{F}_\gamma(h)\right) \tag{1.36}$$

with $C_{\gamma,h} > 0$ the normalization constant. Moreover the law of τ under $\mathbf{P}^f_{N,M,h}$– a probability on the subsets of $(\{0\} \cup \mathbb{N})^2$– converges in the same limit to the law of a positive recurrent two-dimensional renewal with inter-arrival law given by the function from $\mathbb{N}^2$ to $[0, 1)$

$$(i, j) \mapsto K(i+j) \exp\left(-i\left(\mathrm{F}_\gamma(h) - \gamma\partial_\gamma \mathrm{F}_\gamma(h)\right) - j\partial_\gamma \mathrm{F}_\gamma(h)\right). \tag{1.37}$$

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

We observe that exploiting the symmetry of the model under the exchange of N and M, from the definition (1.12) of $\mathrm{F}_\gamma(h)$ we directly obtain $\mathrm{F}_\gamma(\cdot) = \gamma \mathrm{F}_{1/\gamma}(\cdot)$. The same symmetry can be appreciated in (1.37) if we make the replacement (i, j, γ) with $(j, i, 1/\gamma)$.

Theorem 1.7 can be summed up as:

(1) In the delocalized phase, $h < 0$ (and $\overline{\alpha} \neq 1 + \alpha/2$, see below), there is no contact in the bulk of the system and, according to whether the $K_f(\cdot)$ exponent $\overline{\alpha}$ is larger or smaller than $1 + \alpha/2$ the two strands are free except for $O(1)$ contacts all close to the origin, or the two strands get detached after finitely many contacts (all close to the origin) and they meet again at a $O(1)$ distance from (N, M), terminating with two free ends of length $O(1)$. In the case $\overline{\alpha} = 1 + \alpha/2$ the slowly varying corrections $L(\cdot)$ and $\overline{L}(\cdot)$ matter and we leave out this rather cumbersome analysis.

(2) In the localized phase ($h > 0$) and for γ in the Cramér region the process converges to a persistent renewal that we determine: this is similar to what happens in the one-dimensional case, but in this new set-up the limit process has the expression (1.37) which is much less straightforward than the corresponding one-dimensional case. A number of other results can be proven, in the spirit of the one-dimensional analogs (see [8] and [16, Ch. 2]), but we have chosen to limit ourselves to Theorem 1.7(2) and we signal that the proof of (1.35), see Section 4.3, is much richer than (1.35).

1.8. *Open issues and perspectives*

We do not treat a number of natural issues: we list and discuss them here.

The non Cramér regime

For $h > 0$ and $\gamma \notin (1/\gamma_c(h), \gamma_c(h))$ we do not give sharp estimates. To our knowledge sharp estimates on the renewal function in this regime are for the moment not available (the most advanced reference available appears to be [5]). The issue is not a secondary one: it is at the heart of understanding the transitions and the different phases that one observes in the localized regime. And what one expects is rather clear: for γ in the Cramér region we have seen that the free ends are microscopic, i.e. $O(1)$, and the limit process is just a recurrent renewal; for γ in the interior of the complementary of the Cramér region, and for the free case, instead a big loop should appear (showing that there is only one) or the free ends, probably only one, should become macroscopic; which of these two phenomenologies prevail should depend on the exponents α and $\overline{\alpha}$. The analysis is certainly different for the constrained case, because the expected big loop can only be along the chain (and, by exchangeability, its location is going to be uniformly distributed along the chain). This and very similar issues are widely discussed in the physical literature [20,24] and the analogy with Bose condensation is regularly invoked, but the analysis is far from being rigorous.

Counting the transitions in the localized regime

We present examples with zero, one or two transitions. Can there be more than two? Are they always finitely many?

Sharp estimates at criticality

If $\sum_n n^3 K(n) < \infty$ sharp estimates for $h = 0$ are covered by (1.28). We have not treated this case here because it would be natural to consider the complete spectrum of loop exponents, but we meet again with the limitations of multivariate renewal theory. The gPS model demands control only on the special class of renewals with inter-arrivals $\mathrm{K}(n, m) = K(n + m)$ and we hope that an ad-hoc treatment will lead to progress. And of course the gPS model is one more motivation for a more systematic study of multivariate renewals.

Disordered interactions

Here the issues are several: we stick to the one of disorder relevance at criticality (see the review of the literature in [18, Ch. 4]), but there are most probably intriguing questions also away from criticality (in analogy with [19]). The effect of disorder on the critical point is directly related to obtaining sharp estimates on the renewal function of the underlying renewal process, at least if the disorder is introduced via an IID family $\{h_{n,m}\}_{(n,m)\in\mathbb{N}^2}$ of random variables. In fact the tools developed for the basic disordered PS model ([18] and references therein) can be applied, but the problem is that sharp estimates are available only for the very particular case of $\sum_n n^3 K(n) < \infty$. And there is the issue that such an IID disorder is not the most suited for DNA modeling, but if the aim is understanding the effect of noise on critical behaviors this way of introducing the disorder is certainly acceptable (and it is what has been done also in the biophysical literature, even sticking to the DNA/RNA set-up! See for example [7]). A more natural disorder is however obtained by assigning to each strand a sequence of, possibly IID, potentials $\{h_j^{(1)}\}_{j\in\mathbb{N}}$ and $\{h_j^{(2)}\}_{j\in\mathbb{N}}$ – one can imagine the case in which the two sequences are independent or the case in which they are (strongly) correlated – and $h_{n,m} = h_n^{(1)} h_m^{(2)}$: note that correlations are introduced with this product choice even if the two sequences are independent. This appears to be a very challenging model (see [1] and references therein for the issues that arise when correlations are introduced in the disorder sequence for the basic PS model).

Related models

The gPS model is intimately related to the more complex RNA models for secondary structure: [12], where the vast literature is cited, is particularly interesting for us because RNA models are linked with the gPS model. Models for circular DNA [29] are also very much related to gPS, as pointed out for example in [15]. In [20] the authors focus on an issue (existence of one macroscopic loop) for circular DNA that is precisely the one that we face outside of the Cramér regime. Finally, the gPS model can be seen as a toy model for interacting self-avoiding walks. In the related direction of simplified models for a self-interacting self-avoiding walk we signal the Zwanzig–Lauritzen model that has been tackled first by generating function techniques (e.g. [6]) and recently by probabilistic methods in [9,25].

1.9. Organization of the paper

In Section 2, we present the results on Large and Sharp Deviations for bivariate renewal processes that we use. In Section 3, we introduce the constrained model and study the free energy in the localized and delocalized regime proving Theorem 1.2, Lemma 1.4 and Theorem 1.5. At the end of this section, we work out explicitly some examples. Finally in Section 4, we prove Proposition 1.1 and we compute the sharp estimates of the free partition function and the path properties proving Theorems 1.6 and 1.7. For the Dominated Converge Theorem we use the shortcut (DOM).

2. Large deviations and local limit theorems for bivariate renewals

We now give a number of results on the renewal $\widetilde{\tau}_h$, $h > 0$, defined at the beginning of Section 1.6. As it will be clear, they follow directly from various results that one can find in

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

[3,4] where a more general case is treated (starting from the fact that we limit ourselves to the two-dimensional case). Let us start by introducing the exponential moment generating function of $\xi := (\widetilde{\tau}_h)_1$ (we recall: $\mathbf{P}(\xi = (n,m)) = \widetilde{K}_h(n,m)$)

$$q_h(\lambda) := \mathbf{E}\left[\exp\left(\langle \lambda, \xi \rangle\right)\right], \tag{2.1}$$

where $\lambda \in \mathbb{R}^2$, $\langle \lambda, \xi \rangle = \lambda_1 \xi_1 + \lambda_2 \xi_2$. From the definition of $\widetilde{K}_h(\cdot,\cdot)$ one readily sees that $q_h(\lambda) < \infty$ if and only if both $\lambda_1 \leq \textsc{g}$ and $\lambda_2 \leq \textsc{g}$. The Large Deviation function $\Lambda(\cdot)$ corresponding to the random vector ξ is the Legendre transform of the function $\log q_h(\cdot)$

$$\Lambda(\theta) = \sup_{\lambda} \left\{ \langle \lambda, \theta \rangle - \log q_h(\lambda) \right\}, \tag{2.2}$$

where $\theta \in \mathbb{R}^2$. Since we are after the renewal function of $\widetilde{\tau}_h$ the Large Deviation function of the inter-arrival random vector is just an intermediate step. The asymptotic behavior of the renewal function is directly related to the so-called *second deviation function*, introduced and investigated in [3]:

$$D_h(\theta) = D(\theta) = \inf_{s>0} \frac{\Lambda(s\theta)}{s}, \tag{2.3}$$

where $\theta \in \mathbb{R}^2$. The notation with the subscript h will be useful further on to remind the dependence on the parameter but at this stage it is rather superfluous.

Remark 2.1. From (2.3) one can see that $D(s\theta) = sD(\theta)$ for every $s \geq 0$. In [3, pp. 652–653] a detailed analysis of $D(\cdot)$ is given, notably the fact that it is convex: for $p \geq 0$, $q \geq 0$, $p + q = 1$ and $\theta, \eta \in \mathbb{R}^2$,

$$D(p\theta + q\eta) \leq pD(\theta) + qD(\eta). \tag{2.4}$$

We can immediately deduce from these properties that for every $\theta, \eta \in \mathbb{R}^2$, we have

$$D(\theta + \eta) \leq D(\theta) + D(\eta). \tag{2.5}$$

Here is an important step:

Proposition 2.2 (*[3, Theorem 1]*)**.** *For every $\theta = (\theta_1, \theta_2) \in \mathbb{R}^2$*

$$D(\theta) = \sup_{\lambda \in A} \langle \lambda, \theta \rangle = \sup_{\lambda \in \partial A} \langle \lambda, \theta \rangle, \tag{2.6}$$

where A is the closed convex set $\left\{\lambda \in \mathbb{R}^2 : q_h(\lambda) \leq 1\right\}$ and ∂A is the boundary of A.

It is now practical to focus on the specific case we are considering, notably the fact that $D(\theta) = \infty$ if θ is not in the first quadrant is an intuitive consequence of the fact that our process has increments that have positive components and can be read out of the structure of A. Let us make A more explicit

$$A = \left\{ \lambda \in \mathbb{R}^2 : q_h(\lambda_1, \lambda_2) \leq 1 \right\}, \tag{2.7}$$

where $q_h(\cdot)$, defined in (2.1), corresponds to (1.21) in our special case and we recall that $h > 0$ and $\textsc{g} = \textsc{g}(h) > 0$ is chosen so that $(0, 0) \in \partial A$. Note that $q_h(\lambda_1, \lambda_2) = q_h(\lambda_2, \lambda_1)$ and that $q_h(\cdot)$ is convex (this of course implies the convexity of A) and it is symmetric with respect to the diagonal of the first and third quadrant.

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

Lemma 2.3. *We have*

$$A \subset \left\{ \lambda \in \mathbb{R}^2 : \lambda_1 \le \mathrm{G} \text{ and } \lambda_2 \le \mathrm{G} \right\} \cap \left\{ \lambda \in \mathbb{R}^2 : \lambda_2 \le -\lambda_1 \right\}. \tag{2.8}$$

Moreover $\overline{\lambda}_1 \overset{(1.22)}{=} \sup\{\lambda_1 < 0 : q_h(\lambda_1, \mathrm{G}) = 1\} < -\mathrm{G}$ *and the equation* $q_h(\lambda_1, \lambda_2) = 1$ *is uniquely solvable for* $(\lambda_1, \lambda_2) \in [\overline{\lambda}_1, \mathrm{G}]^2$, *defining the curve* $\mathcal{W}_h$, *symmetric with respect to the diagonal of the first and third quadrant.* $\mathcal{W}_h$ *is the graph of a concave and decreasing function* $\widetilde{\lambda}_2 : [\overline{\lambda}_1, \mathrm{G}] \to [\overline{\lambda}_1, \mathrm{G}]$ *which satisfies* $\widetilde{\lambda}_2(\overline{\lambda}_1) = \mathrm{G}$, $\widetilde{\lambda}_2(\mathrm{G}) = \overline{\lambda}_1$ *and* $\widetilde{\lambda}_2(0) = 0$. *Moreover* $\widetilde{\lambda}_2$ *is analytic in the interior of its domain and*

$$\partial A = \left\{ (\lambda_1, \mathrm{G}) : \lambda_1 < \overline{\lambda}_1 \right\} \cup \left\{ (\mathrm{G}, \lambda_2) : \lambda_2 < \overline{\lambda}_1 \right\} \cup \mathcal{W}_h. \tag{2.9}$$

Finally, $q_h(\lambda_1, \lambda_2) < 1$ *for* $(\lambda_1, \lambda_2) \in \partial A \setminus \mathcal{W}_h$.

Proof. For this proof it is practical to keep at hand part (LEFT) of Fig. 3. The fact that $A \subset \left\{ \lambda \in \mathbb{R}^2 : \lambda_1 \le \mathrm{G} \text{ and } \lambda_2 \le \mathrm{G} \right\}$ is just the fact that $q_h(\lambda_1, \lambda_2) = \infty$ if $\lambda_1 \wedge \lambda_2 > \mathrm{G}$. On the other hand this last observation, coupled with $q_h(\lambda_1, \lambda_2) = q_h(\lambda_2, \lambda_1)$, $q_h(0, 0) = 1$ and convexity of A, tells us that A does not go above the line $\lambda_2 = -\lambda_1$ and (2.8) is established. Now, since $q_h(\cdot, \cdot)$ is a separately non-decreasing function (and even increasing where it is bounded), ∂A contains $\{\lambda : q_h(\lambda) = 1\}$. So the issue is the solvability of $q_h(\lambda) = 1$ and it is straightforward to see that $q_h(\lambda) = 1$ has a (unique, by monotonicity) solution if $\lambda_1 \in [\overline{\lambda}_1, \mathrm{G}]$ and this way we define a function $\widetilde{\lambda}_2(\cdot) : [\overline{\lambda}_1, \mathrm{G}] \mapsto [\overline{\lambda}_1, \mathrm{G}]$ (note that at this stage it is already clear that $\overline{\lambda}_1 \le -\mathrm{G}(h)$) which is analytic in the interior of its domain, by the analytic Implicit Function Theorem. It is actually immediate to check that this function is not linear (for example, compute the second derivative at the origin), so $\overline{\lambda}_1 < -\mathrm{G}(h)$. Finally $q_h(\lambda_1, \mathrm{G}) < 1$ for $\lambda_1 < \overline{\lambda}_1$ and, by symmetry, $q_h(\mathrm{G}, \lambda_2) < 1$ for $\lambda_2 < \overline{\lambda}_1$. The proof is therefore complete. □

Remark 2.4. Lemma 2.3 is given for fixed $h > 0$, but $\overline{\lambda}_1$ depends also on h and when we need to make this dependence explicit we write $\overline{\lambda}_1(h)$. The same is true for the function $\widetilde{\lambda}_2(\cdot)$ and we write $\widetilde{\lambda}_{2,h}(\cdot)$. Note that the analyticity of $\overline{\lambda}_1(\cdot)$ on the positive semi-axis is a direct consequence of the Analytic Implicit Function Theorem ([22, Sec. 2.3], and it is just a matter of analyticity in one variable). If instead we consider the function $(\lambda_1, h) \mapsto \widetilde{\lambda}_{2,h}(\lambda_1)$, with $\widetilde{\lambda}_{2,h}(\lambda_1)$ which is obtained by solving for λ_2 the equation $q_h(\lambda_1, \lambda_2) = 1$ and we have seen that this requires $\lambda_1 \in [\overline{\lambda}_1(h), \mathrm{G}(h)]$. Therefore, by the Analytic Implicit Function Theorem [22, Sec. 2.3], the function $(\lambda_1, h) \mapsto \widetilde{\lambda}_{2,h}(\lambda_1)$ is analytic (in two variables this time) in the domain $h > 0$ and $\lambda_1 \in (\overline{\lambda}_1(h), \mathrm{G}(h))$.

From Proposition 2.2 and Lemma 2.3 we can derive a number of consequences, like the fact that $D(\theta) < \infty$ for every θ in the first quadrant (that here includes the two axes) and that, when $D(\theta) < \infty$, that is in the first quadrant, there are only two possibilities: either the supremum in the rightmost term in (2.6) is reached in the interior of $\mathcal{W}_h$ or *at the boundary*, that is in $\{(\overline{\lambda}_1, \mathrm{G}), (\mathrm{G}, \overline{\lambda}_1)\}$. We observe by direct inspection that if $\theta_2 > \theta_1$ then if the supremum is not achieved in the interior, then it is achieved at $(\overline{\lambda}_1, \mathrm{G})$. Moreover if $\theta_1 = \theta_2 > 0$ the supremum is always achieved in the interior and, more precisely, at $(0, 0)$ and therefore $D(\theta_1, \theta_1) = 0$. This induces a partition of the first quadrant: $\theta \in E_h$ if the supremum is achieved in the interior and $\theta \in E_h^{\complement}$ if it is achieved at the boundary. It is also useful to remark that E_h is an open sector ([3, p. 653], or it can be seen directly in our specific set-up): in our case it is also symmetric with respect to the diagonal of the first quadrant, that is there exists $\varphi \in (0, \pi/4)$ such that $E_h = \{\theta \in [0, \infty)^2 : \theta_2/\theta_1 \in (\tan(-\varphi + \pi/4), \tan(\varphi + \pi/4))\}$.

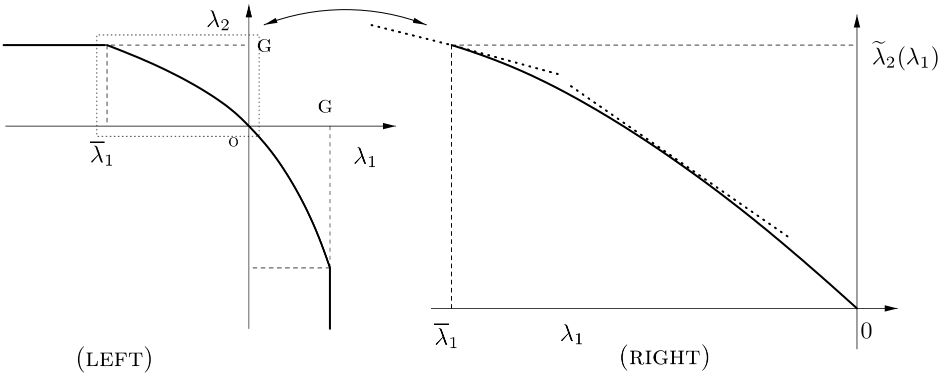

Fig. 3. In part (LEFT) ∂A, with $A = \{\lambda : q_h(\lambda_1, \lambda_2) \le 1\}$, is represented by the thick line. Notice the symmetries and the convexity of A. In part (RIGHT) we zoom into the relevant part for our analysis.

Remark 2.5. If $\theta \in E_h$ then [3, Theorem 2] tells us that the infimum in (2.3) is attained at a unique point $s(\theta)$, so $D(\theta) = \Lambda(\theta s(\theta))/s(\theta)$, and $\Lambda(\cdot)$ is analytic and strictly convex at $s(\theta)$. This is what we may call the *Cramér region* of parameters: such a region being the set of θ's for which the large deviation trajectories can be made typical by a suitable change of measure (*tilting*).

We are now ready to state the result that links $D(\cdot)$ and the renewal function. Once again it is obtained by restricting to our context a result – [3, Theorem 5] – see however Remark 2.7. We employ the notation $\Lambda''(\theta)$ for the Hessian matrix of Λ.

Proposition 2.6. *For every $v = (v_1, v_2) = (N, M)/t \in E_h \cap \{\theta : |\theta| = 1\}$ with $t = \sqrt{N^2 + M^2}$, the following representation holds:*

$$\mathbf{P}\left((N, M) \in \widetilde{\tau}_h\right) = \frac{1}{t^{1/2}} \left(\sqrt{\frac{\det\left(\Lambda''(s(v)v)\right)}{2\pi s(v) \langle v, \Lambda''(s(v)v)v \rangle}} + \varepsilon(N, M) \right) \exp\left(-t D(v)\right), \quad (2.10)$$

and for every compact $C \subseteq E_h \cap \{\theta : |\theta| = 1\}$

$$\lim_{t \to \infty} \sup_{\substack{N,M: \sqrt{N^2+M^2}=t \\ v \in C}} \varepsilon(N, M) = 0. \quad (2.11)$$

Recalling (1.17), we see that Proposition 2.6 implies the sharp estimate

$$Z^c_{N,M,h} = \frac{A(M/N) + \widetilde{\varepsilon}(N, M)}{\sqrt{N}} \exp\left((N + M)\textsc{g}(h) - N D_h(1, M/N)\right), \quad (2.12)$$

where $A(M/N)$ is equal to $(1 + (M/N)^2)^{-1/4}$ times the square root term in (2.10) (which depends only on v, which in turn is just a function of M/N) and $\widetilde{\varepsilon}(N, M) = (1 + (M/N)^2)^{-1/4} \varepsilon(N, M)$.

Remark 2.7. In [3] the factor $s(v)$ that appears just after 2π in (2.10) has the exponent $d+3 = 5$. This formula appears also in [4, p. 11], with an additional oversight. The formula we give is in agreement with [11] who covers only the case $v \propto \mu_h$ (cf. (1.20)), but in greater

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

generality than [3,4]. Formula (2.10) a priori requires some exponential decay of the inter-arrival law to make sure that the Hessian is computed at an analyticity point of $\Lambda(\cdot)$. Actually, the analysis in [11] shows that this is not necessary and (2.10) still holds true for $h = 0$ if $\sum_{n,m\geq 1}(n+m)^2 K(n+m) = \sum_{t\geq 2} t^2(t-1)K(t) < \infty$. With the help of the notation in [11] we remark that (2.10) can be made slightly more readable if we observe that, as it is well known, the inverse $B(\theta)$ of $\Lambda''(\theta)$ is the covariance matrix of the *tilted* random vector $X = (X^1, X^2)$ with $\mathbf{P}(X = (n,m)) = \widetilde{K}_h(n,m)\exp(n\lambda_1(\theta)+m\lambda_2(\theta))/C_\theta$ and $C_\theta = \sum_{n,m}\widetilde{K}_h(n,m)\exp(n\lambda_1(\theta)+m\lambda_2(\theta))$ and $\lambda(\theta)$ is the optimal point of (2.2). Therefore with $v = \theta/|\theta|$

$$\frac{\det\left(\Lambda''(\theta)\right)}{\langle v, \Lambda''(\theta)v\rangle} = \frac{1}{\det\left(B(\theta)\right)\left\langle v, (B(\theta))^{-1}v\right\rangle}$$
$$= \left\langle v, \begin{pmatrix} \mathbf{E}[(X^2)^2] - \mathbf{E}[X^2]^2 & -\mathbf{E}[X^1X^2] + \mathbf{E}[X^1]\mathbf{E}[X^2] \\ -\mathbf{E}[X^1X^2] + \mathbf{E}[X^1]\mathbf{E}[X^2] & \mathbf{E}[(X^1)^2] - \mathbf{E}[X^1]^2 \end{pmatrix} v \right\rangle^{-1}. \tag{2.13}$$

A Local Limit Theorem, analogous to Proposition 2.6, for $v \in E_h^{\complement}$ is available at the moment (see [5, Theorem 2.1]) only if the entries of $\Lambda''(v)$ are all finite, and this is not always the case in our set-up, notably it is not the case if the exponent α entering the definition of $K(\cdot)$ is smaller than two. Nevertheless, the following weaker result will suffice for our purposes:

Proposition 2.8. *For every θ*

$$\limsup_{t\to\infty} \frac{1}{t}\log \mathbf{P}\left(\lceil t\theta\rceil \in \widetilde{\tau}_h\right) \leq -D(\theta), \tag{2.14}$$

and if $\theta \in E_h^{\complement}$, with $\theta_1 > 0$ and $\theta_2 > 0$, then

$$\lim_{t\to\infty} \frac{1}{t}\log \mathbf{P}\left(\lceil t\theta\rceil \in \widetilde{\tau}_h\right) = -D(\theta). \tag{2.15}$$

Of course (2.15) holds also for $\theta \in E_h$ as an immediate consequence of Proposition 2.6. In Proposition 2.8, we have used the notation $\lceil(x,y)\rceil = (\lceil x\rceil, \lceil y\rceil)$.

Proof. The upper bound is a direct consequence of the Large Deviations Principle [3, Theorem 4]:

$$\lim_{\varepsilon\searrow 0}\lim_{t\to\infty} \frac{1}{t}\log \mathbf{P}\left(t\{v\in\mathbb{R}^2 : |v-\theta|\leq\varepsilon\}\cap\widetilde{\tau}_h \neq \emptyset\right) = -D(\theta). \tag{2.16}$$

For the lower bound we assume without loss of generality that $\theta_2 > \theta_1$ and we observe that if $\theta \in E_h^{\complement}$ – a closed set (recall Remark 2.5 and the explanation that precedes it) – then either θ is in the boundary or in the interior of $E_h^{\complement}$. If it is in the interior then there exists $\theta_2^\star < \theta_2$ with $(\theta_1, \theta_2^\star)$ in the boundary of $E_h^{\complement}$, so that for every $\varepsilon > 0$ small we have $(\theta_1, \theta_2^\star - \varepsilon) \in E_h$. If θ is in the boundary of $E_h^{\complement}$ we directly set $\theta_2^\star = \theta_2$. By Proposition 2.6

$$\lim_{t\to\infty} \frac{1}{t}\log \mathbf{P}\left(\left(\lceil t\theta_1\rceil - 1, \left\lceil t(\theta_2^\star-\varepsilon)\right\rceil\right) \in \widetilde{\tau}_h\right) = -D_h\left((\theta_1, \theta_2^\star-\varepsilon)\right). \tag{2.17}$$

But

$$\mathbf{P}\left(\lceil t\theta\rceil \in \widetilde{\tau}_h\right) \geq \mathbf{P}\left(\left(\lceil t\theta_1\rceil - 1, \left\lceil t(\theta_2^\star-\varepsilon)\right\rceil\right) \in \widetilde{\tau}_h\right) \widetilde{K}_h\left(1, \lceil t\theta_2\rceil - \left\lceil t(\theta_2^\star-\varepsilon)\right\rceil\right), \tag{2.18}$$

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

and therefore we have

$$\liminf_{t\to\infty} \frac{1}{t} \log \mathbf{P}\left(\lceil t\theta \rceil \in \widetilde{\tau}_h\right) \geq -D_h\left(\left(\theta_1, \theta_2^\star - \varepsilon\right)\right) - \mathrm{G}(h)\left(\theta_2 - \theta_2^\star + \varepsilon\right), \tag{2.19}$$

and, by continuity of $D_h(\cdot)$, we get to

$$\liminf_{t\to\infty} \frac{1}{t} \log \mathbf{P}\left(\lceil t\theta \rceil \in \widetilde{\tau}_h\right) \geq -D_h\left(\left(\theta_1, \theta_2^\star\right)\right) - \mathrm{G}(h)\left(\theta_2 - \theta_2^\star\right). \tag{2.20}$$

Since $\left(\theta_1, \theta_2^\star\right) \in E_h^{\complement}$ we have that

$$D_h\left(\left(\theta_1, \theta_2^\star\right)\right) + \mathrm{G}(h)\left(\theta_2 - \theta_2^\star\right) = \overline{\lambda}_1\theta_1 + \mathrm{G}(h)\theta_2^\star + \mathrm{G}(h)\left(\theta_2 - \theta_2^\star\right) = \overline{\lambda}_1\theta_1 + \mathrm{G}(h)\theta_2, \tag{2.21}$$

and since $\theta \in E_h^{\complement}$ the rightmost term is $D_h(\theta)$ and we are done. □

3. The constrained model

3.1. The free energy of the constrained model

We assume $M \geq N$. The first result we present is:

Proposition 3.1. *For every $\gamma \geq 1$*

$$\widetilde{\mathrm{F}}_\gamma(h) := \lim_{\substack{N,M\to\infty:\\ \frac{M}{N}\to\gamma}} \frac{1}{N} \log Z^c_{N,M,h} = \begin{cases} 0 & \text{if } h \leq 0, \\ (1+\gamma)\mathrm{G}(h) - D_h(1,\gamma) & \text{if } h > 0, \end{cases} \tag{3.1}$$

and in fact we have also that for every $L > 1$

$$\lim_{\varepsilon\searrow 0} \lim_{N\to\infty} \sup_{\gamma\in[1,L]} \sup_{M:|(M/N)-\gamma|\leq\varepsilon} \left| \frac{1}{N} \log Z^c_{N,M,h} - \widetilde{\mathrm{F}}_\gamma(h) \right| = 0. \tag{3.2}$$

Moreover we have

$$D_h(1,\gamma) = \max_{\lambda\in B_h} \left(\lambda_1 + \gamma\lambda_2\right), \tag{3.3}$$

with

$$B_h = \left\{ \lambda : \overline{\lambda}_1 \leq \lambda_1 \leq 0,\ 0 \leq \lambda_2 \leq \mathrm{G},\ \sum_{n,m} \widetilde{K}_h(n,m)\exp(\lambda_1 n + \lambda_2 m) = 1 \right\}. \tag{3.4}$$

Finally $\widetilde{\mathrm{F}}_\gamma(h) > 0$ for $h > 0$ and in fact

$$2\mathrm{G}(h) \leq \widetilde{\mathrm{F}}_\gamma(h) \leq (1+\gamma)\mathrm{G}(h). \tag{3.5}$$

Proof. Let us start by observing that

$$\mathrm{G}(h) = \lim_{N\to\infty} \frac{1}{2N} \log Z^c_{N,N,h} = \frac{1}{2}\widetilde{\mathrm{F}}_1(h). \tag{3.6}$$

This can be seen, for $h \leq 0$, by the elementary argument presented just after (1.14) and, for $h > 0$, by (1.17) and Proposition 2.6: in fact $D(1,1) = 0$. It is slightly more practical to establish first a result which is a priori weaker than (3.1). Consider first the limit for $N \to \infty$

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

and $M = \lfloor \gamma N \rfloor$ and let us establish the rightmost equality in (3.1) with this notion of limit: call $\breve{\mathrm{F}}_\gamma(h)$ this expression. Again, the case $h \le 0$ of (3.1) is treated by elementary methods just after the statement of Proposition 1.1. For the case $h > 0$ we observe that by (1.17) it suffices to show that

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbf{P}\left((N, \lfloor \gamma N \rfloor) \in \widetilde{\tau}_h\right) = -D_h(1, \gamma). \tag{3.7}$$

But (3.7) is a direct consequence of Propositions 2.6 and 2.8, so the weaker version of (3.1) is established. Now we step to (3.2): if we establish (3.2) with $\widetilde{\mathrm{F}}_\gamma(h)$ replaced by $\breve{\mathrm{F}}_\gamma(h)$, then the original version of (3.1) holds, which implies that $\breve{\mathrm{F}}_\gamma(h) = \widetilde{\mathrm{F}}_\gamma(h)$ and, in turn, that also the original version of (3.2) holds. It is therefore a matter of comparing (uniformly) the limits along all sequences of (M, N), with $|(M/N) - \gamma| \le \varepsilon$, with the special case $M = \lfloor \gamma N \rfloor$. But this is obtained for example by exploiting that if $M' > M$ we have that for any $C_1 > 2 + \alpha$ there exists $C_2 > 0$ such that for every $N \ge 1$

$$Z^c_{N,M,h} \le C_2 {M'}^{C_1} Z^c_{N,M',h}. \tag{3.8}$$

(3.8) can be established by using $K(n)/K(n+m) \le 1/K(n+m) \le C_2 m^{C_1}$ for $m \ge n$. More precisely we apply this inequality to bound $1/K(l_n + t_n + M' - M)$ (using the fact that $l_n + t_n + M' - M \le N + M' \le (1 + 1/\gamma)M'$) in (1.9) to stretch the last renewal so that it matches the boundary constraint (N, M'), and then we allow n to go up to $N \wedge M'$, which in this case is N anyways, and the new constraint on $\underline{t}$ is $|\underline{t}| = M'$. Inequality (3.8) can then be used to sandwich the partition functions $Z^c_{N,M,h}$, with $|(M/N) - \gamma| \le \varepsilon$ and N sufficiently large, between $Z^c_{N,\lfloor(\gamma-2\varepsilon)N\rfloor,h}/N^{2C_1}$ and $Z^c_{N,\lfloor(\gamma+2\varepsilon)N\rfloor,h}N^{2C_1}$. The continuity of $\gamma \mapsto \breve{\mathrm{F}}_\gamma(h)$, and hence the uniform continuity and boundedness on compact sets – in our case $[1, L]$ – completes the argument and the proof of (3.1) and (3.2).

Let us check (3.3): it is of course a matter of replacing ∂A in (2.6), Proposition 2.2, by B_h. Much of the work has been done in Lemma 2.3: we are just left with showing that we can restrict the supremum to B_h. First of all the symmetry of ∂A tells us that $\gamma \ge 1$ implies that $\sup_{\partial A}(\lambda_1 + \gamma\lambda_2)$ does not change if we restrict ∂A to $\lambda_1 \le 0$. Moreover, by Lemma 2.3, if $\lambda_1 < \overline{\lambda}_1$, then $(\lambda_1, \mathrm{G}) \in \partial A$, but since $\lambda_1 + \gamma \mathrm{G} < \overline{\lambda}_1 + \gamma \mathrm{G}$ for $\lambda_1 < \overline{\lambda}_1$ we can actually neglect these points in taking the supremum.

We are left with the positivity of $\widetilde{\mathrm{F}}_\gamma(h)$ for $h > 0$. This follows directly by observing that

$$Z^c_{N,M,h} \ge Z^c_{N-1,N-1,h} \exp(h) K(M - N + 2), \tag{3.9}$$

which implies more than the positivity, that is $\widetilde{\mathrm{F}}_\gamma(h) \ge 2\mathrm{G}(h)$. Finally, by exploiting also that $D_h(\cdot) \ge 0$ (for $h > 0$) and that in any case $\widetilde{\mathrm{F}}_\gamma(h) = 0$ for $h \le 0$, we see that (3.5) holds and the proof of Proposition 3.1 is therefore complete. □

We can go beyond Proposition 3.1 by exploiting the variational problem (3.3)–(3.4). Note that since $\widetilde{\mathrm{F}}_1(h) = 2\mathrm{G}(h)$ we know that $\widetilde{\mathrm{F}}_1(\cdot)$ is analytic except at the origin, but for $\gamma > 0$ the situation is more involved.

Proposition 3.2. *The function $\gamma_c : (0, \infty) \to (1, \infty)$, defined in (1.23), is real analytic. Moreover $\widetilde{\mathrm{F}}_\gamma(\cdot)$ is analytic on the positive semi-axis out of the set $\{h : \gamma_c(h) - \gamma = 0\}$, but $\widetilde{\mathrm{F}}_\gamma(\cdot)$ is not analytic at the values h at which $\gamma_c(h) - \gamma$ changes sign.*

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

Remark 3.3. Of course the regularity issue is not completely resolved by Proposition 3.2, both because it does not make clear whether or not all the points in the discrete set $\{h : \gamma_c(h) - \gamma = 0\}$ are non-analyticity points and because it does not specify the type of singularities.

Proof. Let us first prove the first statement (which is also the first statement in Lemma 1.4). Recall the function $q_h(\cdot)$ and $\overline{\lambda}_1$ from (1.21) and (1.22). Recall also the definition of $\lambda_1 \mapsto \widetilde{\lambda}_2(\lambda_1)$ in Lemma 2.3: $\widetilde{\lambda}_2(\lambda_1)$ is the only solution in λ_2 to $q_h(\lambda_1, \lambda_2) = 1$, for $\lambda_1 \in [\overline{\lambda}_1, \mathrm{G}]$ (see Fig. 3). Recall also that $\widetilde{\lambda}_2(\cdot)$ is concave – i.e. concave down – and analytic in $(\overline{\lambda}_1, 0)$. We have that $\widetilde{\lambda}_2'(\lambda_1)$ equals $-\partial_{\lambda_1} q_h(\lambda_1, \lambda_2)/\partial_{\lambda_2} q_h(\lambda_1, \lambda_2)$ evaluated at $\lambda_2 = \widetilde{\lambda}_2(\lambda_1)$. Therefore

$$\lim_{\lambda_1 \searrow \overline{\lambda}_1} \widetilde{\lambda}_2'(\lambda_1) = \widetilde{\lambda}_2'\left(\overline{\lambda}_1^+\right) = -\frac{\sum_{n,m} nK(n+m) \exp\left(h - n\left(\mathrm{G}(h) + \left|\overline{\lambda}_1\right|\right)\right)}{\sum_{n,m} mK(n+m) \exp\left(h - n\left(\mathrm{G}(h) + \left|\overline{\lambda}_1\right|\right)\right)} \in (-1, 0). \tag{3.10}$$

The denominator is bounded because of (1.24) and the fact that the ratio is bounded below by -1 is just a consequence of concavity and the fact that $\widetilde{\lambda}_2'(0) = -1$. Moreover $\widetilde{\lambda}_2'\left(\overline{\lambda}_1^+\right)$ is a function of h: with the notations in Remark 2.4 we write rather $\widetilde{\lambda}_{2,h}'\left(\overline{\lambda}_1(h)^+\right)$ and a look at the right-hand side of (3.10), recalling that $\overline{\lambda}_1(\cdot)$ is analytic (cf. Remark 2.4), suffices to see that $h \mapsto \widetilde{\lambda}_2'\left(\overline{\lambda}_1^+\right)$ is analytic. Since $\gamma_c(h) = -1/\widetilde{\lambda}_2'\left(\overline{\lambda}_1^+\right)$, the properties of $\gamma_c(\cdot)$ claimed in the statement are proven.

We are now at the heart of the argument: since the function to be maximized, $\lambda_1 + \gamma\lambda_2$ is constant on lines with slope $-1/\gamma$, the maximum is achieved

(1) at a value $\lambda_1 \in (\overline{\lambda}_1, 0)$ if $1/\gamma > -\widetilde{\lambda}_2'\left(\overline{\lambda}_1^+\right)$: the value of λ_1 is of course found by solving $\widetilde{\lambda}_2'(\lambda_1) = -1/\gamma$: this is what we call the Cramér regime (we have a tilted measure that makes typical the Large Deviation event that the renewal follows the slope γ);
(2) at $\lambda_1 = \overline{\lambda}_1$ if $1/\gamma \le -\widetilde{\lambda}_2'\left(\overline{\lambda}_1^+\right)$: we are outside of the Cramér regime.

The situation is therefore that varying h one may switch in and out of the Cramér regime. Out of the Cramér regime the free energy is actually equal to $\mathrm{G}(h) - \overline{\lambda}_1(h)$, and the function $h \mapsto \mathrm{G}(h) - \overline{\lambda}_1(h)$ is analytic for every $h > 0$ (but it is not necessarily equal to the free energy $\widetilde{\mathrm{F}}_\gamma(h)$!). In the Cramér regime the free energy is strictly smaller than $\mathrm{G}(h) - \overline{\lambda}_1(h)$, just because the maximum of $\lambda_1 + \gamma\lambda_2$ is achieved on the boundary and $D_h(1, \gamma)$ contributes to the free energy with a negative sign, cf. (3.1). This explains why the changes of sign of $h \mapsto 1/\gamma + \widetilde{\lambda}_2'\left(\overline{\lambda}_1^+\right)$ are non-analyticity points. □

3.2. Analysis of the free energy singularities in the localized regime

We now go deeper into the analysis of the singularities for the non-analyticity points we have found, that is the values of $h > 0$ for which $\gamma_c(h) - \gamma$ (cf. (1.23)) changes sign. At the same time we tackle also the case in which $\gamma_c(h) - \gamma$ hits zero without crossing it.

For this and referring to the list of two items of the previous proof we introduce some notations.

(1) In the Cramér regime we introduce the notation $\left(\hat{\lambda}_1(h, \gamma), \hat{\lambda}_2(h, \gamma)\right) \in B_h$ for the optimal point (for the sake of brevity, we omit the dependence on γ). It is of course found by

solving $\widetilde{\lambda}_2'(\lambda_1) = -1/\gamma$ that yields $\hat{\lambda}_1(h)$ and $\hat{\lambda}_2(h) = \widetilde{\lambda}_2\left(\hat{\lambda}_1(h)\right)$. Note that the values of h corresponding to the Cramér regime is a union of open disjoint intervals: we call these intervals $\mathcal{I}_1, \mathcal{I}_2, \ldots$. The analyticity of $\hat{\lambda}_1(\cdot)$ and $\hat{\lambda}_2(\cdot)$ in the intervals $\mathcal{I}_j$ follows by Remark 2.4 and by repeating the arguments in the same remark. We also introduce

$$\hat{c}_\gamma(h) := \left(\mathrm{G}(h) - \hat{\lambda}_1(h)\right) + \gamma\left(\mathrm{G}(h) - \hat{\lambda}_2(h)\right), \tag{3.11}$$

for $h \in \cup_j \mathcal{I}_j$ and of course $\widetilde{\mathrm{F}}_\gamma(h) = \hat{c}_\gamma(h)$ on this set. Note that $\cup_j \mathcal{I}_j$ can be alternatively characterized as $\{h : \gamma_c(h) - \gamma > 0\}$.

(2) Out of the Cramér regime, that is for $h \in \mathcal{I}_b := (0, \infty) \setminus \cup_j \mathcal{I}_j$, we introduce instead

$$\mathrm{N}(h) := \mathrm{G}(h) - \overline{\lambda}_1(h), \tag{3.12}$$

and $\widetilde{\mathrm{F}}_\gamma(h) = \mathrm{N}(h)$ if and only if $h \in \mathcal{I}_b$. Note that $\mathcal{I}_b$ can be alternatively characterized as $\{h : \gamma_c(h) - \gamma \le 0\}$. Note also that (cf. the end of the proof of Proposition 3.2) $\mathrm{N}(h) > \hat{c}_\gamma(h)$ in the interior on the intervals on which $\hat{c}_\gamma(\cdot)$ is defined.

Observe moreover that $\partial \cup_j \mathcal{I}_j = \partial\mathcal{I}_b$, here ∂A denotes the boundary of A seen as a subset of $(0, \infty)$, and if $h \in \partial\mathcal{I}_b$ we have $\hat{c}_\gamma(h) = \mathrm{N}(h)$ and $\gamma_c(h) = \gamma$. By differentiating (1.23) we obtain that if $\gamma_c(h) = \gamma$

$$\gamma_c'(h) = -\frac{\mathrm{N}'(h)}{\sum_{n,m} nK(n+m)\exp(-\mathrm{N}(h)n)} \sum_{n,m} n(m-\gamma n)K(n+m)\exp(-n\mathrm{N}(h)), \tag{3.13}$$

hence (with the convention $\mathrm{sign}(0) = 0$)

$$\mathrm{sign}\left(\gamma_c'(h)\right) = -\,\mathrm{sign}\left(\sum_{n,m} n(m-\gamma n)K(n+m)\exp(-n\mathrm{N}(h))\right), \tag{3.14}$$

which is saying in particular that the sum in the right-hand side is zero if and only if $\gamma_c'(h)$ is zero.

In preparation of the next result, that investigates the regularity of the critical points in the positive semi-axis, it is useful to go through what may happen in the *tangential cases*, namely when $\gamma_c(h_0) = \gamma$ and $\gamma_c'(h_0) = 0$, for $h_0 > 0$. There are three different scenarios

(1) $\gamma_c(h_0 \pm \epsilon) - \gamma < 0$ for every $\epsilon > 0$ small, that is h_0 is a maximum, and in this case at h_0 there is no phase transition simply because $\widetilde{\mathrm{F}}_\gamma(h) = \mathrm{N}(h)$ in a neighborhood of h_0, and $\mathrm{N}(\cdot)$ is analytic in the positive semi-axis;

(2) $\gamma_c(h_0 \pm \epsilon) - \gamma > 0$ for every $\epsilon > 0$ small, that is h_0 is a minimum, and, as we will see in the next statement, $\widetilde{\mathrm{F}}_\gamma(\cdot)$ is at least C^2 at h_0, but we are not sure in general that h_0 is a critical point. In fact higher derivatives may exist, see Remark 3.5, but this depends on fine details and we cannot exclude that all of them exist and even that there is no singularity at all in some very special case;

(3) $\gamma_c(\cdot)$ changes sign at h_0, that is h_0 is a saddle: in this case, as we have seen in Proposition 3.2, there is a transition at h_0 and, as we will point out in Remark 3.5, this transition is smoother than in the case in which the derivative of $\gamma_c'(\cdot)$ is not zero.

Proposition 3.4. *Consider $h_0 \in \partial\mathcal{I}_b$, that is $h_0 > 0$ such that $\gamma_c(h_0) = \gamma$. The function $\widetilde{\mathrm{F}}_\gamma'(\cdot)$ is continuous at h_0, so a transition at $h > 0$ is not of first order. If $\sum_m m^2 K(m) < \infty$,*

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

$\widetilde{\mathrm{F}}''_\gamma(\cdot)$ *has a jump discontinuity at* h_0 *(second order transition) if and only if* $\gamma_c'(h_0) \neq 0$. *If* $\sum_m m^2 K(m) = \infty$, $\widetilde{\mathrm{F}}''_\gamma(\cdot)$ *is continuous at* h_0 *(so the transition is of third order or more).*

Note that this statement says in particular that $\widetilde{\mathrm{F}}''_\gamma(\cdot)$ is continuous in full generality if $\gamma_c'(h_0) = 0$.

Proof. Let us first prove that $\widetilde{\mathrm{F}}'_\gamma(h)$ is continuous at h_0. For $\mathrm{N}(h)$, recall from Lemma 2.3 that $q_h\left(\overline{\lambda}_1(h), \mathrm{G}(h)\right) = 1$ (see (1.21) for the definition of $q_h(\cdot)$), therefore

$$\frac{\partial}{\partial h} q_h\left(\overline{\lambda}_1(h), \mathrm{G}(h)\right) = 0, \tag{3.15}$$

which directly implies that

$$\mathrm{N}'(h) \sum_{n,m} nK(n+m) \exp\left(-n\mathrm{N}(h)\right) = \exp(-h). \tag{3.16}$$

For $\hat{c}_\gamma(h)$, first replace $\mathrm{G}(h) - \hat{\lambda}_1(h)$ by $\hat{c}_\gamma(h) - \gamma\left(\mathrm{G}(h) - \hat{\lambda}_2(h)\right)$ (from (3.11)) in $q_h(\cdot)$ to obtain $q_h\left(\hat{\lambda}_1(h), \hat{\lambda}_2(h)\right) = \sum_{n,m} K(n+m) \exp\left(h - \hat{c}_\gamma(h)n - \left(\mathrm{G}(h) - \hat{\lambda}_2(h)\right)(m - \gamma n)\right)$. Recall that $\widetilde{\lambda}_2'\left(\hat{\lambda}_1(h)\right) = -1/\gamma$, which can be rewritten as (using the argument above (3.10))

$$\sum_{n,m} (m - \gamma n) K(n+m) \exp\left(h - \hat{c}_\gamma(h)n - \left(\mathrm{G}(h) - \hat{\lambda}_2(h)\right)(m - \gamma n)\right) = 0, \tag{3.17}$$

for every $h > 0$. Keeping in mind this equality in evaluating $\frac{\partial}{\partial h} q_h(\hat{\lambda}_1(h), \hat{\lambda}_2(h)) = 0$, we obtain

$$\hat{c}_\gamma'(h) \sum_{n,m} nK(n+m) \exp\left(-\hat{c}_\gamma(h)n - \left(\mathrm{G}(h) - \hat{\lambda}_2(h)\right)(m - \gamma n)\right) = \exp(-h). \tag{3.18}$$

Since we have that $\left(\hat{\lambda}_1(h_0), \hat{\lambda}_2(h_0)\right) = \left(\overline{\lambda}_1(h_0), \mathrm{G}(h_0)\right)$ and $\mathrm{N}(h_0) = \hat{c}_\gamma(h_0)$ (from the fact that $h_0 \in \partial \mathcal{I}_{\mathrm{b}}$), by evaluating (3.16) and (3.18) at h_0 we get

$$\widetilde{\mathrm{F}}'_\gamma(h_0) = \mathrm{N}'(h_0) = \hat{c}_\gamma'(h_0) = \left(\sum_{n,m} nK(n+m) \exp\left(h_0 - n\mathrm{N}(h_0)\right)\right)^{-1}, \tag{3.19}$$

and the continuity of $\widetilde{\mathrm{F}}'_\gamma(h)$ at h_0 is proven.

Now by differentiating once again (3.16) we have

$$\mathrm{N}''(h) = \frac{-\exp(-h) + \left(\mathrm{N}'(h)\right)^2 \sum_{n,m} n^2 K(n+m) \exp\left(-n\mathrm{N}(h)\right)}{\sum_{n,m} nK(n+m) \exp\left(-n\mathrm{N}(h)\right)}, \tag{3.20}$$

and by differentiating (3.18)

$$\hat{c}''_\gamma(h) = \frac{-e^{-h} + \left(\hat{c}'_\gamma(h)\right)^2 \sum_{n,m} n^2 K(n+m) e^{-\hat{c}_\gamma(h)n - (\mathrm{G}(h)-\hat{\lambda}_2(h))(m-\gamma n)}}{\sum_{n,m} nK(n+m) e^{-\hat{c}_\gamma(h)n - \left(\mathrm{G}(h)-\hat{\lambda}_2(h)\right)(m-\gamma n)}}$$
$$+ \frac{\hat{c}'_\gamma(h)\left(\mathrm{G}'(h) - \hat{\lambda}'_2(h)\right) \sum_{n,m} n(m-\gamma n) K(n+m) e^{-\hat{c}_\gamma(h)n - \left(\mathrm{G}(h)-\hat{\lambda}_2(h)\right)(m-\gamma n)}}{\sum_{n,m} nK(n+m) e^{-\hat{c}_\gamma(h)n - \left(\mathrm{G}(h)-\hat{\lambda}_2(h)\right)(m-\gamma n)}}. \quad (3.21)$$

We now observe that $\mathrm{N}''(h_0)$ coincides with the first term in the right-hand side of (3.21) evaluated at $h = h_0$. Therefore $\widetilde{\mathrm{F}}''_\gamma(\cdot)$ is continuous at h_0 if and only if the second term in the right-hand side of (3.21) vanishes at $h = h_0$. To clarify this issue we rewrite $\mathrm{G}'(\cdot) - \hat{\lambda}'_2(\cdot)$ by exploiting the fact that by differentiating (3.17) with respect to h we get

$$\mathrm{G}'(h) - \hat{\lambda}'_2(h)$$
$$= -\hat{c}'_\gamma(h) \frac{\sum_{n,m} n(m-\gamma n) K(n+m) \exp\left(-\hat{c}_\gamma(h)n - \left(\mathrm{G}(h) - \hat{\lambda}_2(h)\right)(m-\gamma n)\right)}{\sum_{n,m} (m-\gamma n)^2 K(n+m) \exp\left(-\hat{c}_\gamma(h)n - \left(\mathrm{G}(h) - \hat{\lambda}_2(h)\right)(m-\gamma n)\right)}. \quad (3.22)$$

Therefore going back to (3.19)–(3.21), we see that if $\sum_m m^2 K(m) < \infty$ then

$$\hat{c}''_\gamma(h_0) - \mathrm{N}''(h_0) = -\left(\widetilde{\mathrm{F}}'_\gamma(h_0)\right)^3 \frac{\left(\sum_{n,m} n(m-\gamma n)K(n+m)\exp\left(h_0 - n\mathrm{N}(h_0)\right)\right)^2}{\sum_{n,m}(m-\gamma n)^2 K(n+m)\exp\left(h_0 - n\mathrm{N}(h_0)\right)}, \quad (3.23)$$

and if $\sum_m m^2 K(m) = \infty$ then $\hat{c}''_\gamma(h_0) - \mathrm{N}''(h_0) = 0$. So, if $\sum_m m^2 K(m) = \infty$ then $\widetilde{\mathrm{F}}''_\gamma(\cdot)$ is continuous at h_0. If $\sum_m m^2 K(m) < \infty$ then (3.23) *generically* tells us that $\widetilde{\mathrm{F}}''_\gamma(\cdot)$ has a jump discontinuity at h_0, but the jump is zero if $\sum_{n,m} n(m-\gamma n)K(n+m)e^{h_0 - n\mathrm{N}(h_0)} = 0$ and, by (3.14), this is equivalent to $\gamma'_c(h_0) = 0$. The proof of Proposition 3.4 is therefore complete. □

Remark 3.5. A sharper analysis of the singularity at $h_0 \in \mathcal{I}_\mathrm{b}$ is possible and one sees that the closer α is to zero the more the transition is regular. The general analysis however is cumbersome due also to the fact that the transition can gain regularity from cancellations that may appear and that depend on fine details: we have already found an instance of this when in the proof of Proposition 3.4 we have seen that if $\sum_m m^2 K(m) < \infty$ then the second derivative of the free energy has a jump at h_0 unless $\gamma'_c(h_0) = 0$. These cancellations are at the origin of the difficulties in resolving the issue in item (2) of the list right before Proposition 3.4.

3.3. Free energy analysis of the delocalization transition

We complete now the proof of Theorem 1.2 by studying the asymptotic behavior near h_c of $\widetilde{\mathrm{F}}_\gamma(h)$: we will show in Section 4 that $\widetilde{\mathrm{F}}_\gamma(h) = \mathrm{F}_\gamma(h)$.

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

We start by observing that, by Proposition 3.1, (3.1) reduces to study the critical behavior of $\mathrm{G}(h)$ and $D_h(1,\gamma)$. In the case $\gamma = 1$, since $D_h(1,1) = 0$, $\mathrm{F}(h) = 2\mathrm{G}(h)$ and, since we have seen that the only singularity of $\mathrm{G}(\cdot)$ is at the origin, Theorem 1.2 reduces in this case to:

Lemma 3.6. *For $\alpha > 0$ we have*

$$\mathrm{G}(h) \overset{h\searrow 0}{\sim} \begin{cases} \frac{1}{2}ch & \text{if } \sum_n n^2 K(n) < \infty, \\ \frac{1}{2}L_\alpha(h)h^{1/\alpha} & \text{if } \sum_n n^2 K(n) = \infty, \end{cases} \tag{3.24}$$

where c is the same as for (1.15) and $L_\alpha(\cdot)$ is a slowly varying function. For $\alpha = 0$, $\mathrm{G}(h)$ vanishes faster than any power of h.

Implicit expressions for $L_\alpha(\cdot)$ as well as $\mathrm{G}(\cdot)$ in terms of the inverse of a suitable slowly varying function in the case $\alpha = 0$ can be found in the proof.

Proof. Actually since (1.18) can be written as

$$\sum_{n\geq 2}(n-1)K(n)\exp(h - n\mathrm{G}(h)) = 1, \tag{3.25}$$

we remark that $\mathrm{G}(h)$ is the free energy of the pinning model based on a one-dimensional renewal process and the proof is therefore just a revisitation of [16, Theorem 2.1]. We give in any case a substantial part of the arguments here.

If $\sum_n n^2 K(n) < \infty$, by (DOM) we have that $\sum_{n\geq 2}(n-1)K(n)\,(1-\exp(-n\mathrm{G}(h))) \sim \mathrm{G}(h)\sum_{n\geq 2} n(n-1)K(n)$ as $h \searrow 0$. Therefore by (3.25)

$$\mathrm{G}(h)\sum_{n\geq 2} n(n-1)K(n) \sim 1 - \exp(-h) \sim h, \tag{3.26}$$

and this of course proves the statement (1.15) for $\widetilde{\mathrm{F}}(h)$.

If $\sum_n n^2 K(n) = \infty$ and $\alpha \in (0,1)$, by (3.25) and by Riemann sum approximation, one has

$$\begin{aligned} 1 - \exp(-h) &= \sum_{n\geq 2}(n-1)K(n)\,(1-\exp(-n\mathrm{G}(h))) \\ &\overset{h\searrow 0}{\sim} (\mathrm{G}(h))^\alpha\, L\left(\frac{1}{\mathrm{G}(h)}\right)\int_0^\infty \frac{1-\exp(-x)}{x^{1+\alpha}}\,\mathrm{d}x, \end{aligned} \tag{3.27}$$

and therefore

$$\frac{1}{\alpha}\Gamma(1-\alpha)\,(\mathrm{G}(h))^\alpha\, L\left(1/\mathrm{G}(h)\right) \overset{h\searrow 0}{\sim} h, \tag{3.28}$$

and $\mathrm{G}(h) \sim \frac{1}{2}\, L_\alpha(h)\, h^{1/\alpha}$ where $L_\alpha(h) = 2\,(\alpha/\Gamma(1-\alpha))^{1/\alpha}\, h^{-1/\alpha}\hat{L}_\alpha(h)$ and $\hat{L}_\alpha(\cdot)$ is asymptotically equivalent to the inverse of $x \longmapsto x^\alpha L(1/x)$.

For the case $\alpha = 1$ we restart with the right-hand side of the first equality in (3.27) which equals, up to an additive term $O(\mathrm{G})$ for $(\mathrm{G} \searrow 0)$, to

$$\begin{aligned}&\sum_{n\ge 2}\frac{L(n)}{n^2}\left(1-\exp(-n\mathrm{G})\right)\\ &\quad=\sum_{n=2}^{\lfloor \varepsilon/\mathrm{G}\rfloor}\frac{L(n)}{n^2}\left(1-\exp(-n\mathrm{G})\right)+\sum_{n>\lfloor \varepsilon/\mathrm{G}\rfloor}\frac{L(n)}{n^2}\left(1-\exp(-n\mathrm{G})\right),\end{aligned} \tag{3.29}$$

with $\varepsilon > 0$. By performing a Riemann sum approximation we see that the second term in the right-hand side is $O(\mathrm{G}L(1/\mathrm{G}))$. For the first one instead we use that, by Taylor formula, for every $\delta > 0$ there exists $\varepsilon > 0$ such that

$$(1-\delta)\mathrm{G}\sum_{n=2}^{\lfloor \varepsilon/\mathrm{G}\rfloor}\frac{L(n)}{n}\le\sum_{n=2}^{\lfloor \varepsilon/\mathrm{G}\rfloor}\frac{L(n)}{n^2}\left(1-\exp(-n\mathrm{G})\right)\le(1+\delta)\mathrm{G}\sum_{n=2}^{\lfloor \varepsilon/\mathrm{G}\rfloor}\frac{L(n)}{n}, \tag{3.30}$$

but the sum on the leftmost and rightmost term is asymptotic to $\int_1^{1/\mathrm{G}}(L(t)/t)\,\mathrm{d}t =: \check{L}(1/\mathrm{G})$, which is slowly varying and $L(x) = o(\check{L}(x))$ for $x \to \infty$ [2, Proposition 1.5.9a]. At this point we go back to the first equality in (3.27) and we have

$$\mathrm{G}(h)\check{L}\left(1/(\mathrm{G}(h))\right) \overset{h\searrow 0}{\sim} h. \tag{3.31}$$

Since the right-hand side of the first equality in (3.27) is an increasing function of G, (3.31) can be asymptotically inverted and the case $\alpha = 1$ is complete.

For the case $\alpha = 0$ the computation is similar. Again we replace the term $(n-1)$ with n in the right-hand side of the first equality in (3.27): the error is $O(\mathrm{G})$. Then we are left with $\sum_{n\ge 2}(L(n)/n)(1-\exp(-n\mathrm{G}))$ and we split the sum into n smaller and larger than $1/(\varepsilon\mathrm{G})$. The first sum can be treated by Riemann approximation yielding a term $O(L(1/\mathrm{G}))$. The other term instead is asymptotic to $\check{L}(x) := \int_x^\infty (L(t)/t)$, which is slowly varying and $L(x) = o(\check{L}(x))$ [2, Proposition 1.5.9b]. The right-hand side of the first equality in (3.27) is an increasing function of G that vanishes as G tends to zero. So $\check{L}(x)$ can be chosen decreasing (to zero) and the slowly varying property implies that $\check{L}(x) \ge x^{-\varepsilon}$ for every $\varepsilon > 0$ and every x sufficiently large. Hence $\check{L}(1/\mathrm{G}(h)) \ge \mathrm{G}(h)^\varepsilon$ for h sufficiently small and $\check{L}(1/\mathrm{G}(h)) \sim h$ readily implies that $\mathrm{G}(h) = O(h^{1/\varepsilon})$ which is what we wanted to prove. The proof of Lemma 3.6 is therefore complete. □

Proof of Lemma 1.4. The analyticity of $\gamma_c(\cdot)$ has been proven in Proposition 3.2 and the second statement of (1.25) is trivial. Let us prove the first statement of (1.25) (keeping in mind that $\alpha > 0$).

Recall the definition of $\gamma_c(\cdot)$ from (1.23). If $\sum_{n\ge 1} n^2K(n) < \infty$ then it is easy to see that

$$\gamma_c(0) = \frac{\sum_{n,m\ge 1} mK(n+m)}{\sum_{n,m\ge 1} nK(n+m)} = 1. \tag{3.32}$$

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

Now if $\sum_{n\geq 1} n^2 K(n) = \infty$ and $\alpha \in (0,1)$, set $\mathrm{G}(h) - \overline{\lambda}_1(h) = x = o(1)$ (as $h \searrow 0$) and remark that

$$\sum_{n,m} nK(n+m)\exp(-xn) = \sum_{t\geq 2} K(t)\, \frac{e^x(1-e^{-xt}) + te^{x(1-t)}(1-e^x)}{(e^x-1)^2}. \tag{3.33}$$

By Riemann sum approximation, the right-hand side of (3.33) is equivalent to (as $x \searrow 0$)

$$x^{\alpha-1} L(1/x) \int_0^\infty \frac{1-\exp(-y)(1+y)}{y^{2+\alpha}}\,\mathrm{d}y, \tag{3.34}$$

therefore

$$\sum_{n,m\geq 1} nK(n+m)\exp(-xn) \stackrel{x\searrow 0}{\sim} x^{\alpha-1}\frac{\Gamma(1-\alpha)}{1+\alpha} L(1/x)\,. \tag{3.35}$$

Repeating the same argument leading to (3.35), we see that

$$\sum_{n,m\geq 1} mK(n+m)\exp(-xn) \stackrel{x\searrow 0}{\sim} x^{\alpha-1}\Gamma(-\alpha-1) L(1/x)\,. \tag{3.36}$$

By (1.23), (3.35) and (3.36), we get

$$\gamma_c(h) \stackrel{h\searrow 0}{\sim} \frac{1}{\alpha}. \tag{3.37}$$

For the case $\alpha = 1$, it suffices to show that

$$\sum_{n,m} nK(n+m)\exp(-xn) \stackrel{x\searrow 0}{\sim} \sum_{n,m} mK(n+m)\exp(-xn). \tag{3.38}$$

In fact, both terms are asymptotic to $\check{L}(x) = \int_1^{1/x} L(t)/t$, slowly varying by [2, Proposition 1.5.9b] and diverging at ∞ because $\sum_{n,m} nK(n+m) = \infty$. This can be seen by restarting from (3.33): the left-hand side of (3.38) is equal to

$$\begin{aligned}
&\frac{e^x}{(e^x-1)^2}\sum_{t\geq 2}\frac{L(t)}{t^3}\left((1-e^{-xt}) + te^{-xt}(1-e^x)\right)\\
&\quad = \frac{1+O(x)}{x^2}\sum_{t\geq 2}\frac{L(t)}{t^3}\left((1-e^{-xt}) - txe^{-xt}\right) + O(1).
\end{aligned} \tag{3.39}$$

A Riemann sum approximation shows that if the sum is limited to $t > \varepsilon/x$, for an arbitrary $\varepsilon > 0$, yields an $O(1)$ contribution. For the term that is left we use that, by Taylor formula, for every $\delta > 0$ one finds a $\varepsilon > 0$ such that

$$\frac{1}{x^2}\sum_{t=2}^{\lfloor\varepsilon/x\rfloor}\frac{L(t)}{t^3}\left((1-e^{-xt}) - txe^{-xt}\right) \leq \frac{1+\delta}{2x^2}\sum_{t=2}^{\lfloor\varepsilon/x\rfloor}\frac{L(t)}{t^3}(tx)^2 \stackrel{x\searrow 0}{\sim} \frac{(1+\delta)}{2}\sum_{t=2}^{\lfloor\varepsilon/x\rfloor}\frac{L(t)}{t}, \tag{3.40}$$

and analogous lower bound with $1-\delta$. Since $\delta > 0$ is arbitrary, the claimed asymptotic behavior for the left-hand side of (3.38) is established. The computation for the right-hand side is very similar and left to the reader. Therefore the proof is complete. □

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

Recall now Proposition 3.1 and in particular the fact that $D_h(1,\gamma)$ is the result of an optimization problem, cf. (3.3), over the set B_h (cf. (3.4)). As widely used and discussed in and right after Proposition 3.2, the maximum can be achieved in the interior of B_h (Cramér regime) or at the boundary (out of Cramér regime).

Proposition 3.7. *Choose $\gamma > 1$. If $\sum_n n^2 K(n) < \infty$ we have that for h small the system is outside of the Cramér regime and*

$$\widetilde{\mathrm{F}}_\gamma(h) \overset{h\searrow 0}{\sim} \widetilde{\mathrm{F}}_1(h) = 2\mathrm{G}(h). \tag{3.41}$$

If instead $\sum_n n^2 K(n) = \infty$ for h small the system is in the Cramér regime if $\gamma < 1/\alpha$ and there exists $c_{\alpha,\gamma} \in (1, \frac{1}{2}((1+\alpha)^{1/\alpha} \wedge (1+\gamma)))$ such that

$$\widetilde{\mathrm{F}}_\gamma(h) \sim c_{\alpha,\gamma}\widetilde{\mathrm{F}}_1(h). \tag{3.42}$$

If $\gamma > 1/\alpha$, the system is outside of this regime and

$$\widetilde{\mathrm{F}}_\gamma(h) \sim \frac{(1+\alpha)^{1/\alpha}}{2}\widetilde{\mathrm{F}}_1(h). \tag{3.43}$$

If $\alpha = 0$ the system is in the Cramér regime for every γ and $\widetilde{\mathrm{F}}_\gamma(h) = O(h^{1/\varepsilon})$ for every $\varepsilon > 0$. The asymptotic behavior of $\mathrm{G}(h)$ is given in Lemma 3.6.

Proof. Since $\gamma > 1$ we have that $D_h(1,\gamma) > 0$. We recall Proposition 3.1 and for $(\lambda_1, \lambda_2) \in B_h$, we make the change of variables $\mathrm{G}(h) - \lambda_1 = a_1\,\mathrm{G}(h)$, so $a_1 \geq 1$, and $\mathrm{G}(h) - \lambda_2 = a_2\,\mathrm{G}(h)$, so $a_2 \in [0,1]$. With these new variables we have

$$D_h(1,\gamma) = \mathrm{G}(h) \max_{a\in\mathcal{B}_h} (1 - a_1 + \gamma(1-a_2)) = \mathrm{G}(h)\left(1 + \gamma - \min_{a\in\mathcal{B}_h} (a_1 + \gamma a_2)\right), \tag{3.44}$$

hence

$$\widetilde{\mathrm{F}}_\gamma(h) = \min_{a\in\mathcal{B}_h} (a_1 + \gamma a_2)\,\mathrm{G}(h), \tag{3.45}$$

with

$$\mathcal{B}_h := \left\{ a : 1 \leq a_1 \leq 1 - \frac{\overline{\lambda}_1(h)}{\mathrm{G}(h)},\ 0 \leq a_2 \leq 1, \sum_{n,m} K(n+m) e^{h-(a_1 n + a_2 m)\mathrm{G}(h)} = 1 \right\}. \tag{3.46}$$

We set $\Psi_h(a,\mathrm{G}) := \sum_{n,m} K(n+m)\exp(-(a_1 n + a_2 m)\mathrm{G})$ and for every $a \in \mathcal{B}_h$, we have that $\Psi_h(a,\mathrm{G}(h)) = \exp(-h)$, which we will use asymptotically as

$$1 - \Psi_h(a,\mathrm{G}(h)) \overset{h\searrow 0}{\sim} h. \tag{3.47}$$

If $\sum_n n^2 K(n) < \infty$, since $\gamma_c(0) = 1$ (from Lemma 1.4) and $\gamma > 1$, the system is outside of the Cramér regime for h small (recall that the system is in the Cramér regime if and only if $\gamma_c(h) > \gamma$) and in terms of the new variable it means that the minimizer satisfies $a_2 = 0$ (see the proof of Proposition 3.2). Set $\overline{a} = (\overline{a}_1, 0)$ with $\overline{a}_1\mathrm{G}(h) = \mathrm{G}(h) - \overline{\lambda}_1(h)$. By (DOM) we have

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

$$1 - \Psi_h(\overline{a}, \mathrm{G}(h)) = 1 - \sum_{n,m\geq 1} K(n+m)\exp\left(-\overline{a}_1 n \mathrm{G}(h)\right) \overset{h\searrow 0}{\sim} \frac{1}{2}\overline{a}_1 \mathrm{G}(h) \sum_{n\geq 2} n(n-1)K(n), \tag{3.48}$$

which, together with (3.26) and (3.47), implies $\overline{a}_1 \overset{h\searrow 0}{\sim} 2$. Therefore, by (3.45), $\widetilde{\mathrm{F}}_\gamma(h) \sim 2\mathrm{G}(h)$.

When $\sum_n n^2 K(n) = \infty$, we first treat the case $\alpha \in (0, 1)$: we have seen in Lemma 1.4 that in this case $\gamma_c(0) = 1/\alpha$ which implies that we are in the Cramér regime for h small if $\gamma < 1/\alpha$ and outside if $\gamma > 1/\alpha$ (and $a_2 = 0$). But let us consider the two regimes together for now and observe that for $a_2 \in [0, 1]$, $a_1 \geq 1$ and $a_1 \neq a_2$

$$\begin{aligned} &1 - \Psi_h(a, \mathrm{G}) \\ &= \sum_{t\geq 2} K(t)\left((t-1) - \left(\frac{\exp\left(-ta_1\mathrm{G}(h)\right) - \exp\left(-ta_2\mathrm{G}(h) - (a_1-a_2)\mathrm{G}(h)\right)}{\exp\left(-(a_1-a_2)\mathrm{G}(h)\right) - 1}\right)\right), \end{aligned} \tag{3.49}$$

and by Riemann sum approximation, the right-hand side is equivalent to (as $h \searrow 0$)

$$\mathrm{G}(h)^\alpha L\left(1/\mathrm{G}(h)\right) \int_0^\infty \frac{x - (\exp(-a_1x) - \exp(-a_2x))/(a_2-a_1)}{x^{2+\alpha}}\,\mathrm{d}x. \tag{3.50}$$

The integral can be made explicit:

$$1 - \Psi_h(a, \mathrm{G}) \sim b_\alpha(a)\Gamma(-\alpha-1)\,\mathrm{G}(h)^\alpha L\left(1/\mathrm{G}(h)\right), \tag{3.51}$$

with

$$b_\alpha(a) := \frac{\left({a_1}^{1+\alpha} - {a_2}^{1+\alpha}\right)}{(a_1-a_2)} = (1+\alpha)\int_0^1 \left((1-t)a_2 + ta_1\right)^\alpha \,\mathrm{d}t. \tag{3.52}$$

By (3.28) and (3.47) we get that for $a \in \mathcal{B}_h$ and $a_1 \neq a_2$

$$b_\alpha(a) \sim 1+\alpha, \tag{3.53}$$

but the rightmost term in (3.52) shows that the singularity in $a_1 = a_2$ is removable and one directly verifies that (3.51), and therefore (3.53), holds also for $a_1 = a_2(=1)$.

A number of considerations are in order:

(1) By recalling the convexity arguments used in Section 3.1, we directly have that the constraint in $\mathcal{B}_h$ can be rewritten as $a_2 = \widetilde{a}_{2,h}(a_1)$, with $\widetilde{a}_2(\cdot)$ a decreasing convex function (this is the curve appearing in Fig. 3, up to the affine change of variable we performed). It will be more practical at this stage to write rather $a_1 = \widetilde{a}_{1,h}(a_2)$, and $\widetilde{a}_{1,h} : [0, 1] \to [1, \infty)$ is also convex and decreasing with $\widetilde{a}_{1,h}(0) = 1 - \overline{\lambda}_1(h)/\mathrm{G}(h) \sim (1+\alpha)^{1/\alpha}$ and $\widetilde{a}_{1,h}(1) = 1$ where we have used (3.53). The choice in favor of $\widetilde{a}_{1,h}(\cdot)$ over $\widetilde{a}_{2,h}(\cdot)$ is because we prefer having a h dependence in the image rather than in the domain.

(2) Since $1 - \Psi_h(\cdot, \mathrm{G})$ is concave, also $b_\alpha(\cdot)$ is (this can also be verified directly) and the constraint in $\mathcal{B}_h$, namely (3.53), becomes in the limit $b_\alpha(a) = 1+\alpha$, with $a_1 \in [1, (1+\alpha)^{1/\alpha}]$ and $a_2 \in [0, 1]$. Such a constraint can be expressed (like in Section 3.1) as $a_1 = \widetilde{a}_1(a_2)$, where $\widetilde{a}_1 : [0, 1] \to [1, (1+\alpha)^{1/\alpha}]$ is a convex decreasing function. This function is smooth (in fact, analytic, by the Implicit Function Theorem [22]) in the interior of the domain and one directly computes $\widetilde{a}_1'(0^+) = -1/\alpha$ and $\widetilde{a}_1'(1^-) = -1$ by expanding (3.52) to second order, which actually coincide with the limits as $h \searrow 0$ respectively of $\widetilde{a}_{1,h}'(0^+)$ (this is precisely (3.37)) and $\widetilde{a}_{1,h}'(1^-) = -1$ (in fact: $\widetilde{a}_{1,h}'(1^-) = -1$ for every $h > 0$ by symmetry).

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

(3) But (3.53), that is the convergence of the constraint function for $h \searrow 0$, implies $\lim_h \widetilde{a}_{1,h}(a_2) = \widetilde{a}_1(a_2)$ for every $a_2 \in [0, 1]$, as well as the convergence of the derivative for $a_2 \in (0, 1)$, because we are dealing with a sequence of convex functions and because the limit is differentiable. Note that we have already pointed out that there is convergence of the derivatives also at the endpoints and that, by analyticity, $\widetilde{a}_1(\cdot)$ is strictly convex. This allows to conclude that the unique minimizer $\hat{a}(h)$ for (3.45) converges to the unique minimizer of $\min_{a\in\mathcal{B}}(a_1 + \gamma a_2)$, with

$$\mathcal{B} := \left\{a : 1 \le a_1 \le (1+\alpha)^{1/\alpha},\ 0 \le a_2 \le 1, b_\alpha(a) = 1+\alpha\right\}. \tag{3.54}$$

The limit problem, like the approaching ones, can of course be rewritten in a one-dimensional form using $\widetilde{a}_1(\cdot)$ and $\widetilde{a}_{1,h}(\cdot)$.

At this point we can treat separately the non Cramér case, in which $a_2 = 0$ (this happens when $\gamma > 1/\alpha$, but also when $\gamma = 1/\alpha$ by a limit procedure), so we have $a_1 = (1+\alpha)^{1/\alpha}$ and $\widetilde{\mathrm{F}}_\gamma(h) \sim (1+\alpha)^{1/\alpha}\mathrm{G}(h)$.

If $\gamma < 1/\alpha$ instead, $\widetilde{\mathrm{F}}_\gamma(h) \sim \min_{a\in\mathcal{B}}(a_1 + \gamma a_2)\mathrm{G}(h)$, so $c_{\alpha,\gamma} = \frac{1}{2}\min_{a\in\mathcal{B}}(a_1 + \gamma a_2)$ and (3.42) is proven. Since $a_1(\cdot)$ is strictly convex and decreases from $a_1(0) = (1+\alpha)^{1/\alpha}$ to $a_1(1) = 1$, one can easily see (just evaluate $a_1 + \gamma a_2$ at $a = (1, 1)$ and $a = ((1+\alpha)^{1/\alpha}, 0)$) that $c_{\alpha,\gamma} \in (1, \frac{1}{2}((1+\alpha)^{1/\alpha} \wedge (1+\gamma)))$.

Remark 3.8. For $\alpha \in (0, 1)$ we can use the Lagrange multiplier method to solve the limit optimization problem $\min_{a\in\mathcal{B}}(a_1 + \gamma a_2)$. With s as multiplier we have

$$\begin{cases} s\,(1+\alpha)\,(1 - {a_1}^\alpha) = 1, \\ s\,(1+\alpha)\,({a_2}^\alpha - 1) = \gamma, \end{cases} \tag{3.55}$$

which implies that

$${a_2}^\alpha = 1 - \gamma({a_1}^\alpha - 1). \tag{3.56}$$

In particular, and consistently with what precedes, if $\gamma > 1/({a_1}^\alpha - 1) \ge 1/\alpha$ no solution is found and we are out of the Cramér regime, and $(a_1, a_2) = ((1+\alpha)^{1/\alpha}, 0)$. If $\gamma < 1/\alpha$ instead a solution is found and in fact we are in the Cramér regime.

For the case $\alpha = 1$, from Lemma 1.4, we have $\gamma_c(0) = 1$, which implies that for every $\gamma > 1$, we are outside of the Cramér regime for h small and $a = \overline{a} = (\overline{a}_1, 0)$. Observe that

$$1 - \Psi_h(\overline{a}, \mathrm{G}) = \sum_{t\ge 2}\frac{L(t)}{t^3}\left((t-1) - \frac{\exp(-\overline{a}_1\mathrm{G}(h)(t-1)) - 1}{1 - \exp(\overline{a}_1\mathrm{G}(h))}\right), \tag{3.57}$$

and follow the same procedure adopted for the case $\alpha = 1$ in Lemma 3.6: split the sum in (3.57) into t larger than $\varepsilon/(\overline{a}_1\mathrm{G})$, for an arbitrary $\varepsilon > 0$ (which yields $O(1)$ contribution) and to t smaller than $\varepsilon/(\overline{a}_1\mathrm{G})$ to obtain

$$1 - \Psi_h(\overline{a}, \mathrm{G}) \overset{h\searrow 0}{\sim} \frac{1}{2}\overline{a}_1\mathrm{G}(h)\widetilde{L}\left(1/(\overline{a}_1\mathrm{G}(h))\right). \tag{3.58}$$

We know that the left-hand side is equivalent to h from (3.47), and using the property of slowly varying functions (A.1) ($\widetilde{L}\,(1/(\overline{a}_1\mathrm{G}(h))) \sim \widetilde{L}\,(1/(\mathrm{G}(h)))$), then by (3.31) we get $\overline{a}_1 \sim 2$. Therefore by (3.45), we obtain $\widetilde{\mathrm{F}}_\gamma(h) \sim 2\mathrm{G}(h)$.

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

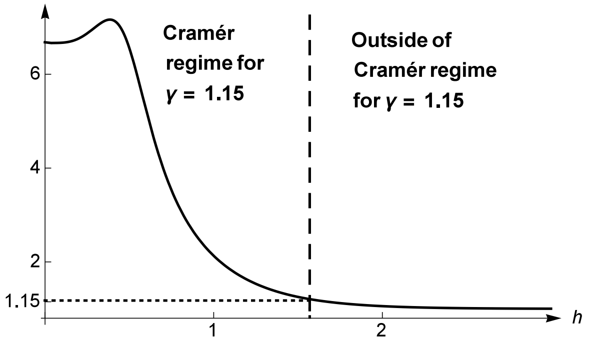

Fig. 4. A representation of the function $\gamma_c(\cdot)$ for the distribution $K_h(\cdot)$ with $c = 2.15$, $E_b = 6$ and $E_l = 3$. The horizontal dashed line corresponds to the value of $\gamma = 1.15$ and the vertical one to the critical point $h_{c,\gamma} \simeq 1.676$. Computing $\gamma_c(h)$ is immediate from (1.23) once $\mathrm{G}(h)$ and $\overline{\lambda}_1(h)$ are known: we have first obtained $\mathrm{G}(h)$ by solving numerically $\sum_{n,m} \widetilde{K}_h(n, m) = 1$ and then we have solved numerically $q_h(\overline{\lambda}_1, \mathrm{G}) = 1$ for $\overline{\lambda}_1(h)$.

For the case $\alpha = 0$, we do not strive for the sharp prefactor, since we did not go for the sharp behavior of $\mathrm{G}(h)$, cf. Lemma 3.6. We just use (3.45) and observe that $\min_{a \in \mathcal{B}_h}(a_1 + \gamma a_2)$ is bounded (because the supremum is over a set that is bounded uniformly in h, with h in a right neighborhood of zero). Hence in this case $\widetilde{\mathrm{F}}_\gamma(h) = O(h^{1/\varepsilon})$ for every $\varepsilon > 0$. The proof is therefore complete. □

3.4. The bio-physics model and other examples

In this section we make $\gamma_c(\cdot)$ explicit for some choices of $K(\cdot)$, starting with [24].

The bio-physics model

We refer to Section 1.2. We have seen that the geometric constant s is irrelevant then we take $s = 1$ and $B(l) = 1/l^c$. Recall that $E_b > 0$ is the binding energy and $E_l > 0$ is the loop initiation cost. Set

$$K_h(n) := \frac{c_K}{n^c} \exp\left(h(E_b - E_l \mathbf{1}_{n>2}) - n\mathrm{G}(h)\right), \tag{3.59}$$

with $c_K = \sum_{n,m\geq 1} 1/(n+m)^c = \sum_{n\geq 2}(n-1)/n^c$ is the normalization constant and $\mathrm{G}(h)$ is the only solution to $\sum_{n,m} K_h(n+m) = 1$.

In Fig. 4, a plot of $\gamma_c(\cdot)$ for values of E_B, E_l, c and γ chosen in [24], shows that the system exhibits a unique transition at $h_{c,\gamma} \simeq 1.676$ (the vertical dashed line). For $\gamma = 1.15$, observe that the system is in the Cramér regime for $h < h_{c,\gamma}$ and outside of the Cramér regime if $h \geq h_{c,\gamma}$. In [24] the transition at $h_{c,\gamma}$ is described as between a phase with microscopic free strands at the end of the polymer and macroscopic free strands.

A more explicitly solvable class: basic case

We exploit the inter-arrival law used in [17] in the one-dimensional case. Recall that Euler's Gamma function $\Gamma(z) := \int_0^\infty t^{z-1}\exp(-t)\,\mathrm{d}t$ defines an analytic function on $\{z \in \mathbb{C} : Re(z) > 0\}$ and that for $z \in \mathbb{C} \setminus \{0, -1, -2, \ldots\}$, it verifies $\Gamma(z+1) = z\Gamma(z)$ (therefore $\Gamma(n+1) = n!$ for $n \in \mathbb{N}$). Recall also that the Taylor coefficients of the function $(1-z)^\alpha$ for $|z| < 1$ and

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

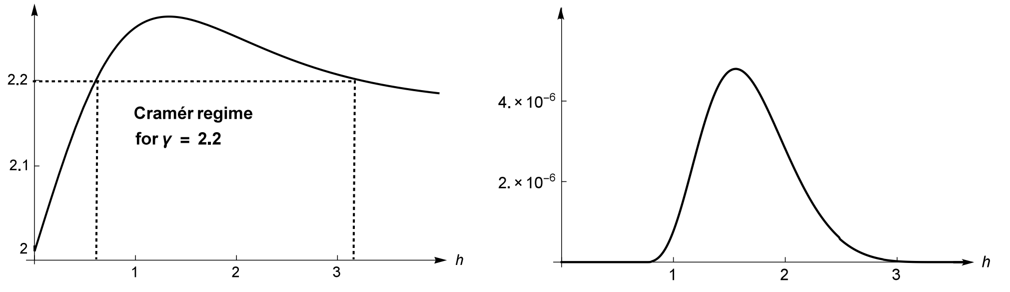

Fig. 5. On the left the function $\gamma_c(\cdot)$ for the inter-arrival distribution $K_2(\cdot)$, with $\alpha = 0.5$ and $\kappa = 0.01$. The horizontal dashed line corresponds to the value of $\gamma = 2.2$. There are two values of h (vertical dashed lines) such that $\gamma_c(h) = \gamma$ and therefore there are two transitions. On the right we plot the difference $\mathrm{N}(\cdot) - \widetilde{\mathrm{F}}_\gamma(\cdot)$, that is $\mathrm{N}(\cdot) - \hat{c}_\gamma(\cdot)$ (recall (3.11) and (3.12)), which makes clear the presence of the two transitions. The resolution of the graph does not allow to appreciate the positivity of such a difference for example at $h = 3$ where it is about 6×10^{-9}.

$\alpha \in \mathbb{R} \setminus \{0, 1, 2, \ldots\}$, is known exactly

$$\sum_{n \geq 0} \frac{\Gamma(n-\alpha)}{n!} z^n = \Gamma(-\alpha)(1-z)^\alpha, \tag{3.60}$$

and asymptotically from Stirling's formula we have $\Gamma(n-\alpha)/n! \stackrel{n\to\infty}{\sim} 1/n^{1+\alpha}$. Note that the first terms of the series in (3.60) have alternating signs.

First, let us suppose that $\alpha \in (0, 1)$ and set

$$K_1(n) := \frac{\Gamma(n-\alpha-1)}{\Gamma(-\alpha-1)n!}, \quad \text{for } n \geq 2 \tag{3.61}$$

and from (3.60) we have that $\sum_{n\geq 2}(n-1)K_1(n) = 1$ and by (3.33) and the fact that $\sum_{n,m} nK(n+m)z^n = z\alpha(1-z)^{\alpha-1}$ and $\sum_{n,m} mK(n+m)z^n = z(1-z)^{\alpha-1}$ we obtain $\gamma_c(h) = 1/\alpha$. This implies that there is no transition in the localized phase and for every $h > 0$, the free energy is either – recall (3.12) – $\mathrm{N}(h)$ (if $\gamma \geq \gamma_c(h)$: out of the Cramér regime) or $\hat{c}_\gamma(h)$ (if $\gamma < \gamma_c(h)$, the Cramér regime, recall (3.11)).

A more explicitly solvable class: some generalizations of the basic case

More general explicit cases can be built by modifying $K_1(\cdot)$ on a finite number of sites. Let us choose now $K_2(2) = K_1(2)$, $K_2(3) = \kappa$ and normalize the rest

$$K_2(n) = c_\kappa K_1(n), \quad \text{for } n \geq 4 \tag{3.62}$$

with $c_\kappa = (1 - K_1(2) - 2\kappa) / (1 - K_1(2) - 2K_1(3))$ and $\kappa \in [0, 1)$. With this choice of the inter-arrival $K(\cdot)$ there are two transitions in the localized phase if and only if $\gamma > 2$, see Fig. 5: the system here is in the Cramér regime only for intermediate values of h.

We then present an example with $\alpha \in (1, 2)$: Proposition 3.7 shows that in this case for h small the system is outside the Cramér regime. We take $K_3(2) = \varrho$ and $K_3(n) = c_\varrho K_1(n)$ for $n \geq 3$ with $c_\varrho = (1-\varrho)/\left(\sum_{n\geq 3}(n-1)K_1(n)\right)$. A look at Fig. 6 shows that, if $\gamma < 2.27$, the system is outside of the Cramér regime when h is below a critical value $h_{c,\gamma}$. For larger values of γ the system is outside of the Cramér regime for every $h > 0$.

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

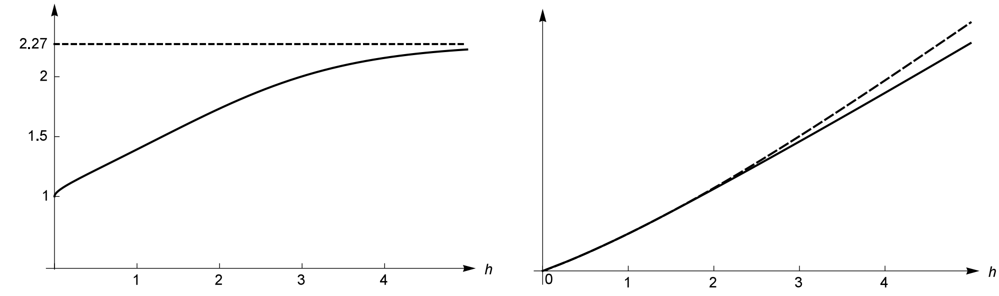

Fig. 6. For the inter-arrival distribution $K_3(\cdot)$, the function $\gamma_c(\cdot)$ (on the left) is strictly increasing (for $\alpha = 1.5, \varrho = 0.02$), this implies that if $\gamma_0 := \lim_{h\to\infty} \gamma_c(h) \le \gamma$ (dashed line represents the value of γ_0 computed using (1.25)), there is no transition in the system. On the right, we present the free energy $\widetilde{\mathrm{F}}_\gamma(h)$ and $\mathrm{N}(h)$ (dashed line, recall (3.12)) for $\gamma = 1.5$.

4. The free model

In this section we use the notation $\gamma_N := M/N$ and $\gamma'_N := M'/N'$. Recall that we assume $M \ge N$, but $M' \in \{0, \dots, M\}$ may be smaller than $N' \in \{0, \dots, N\}$. We use the short-cut of saying that γ is in the Cramer region if $(1, \gamma) \in E_h$, i.e. if $(1, \gamma)$ is in the Cramer region. Fig. 7 and its caption sum up properties of $D_h(1, \cdot)$ and of $\widetilde{\mathrm{F}}_\cdot(h)$ that will come handy in the remainder.

4.1. A preliminary result and free energy equivalence

We start with a preliminary important lemma on the constrained free energy: a minimal part of its strength will be used to prove, just below (Proposition 4.2), that this free energy coincides with the one of the free model. The full strength of this lemma is used in Section 4.2.

Lemma 4.1. *There exists $a(h) > 0$ such that for $\gamma'_N \ge \gamma_N$ we have*

$$\widetilde{\mathrm{F}}_{\gamma_N}(h) - \frac{N'}{N}\widetilde{\mathrm{F}}_{\gamma'_N}(h) \ge a(h)\left(1 - \frac{\gamma_N}{\gamma'_N}\right) + \frac{\gamma_N}{\gamma'_N}\widetilde{\mathrm{F}}_{\gamma'_N}(h)\left(1 - \frac{M'}{M}\right). \tag{4.1}$$

For $\gamma'_N \le \gamma_N$ we have

$$\widetilde{\mathrm{F}}_{\gamma_N}(h) - \frac{N'}{N}\widetilde{\mathrm{F}}_{\gamma'_N}(h) \ge \partial_\gamma \widetilde{\mathrm{F}}_\gamma(h)\big|_{\gamma=\gamma_N}(\gamma_N - \gamma'_N) + \widetilde{\mathrm{F}}_{\gamma'_N}(h)\left(1 - \frac{N'}{N}\right), \tag{4.2}$$

where $\partial_\gamma \widetilde{\mathrm{F}}_\gamma(h) \ge 0$ and for every $\varepsilon \in (0, \gamma_c(h))$ we have $\inf_{\gamma \le \gamma_c(h)-\varepsilon} \partial_\gamma \widetilde{\mathrm{F}}_\gamma(h) > 0$.

In particular for every ε as above and every $L > 0$ there exists $a_{\varepsilon,L}(h) > 0$ such that

$$\begin{aligned} &\widetilde{\mathrm{F}}_{\gamma_N}(h) - \frac{N'}{N}\widetilde{\mathrm{F}}_{\gamma'_N}(h) \\ &\quad \ge a_{\varepsilon,L}(h)\left(\left|\gamma_N - \gamma'_N\right| + \mathbf{1}_{\gamma'_N \ge \gamma_N}\gamma_N\left(1 - \frac{M'}{M}\right) + \mathbf{1}_{\gamma'_N < \gamma_N}\left(1 - \frac{N'}{N}\right)\right), \end{aligned} \tag{4.3}$$

for $\gamma_N \le \gamma_c(h) - \varepsilon$ and $\gamma'_N \in [1/L, L]$.

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

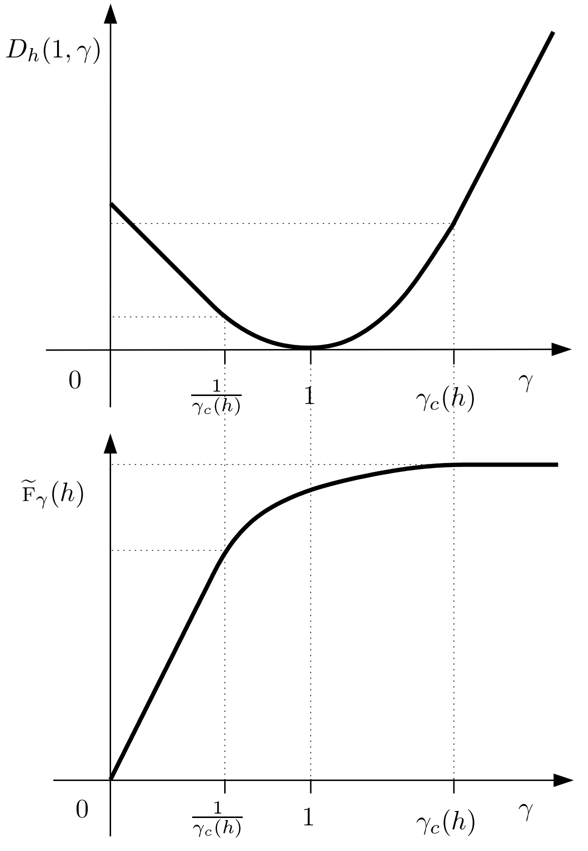

Fig. 7. We plot the Large Deviations functional $D_h(1, \gamma)$ and the free energy $\widetilde{\mathrm{F}}_\gamma(h)$, that coincides with $\mathrm{F}_\gamma(h)$, as functions of γ for a given $h > 0$. Relevant features are the convexity of the first and concavity of the second, which become strict in the Cramer region or interval $(1/\gamma_c(h), \gamma_c(h))$. Both functions are real analytic except at the boundary of the Cramer interval and they are affine functions outside of this interval. This and a number of other properties can be extracted from the variational expression (2.6) for $D_h(1, \gamma)$, the analysis in Section 3 and formula (3.1) for $\mathrm{F}_\gamma(h)$. In particular we have that $\partial_\gamma D_h(1, \gamma)$ is equal to $\mathrm{G}(h)$ for $\gamma > \gamma_c(h)$ – see (1.23) for an explicit expression of $\gamma_c(h)$ – and to $-\mathrm{G}(h)/\gamma_c(h) + D_h(1, \gamma_c(h))$ for $\gamma < 1/\gamma_c(h)$. Note also that the fact that the lower endpoint of Cramer interval is $1/\gamma_c(h)$ follows from the symmetry $D_h(v_1, v_2) = D_h(v_2, v_1)$ and the scaling behavior $D_h(cv) = cD_h(v)$ for $c > 0$ and $v = (v_1, v_2)$, for every v_1 and v_2 positive. The free energy becomes constant above $\gamma_c(h)$ and it is equal to $\mathrm{G}(h)+|\overline{\lambda}_1(h)|$ (recall (1.22)).

Proof. If $\gamma'_N \geq \gamma_N$ we write

$$\widetilde{\mathrm{F}}_{\gamma_N}(h) - \frac{N'}{N}\widetilde{\mathrm{F}}_{\gamma'_N}(h) = \left(\widetilde{\mathrm{F}}_{\gamma_N}(h) - \frac{\gamma_N}{\gamma'_N}\widetilde{\mathrm{F}}_{\gamma'_N}(h)\right) + \frac{\gamma_N}{\gamma'_N}\widetilde{\mathrm{F}}_{\gamma'_N}(h)\left(1 - \frac{M'}{M}\right), \tag{4.4}$$

and it suffices to focus on the first term between parentheses in the right-hand side and by using $\gamma = \gamma_N$ and $\gamma' = \gamma'_N$ to keep expressions short, since $\gamma \mapsto \widetilde{\mathrm{F}}_\gamma(h)$ is concave we have $\widetilde{\mathrm{F}}_{\gamma'}(h) \leq \widetilde{\mathrm{F}}_\gamma(h) + \partial_\gamma\widetilde{\mathrm{F}}_\gamma(h)(\gamma' - \gamma)$ so

$$\widetilde{\mathrm{F}}_\gamma(h) - \frac{\gamma}{\gamma'}\widetilde{\mathrm{F}}_{\gamma'}(h) \geq \left(1 - \frac{\gamma}{\gamma'}\right)\left(\widetilde{\mathrm{F}}_\gamma(h) - \gamma\partial_\gamma\widetilde{\mathrm{F}}_\gamma(h)\right), \tag{4.5}$$

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

and the right-hand side is non-negative because $\gamma \mapsto \widetilde{\mathrm{F}}_\gamma(h)$ is concave and $\widetilde{\mathrm{F}}_0(h) = 0$, so $\widetilde{\mathrm{F}}_\gamma(h) - \gamma \partial_\gamma \widetilde{\mathrm{F}}_\gamma(h)$ is non decreasing in γ and it is even increasing if $\gamma \in (1/\gamma_c(h), \gamma_c(h))$. This implies in particular that

$$a(h) := \inf_{\gamma \geq 1} \left(\widetilde{\mathrm{F}}_\gamma(h) - \gamma \partial_\gamma \widetilde{\mathrm{F}}_\gamma(h)\right) = \widetilde{\mathrm{F}}_1(h) - \partial_\gamma \widetilde{\mathrm{F}}_\gamma(h)\big|_{\gamma=1} > 0, \tag{4.6}$$

and (4.1) is proven.

Let us turn then to $\gamma'_N \leq \gamma_N$ and (4.2). This time we write

$$\widetilde{\mathrm{F}}_{\gamma_N}(h) - \frac{N'}{N}\widetilde{\mathrm{F}}_{\gamma'_N}(h) = \left(\widetilde{\mathrm{F}}_{\gamma_N}(h) - \widetilde{\mathrm{F}}_{\gamma'_N}(h)\right) + \left(1 - \frac{N'}{N}\right)\widetilde{\mathrm{F}}_{\gamma'_N}(h), \tag{4.7}$$

so it suffices to bound from below the first term in parentheses in the right-hand side as claimed in (4.2), but this is a direct consequence of concavity in γ of the free energy.

The validity of (4.3) follows from (4.1) and (4.2) by elementary considerations. □

We are now ready to show that the free and constrained models have the same free energy:

Proposition 4.2. *For every $h \in \mathbb{R}$ we have that the limit that defines $\mathrm{F}_\gamma(h)$, i.e. (1.12), exists and*

$$\mathrm{F}_\gamma(h) = \widetilde{\mathrm{F}}_\gamma(h). \tag{4.8}$$

Proof. Recall that we treat a result for $N \to \infty$ and $M \sim \gamma N$. First of all observe that, since $Z^f_{N,M,h} \geq Z^c_{N,M,h}$ we just need to worry about the upper bound. For the upper bound we use Proposition 3.1 and, more precisely the uniform estimate (3.2). As a matter of fact the supremum with respect to γ in (3.2) can be extended to $\gamma \in [1/L, L]$: the choice to restrict to $\gamma \geq 1$ in that proof was just to conform to the convention chosen from the beginning (but too restrictive for this proof as we will see). But the extension to $\gamma \in [1/L, L]$ can also be obtained from the result for $\gamma \in [1, L]$ by exploiting the symmetry of the expressions when exchanging M and N. Therefore we have that for every $\delta > 0$ and every $L > 1$ there exists ℓ_0 such that

$$Z^c_{N',M',h} \leq \exp\left(N'\left(\widetilde{\mathrm{F}}_{\gamma'_N}(h) + \frac{\delta}{2}\right)\right), \tag{4.9}$$

for every N' and M' such that both M' and $N' \geq \ell_0$ and $\gamma'_N = M'/N' \in [1/L, L]$. At the same time we require that

$$\frac{2\gamma h}{L} \leq \widetilde{\mathrm{F}}_\gamma(h), \tag{4.10}$$

where γ is the one that appears in the final statement, so $\gamma \geq 1$ without loss of generality, and for (4.10) to hold is just a matter of choosing L sufficiently large and this, in turn, just affects the choice of ℓ_0. But then thanks to Lemma 4.1

$$\widetilde{\mathrm{F}}_{\gamma_N}(h) - \frac{N'}{N}\widetilde{\mathrm{F}}_{\gamma'_N}(h) \geq 0, \tag{4.11}$$

so

$$Z^c_{N',M',h} \leq \exp\left(N\left(\widetilde{\mathrm{F}}_{\gamma_N}(h) + \frac{\delta}{2}\right)\right) \leq \exp\left(N\left(\widetilde{\mathrm{F}}_\gamma(h) + \delta\right)\right), \tag{4.12}$$

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

still for M' and $N' \geq \ell_0$ and $\gamma'_N = M'/N' \in [1/L, L]$: we have used the continuity of $\gamma \mapsto F_\gamma(h)$, see the caption of Fig. 7, and we have chosen N sufficiently large. Therefore by using $K_f(n) \leq n^C$ (for some $C > 0$) we have

$$Z^f_{N,M,h} \leq N^{2C} \sum_{N',M'} Z^c_{N',M',h} \leq (1+\gamma) N^{2C+2} \exp\left(N\left(\widetilde{\mathrm{F}}_{\gamma_N}(h) + \delta\right)\right)$$
$$+ N^{2C} \ell_0^2 \max_{N',M' \leq \ell_0} Z^c_{N',M',h} + N^{2C} \sum_{\substack{N',M' : N' \vee M' \geq \ell_0 \\ N' \leq N, M' \leq M, \gamma'_N \notin [1/L,L]}} Z^c_{N',M',h}. \tag{4.13}$$

Since the maximum in the second line is just a constant we are left with controlling the last sum. About this last sum, which we control without exploiting the condition $N' \vee M' \geq \ell_0$, we just remark that it can be split into the two cases $\gamma'_N < 1/L$ and $\gamma'_N > L$. The two terms are very similar and the second is larger than the first when $\gamma > 1$ (and N large) and if $\gamma = 1$ the bound is the same for both terms. So we focus on the second one and remark that the rough bound $Z^c_{N',M',h} \leq \exp(hN')$ and the observation that $\gamma_N N \geq M' = \gamma'_N N' > LN'$, hence $N' \leq 2\gamma N/L$ for N large, yield $Z^c_{N',M',h} \leq \exp(2\gamma h N/L)$. Therefore, by (4.10), this term is negligible with respect to the first term in the right-hand side of (4.13) and we are done. □

4.2. Sharp estimates on $Z^f_{N,M,h}$

The case $h > 0$: Proof of (1) in Theorem 1.6

Recall the definition of $Z^f_{N,M,h}$ in (1.10) and that we work with $M \sim \gamma N$, γ in the Cramer region. For every $a \in (0,1)$

$$\sum_{i=\lfloor aN \rfloor}^{N} \sum_{j=0}^{M} K_f(i) K_f(j) Z^c_{N-i,M-j,h}$$
$$= \sum_{i=\lfloor aN \rfloor}^{N} \sum_{j=0}^{M} K_f(i - \lfloor aN \rfloor) \frac{K_f(i)}{K_f(i - \lfloor aN \rfloor)} K_f(j) Z^c_{N-i,M-j,h}, \tag{4.14}$$

and since $K_f(\cdot)$ is positive and it has an asymptotic power law behavior there exists $C > 0$ such that (for N sufficiently large)

$$\sum_{i=\lfloor aN \rfloor}^{N} \sum_{j=0}^{M} K_f(i) K_f(j) Z^c_{N-i,M-j,h} \leq N^C \sum_{i=0}^{N-\lfloor aN \rfloor} \sum_{j=0}^{M} K_f(i) K_f(j) Z^c_{N-\lfloor aN \rfloor - i, M-j,h}$$
$$= N^C Z^f_{N-\lfloor aN \rfloor, M, h}. \tag{4.15}$$

Note that the leftmost term in (4.14) and (4.15) is $Z^f_{N,M,h}$ if $a = 0$. On the other hand by (3.2) and Proposition 4.2

$$\log\left(N^C Z^f_{N-\lfloor aN \rfloor,M,h}\right) \overset{N\to\infty}{\sim} N(1-a) \mathrm{F}_{\gamma_N/(1-a)}(h), \tag{4.16}$$

and we observe that

$$\mathrm{F}_{\gamma_N}(h) > (1-a) \mathrm{F}_{\gamma_N/(1-a)}, \tag{4.17}$$

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

which follows from (4.1) by choosing $M' = M$ so $N'/N = (\gamma_N/\gamma'_N)$ so that one obtains $\mathrm{F}_\gamma(h) - (\gamma/\gamma')\mathrm{F}_{\gamma'}(h) > 0$ for every $\gamma' > \gamma \geq 1$, and this inequality becomes (4.17) if we choose $\gamma' = \gamma_N/(1-a)$.

At this point we observe that, since $K_f(0) = 1$, we have $Z^f_{N,M,h} \geq Z^c_{N,M,h}$ and $\log Z^c_{N,M,h} \sim N\mathrm{F}_{\gamma_N}(h)$ (by (3.2)) and therefore, by (4.15)–(4.17) we see that for every $a \in (0,1)$ there exists $q > 0$ such that

$$Z^f_{N,M,h} = (1 + O(\exp(-qN))) \sum_{i=0}^{\lfloor aN \rfloor} \sum_{j=0}^{M} K_f(i)K_f(j)Z^c_{N-i,M-j,h}. \tag{4.18}$$

A parallel, somewhat easier, argument can be put at work when we restrict the summation in the definition of $Z^f_{N,M,h}$ in (1.10) to $j \geq \lfloor aN \rfloor$. We have to use again Lemma 4.1: (4.1) for $N' = N$ simply becomes the fact that $\gamma \mapsto \mathrm{F}_\gamma(h)$ is (strictly) increasing for $\gamma < \gamma_c(h)$ and this allows to conclude that for every $a \in (0,1)$ there exists $q > 0$ such that

$$Z^f_{N,M,h} = (1 + O(\exp(-qN))) \sum_{i=0}^{\lfloor aN \rfloor} \sum_{j=0}^{\lfloor aN \rfloor} K_f(i)K_f(j)Z^c_{N-i,M-j,h}. \tag{4.19}$$

With (4.19) we now want to show that we can restrict the sum in (1.10) to a small (since we can choose $a > 0$ small) macroscopic square. We want now to show that we can restrict almost to a microscopic square: a microscopic square would be a square of size that does not diverge with N. The result we shall now prove is

$$Z^f_{N,M,h} = \left(1 + O\left(\exp(-(\log N)^{3/2})\right)\right) \sum_{i=0}^{\ell_N} \sum_{j=0}^{\ell_N} K_f(i)K_f(j)Z^c_{N-i,M-j,h}, \tag{4.20}$$

where

$$\ell_N := \lfloor (\log N)^2 \rfloor. \tag{4.21}$$

For this choose a small so that $(M-j)/(N-i)$ is in the Cramer region for all values of i and j in the summation in (4.19). We can then apply (2.12) and, more precisely, the following consequence of (2.12): for every N sufficiently large

$$Z^c_{N',M',h} \leq C \exp\left(N'\mathrm{F}_{\gamma'_N}(h),\right), \tag{4.22}$$

where $C > 0$ and $N - \lfloor aN \rfloor \leq N' \leq N$, $M - \lfloor aN \rfloor \leq M' \leq M = \gamma_N N$. For this we exploit (4.3) of Lemma 4.1: since i and j are (macroscopically) small we have that there exists $c = c(h,a) > 0$ such that

$$N\mathrm{F}_{\gamma_N}(h) - N'\mathrm{F}_{\gamma'_N}(h) \geq c\left(N\left|\gamma'_N - \gamma_N\right| + \mathbf{1}_{\gamma'_N \geq \gamma_N}\left(M - M'\right) + \mathbf{1}_{\gamma'_N < \gamma_N}\left(N - N'\right)\right), \tag{4.23}$$

where $N' = N - i$, $M' = M - j$, N is sufficiently large and both i and j in their range of summation. But now we recall that we aim at (4.20) and therefore if $i < \ell_N$ then $j \geq \ell_N$ and the same is true if we exchange i and j. Therefore, omitting the constant c, the right-hand side of (4.23) is equal to $\gamma'_N(N - N') \geq (N - N')$ for $\gamma'_N \geq \gamma_N$ and since $N - N' = i$ and either $i \geq \ell_N$, or $i < \ell_N$ and then $j \geq \ell_N$, we get that $N - N' \geq (M - M')/\gamma_N \geq \ell_N/(2\gamma)$. For

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

$\gamma_N' \leq \gamma_N$ instead the right-hand side of (4.23) is equal to $M - N\gamma_N' + N - N'$ which, on the one hand, is bounded from below by $M - N\gamma_N + N - N' = N - N'$. On the other hand it is equal to $(M - M') - (N - N')(\gamma_N' - 1)$ which is bounded from below by $(M - M') - (N - N')(\gamma_N - 1)$. For $\gamma_N' \leq \gamma_N$ we have $M - M' \geq \ell_N$ since if $M - M' = j < \ell_N$ then $N - N' = i \geq \ell_N$ and we get that $\gamma_N' = (M - j)/(N - i) \geq (\gamma_N - \ell_N/N)/(1 - \ell_N/N)$ which is strictly larger than γ_N for N large. So either $N - N' \leq \frac{1}{2(\gamma_N - 1)}\ell_N$, so $(M - M') - (N - N')(\gamma_N - 1) \geq \frac{1}{2}\ell_N$, or $N - N' > \frac{1}{2(\gamma_N - 1)}\ell_N$. Hence the right-hand side of (4.23) is bounded from below for $\gamma_N' \leq \gamma_N$ by

$$\frac{c}{2}\ell_N \min\left(\frac{1}{\gamma_N - 1}, 1\right). \tag{4.24}$$

Recalling the lower bound found for $\gamma_N' \geq \gamma_N$, we see that if we set $c_\gamma := \frac{1}{2}\min(1/\gamma, 1)$ we have

$$N\mathrm{F}_{\gamma_N}(h) - N'\mathrm{F}_{\gamma_N'}(h) \geq c_{\gamma_N}\ell_N, \tag{4.25}$$

for $(N - N', M - M') \in ([0, aN]^2 \setminus [0, \ell_N)^2) \cap \mathbb{Z}^2$. But then (4.20) becomes evident from (4.22), (4.25) and the fact the summation is over less than $a^2N^2 = O(N^2)$ sites: so the total contribution by summing over the sites in small macroscopic square minus the almost microscopic square is $O(N^2 \exp(N\mathrm{F}_{\gamma_N}(h) - c_{\gamma_N}\ell_N))$. On the other hand $Z^f_{N,M,h} \geq Z^c_{N,M,h} \geq AN^{-1/2}\exp(N\mathrm{F}_{\gamma_N}(h))$ for some $A > 0$ and N large, cf. (2.12), and (4.20) is proven.

The question of the sharp estimates on Z^f is then reduced to find the leading behavior of

$$\sum_{i=0}^{\ell_N}\sum_{j=0}^{\ell_N} K_f(i)K_f(j)Z^c_{N-i,M-j,h}. \tag{4.26}$$

Observe that $(N - i, M - j)/|(N - i, M - j)|$ is close to $(1, \gamma)/|(1, \gamma)|$ and hence it is in a compact subset J of E_h for every $i, j \in [0, \ell_N]$. So we have by (2.12)

$$\begin{aligned}\sum_{i=0}^{\ell_N}\sum_{j=0}^{\ell_N} K_f(i)K_f(j)Z^c_{N-i,M-j,h} &\overset{N\to\infty}{\sim} \frac{A(\gamma)\exp\left(\mathrm{F}_{\gamma_N}(h)N\right)}{\sqrt{N}}\sum_{i=0}^{\ell_N}\sum_{j=0}^{\ell_N} K_f(i)K_f(j)\\ &\quad\times \exp\left(-N\left(\mathrm{F}_{\gamma_N}(h) - \mathrm{F}_{(M-j)/(N-i)}(h) + \frac{i}{N}\mathrm{F}_{(M-j)/(N-i)}(h)\right)\right).\end{aligned} \tag{4.27}$$

A Taylor expansion yields that there exists $\widetilde{\gamma} \in J$ such that

$$\begin{aligned}\mathrm{F}_{\gamma_N}(h) - \mathrm{F}_{(M-j)/(N-i)}(h) &= \left(\gamma_N - \frac{M-j}{N-i}\right)\partial_\gamma \mathrm{F}_\gamma(h)\big|_{\gamma=(M-j)/(N-i)}\\ &\quad + \frac{1}{2}\left(\gamma_N - \frac{M-j}{N-i}\right)^2 \partial^2_\gamma \mathrm{F}_\gamma(h)\big|_{\gamma=\widetilde{\gamma}},\end{aligned} \tag{4.28}$$

and since

$$\gamma_N - \frac{M-j}{N-i} = \frac{j}{N} - \gamma_N\frac{i}{N} + O\left(\frac{\ell_N^2}{N^2}\right) = O\left(\frac{\ell_N}{N}\right), \tag{4.29}$$

the second term in the right-hand side of (4.28) is $O\left(\ell_N^2/N^2\right)$ and finally

$$\begin{aligned}&\mathrm{F}_{\gamma_N}(h)-\mathrm{F}_{(M-j)/(N-i)}(h)+\frac{i}{N}\mathrm{F}_{(M-j)/(N-i)}(h)\\&\quad=\frac{j\partial_\gamma \mathrm{F}_\gamma(h)+i\left(\mathrm{F}_\gamma(h)-\gamma\partial_\gamma \mathrm{F}_\gamma(h)\right)}{N}+\frac{i}{N}O\left(\gamma_N-\gamma\right)+O\left(\frac{\ell_N^2}{N^2}\right).\end{aligned}\tag{4.30}$$

Recall now that $\mathrm{F}_\gamma(h)-\gamma\partial_\gamma \mathrm{F}_\gamma(h)$ is positive for $\gamma<\gamma_c(h)$, so that (4.30) implies that the double sum in the right-hand side of (4.27) converges to

$$C_{\gamma,h}:=\left(\sum_{i=0}^{\infty}K_f(i)\exp\left(-i\left(\mathrm{F}_\gamma(h)-\gamma\partial_\gamma \mathrm{F}_\gamma(h)\right)\right)\right)\left(\sum_{j=0}^{\infty}K_f(j)\exp\left(-j\partial_\gamma \mathrm{F}_\gamma(h)\right)\right),\tag{4.31}$$

and

$$Z^f_{N,M,h}\sim A(\gamma)C_{\gamma,h}\frac{\exp\left(\mathrm{F}_{\gamma_N}(h)N\right)}{\sqrt{N}}.\tag{4.32}$$

This completes the proof of (1) in Theorem 1.6 with $c_{\gamma,h}=A(\gamma)C_{\gamma,h}$. □

The case $h<0$: Proof of (2) in Theorem 1.6

In the delocalized phase, $\widetilde{\tau}_h$ is terminating and $Z^c_{N,M,h}=\mathbf{P}\left((N,M)\in\widetilde{\tau}_h\right)$. Recall that τ is recurrent, i.e. $\sum_{n,m}\mathrm{K}(n,m)=1$, then

$$\widetilde{K}_h(\infty)=1-\sum_{(n,m)\in\mathbb{N}^2}\widetilde{K}_h(n,m)=1-\exp(h)>0.\tag{4.33}$$

For the renewal function we write

$$\mathbf{P}\left((N,M)\in\widetilde{\tau}_h\right)=\sum_{j=0}^{\infty}\widetilde{K}_h^{j*}(N,M)=\sum_{j=0}^{\infty}\exp(jh)\mathrm{K}^{j*}(N,M),\tag{4.34}$$

where $\widetilde{K}_h^{j*}(\cdot,\cdot)$ is the j-fold convolution of $\widetilde{K}_h(\cdot,\cdot)$.

For the rest of the section, c_i, $i=1,2,\ldots$ denote positive constants. We have the following estimate:

Lemma 4.3. *Let $\alpha>0$, there exists $c>0$ such that for every $j\in\mathbb{N}$ and for every (N,M)*

$$\mathrm{K}^{j*}(N,M)\le j^c\,\mathrm{K}(N,M).\tag{4.35}$$

Moreover,

$$\lim_{N,M\to\infty}\frac{\mathrm{K}^{j*}(N,M)}{\mathrm{K}(N,M)}=j.\tag{4.36}$$

Proof. It is clear that (4.35) holds for $j=1$. Assume that it is true for $j<2s$ and we want to show it for $j=2s$. Observe that

$$
\begin{aligned}
\mathrm{K}^{2s*}(N, M) &\le 2 \sum_{n=1}^{\lfloor N/2\rfloor} \sum_{m=1}^{\lfloor M/2\rfloor} \mathrm{K}^{s*}(n, m)\mathrm{K}^{s*}(N-n, M-m) \\
&\quad + \left(\sum_{n=1}^{\lfloor N/2\rfloor} \sum_{m=\lfloor M/2\rfloor+1}^{M} + \sum_{n=\lfloor N/2\rfloor+1}^{N} \sum_{m=1}^{\lfloor M/2\rfloor} \right) \mathrm{K}^{s*}(n, m)\mathrm{K}^{s*}(N-n, M-m) \\
&:= Q_1 + (Q_2 + Q_3).
\end{aligned}
\tag{4.37}
$$

First we have by induction hypothesis

$$
Q_1 \le 2s^c \sum_{n=1}^{\lfloor N/2\rfloor} \sum_{m=1}^{\lfloor M/2\rfloor} \mathrm{K}^{s*}(n, m)\mathrm{K}(N-n, M-m), \tag{4.38}
$$

and from (A.1), we see that there exists $c_1 > 0$ such that $L(ux) \le c_1 L(x)$ for every $u \in [1/2, 1]$ and $x \ge 1$, therefore $\mathrm{K}(N-n, M-m) \le c_1 2^{2+\alpha}\mathrm{K}(N, M)$ and

$$
Q_1 \le c_1 2^{3+\alpha-c}(2s)^c \mathrm{K}(N, M) \sum_{n,m} \mathrm{K}^{s*}(n, m). \tag{4.39}
$$

Now observe that $Q_2(N, M) = Q_3(M, N)$, so in view of the bound we are after it suffices to consider Q_2. Applying (4.35) to the first term in the sum in Q_2, there exists $c_2 > 0$ such that

$$
Q_2 \le s^c \sum_{n=1}^{\lfloor N/2\rfloor} \sum_{m=\lfloor M/2\rfloor+1}^{M} \mathrm{K}(n, m)\mathrm{K}^{s*}(N-n, M-m) \le c_2 2^{-c}(2s)^c K(M). \tag{4.40}
$$

On the other hand, applying (4.35) to the second term in the sum in Q_2, we get

$$
Q_2 \le s^c \sum_{n=1}^{\lfloor N/2\rfloor} \sum_{m=\lfloor M/2\rfloor+1}^{M} \mathrm{K}^{s*}(n, m)\mathrm{K}(N-n, M-m) \le c_2 2^{-c}(2s)^c K(N). \tag{4.41}
$$

It suffices then to prove that there exists $c_3 > 0$ such that $K(N) \wedge K(M) \le c_3 \mathrm{K}(N, M)$. By elementary arguments we see that this follows if we can show that for every $x, y \ge 0$ and for every slowly varying function $L(\cdot)$, there exists $c_{L(\cdot)} > 0$ such that

$$
xL(x) \vee yL(y) \ge c_{L(\cdot)}(x+y)L(x+y). \tag{4.42}
$$

By symmetry it suffices to consider the case $y \ge x$ and in this case it suffices to show that $yL(y) \ge c_{L(\cdot)}(x+y)L(x+y)$. Since $(x+y) \in [y, 2y]$ we can apply (A.1) to see that there exists $c'_{L(\cdot)} > 0$ such that $L(y) \ge c'_{L(\cdot)}L(x+y)$ for every $y \ge x$ and this implies the desired inequality, and therefore (4.42), with $c_{L(\cdot)} = c'_{L(\cdot)}/2$.

Therefore $Q_2 + Q_3 \le c_4 2^{1-c}(2s)^c \mathrm{K}(N, M)$ and

$$
\mathrm{K}^{2s*}(N, M) \le (c_1 2^{3+\alpha-c} + c_4 2^{1-c})(2s)^c \mathrm{K}(N, M), \tag{4.43}
$$

and if $c = 1 + \log_2(c_1 2^{2+\alpha} + c_4)$, we obtain (4.35) for $j = 2s$. The procedure can be repeated for $j = 2s+1$ with minor changes. Therefore (4.35) is proven.

For what concerns (4.36), observe that it holds for $j = 1$. Assume that it is still valid up to $j = s$ and write

$$
\frac{\mathrm{K}^{(s+1)*}(N, M)}{\mathrm{K}(N, M)} = \sum_{n=1}^{N-1} \sum_{m=1}^{M-1} \frac{\mathrm{K}^{s*}(n, m)\mathrm{K}(N-n, M-m)}{\mathrm{K}(N, M)}. \tag{4.44}
$$

Split the double sum in (4.44) to four terms:

$$S_1 = \sum_{n=1}^{\lfloor N/2 \rfloor} \sum_{m=1}^{\lfloor M/2 \rfloor} \mathrm{K}^{s*}(n,m) \frac{\mathrm{K}(N-n, M-m)}{\mathrm{K}(N,M)} \tag{4.45}$$

$$S_2 = \sum_{n=1}^{\lfloor N/2 \rfloor -1} \sum_{m=1}^{\lfloor M/2 \rfloor -1} \mathrm{K}(n,m) \frac{\mathrm{K}^{s*}(N-n, M-m)}{\mathrm{K}(N,M)} \tag{4.46}$$

$$S_3 = \sum_{n=1}^{\lfloor N/2 \rfloor} \sum_{m=1}^{\lfloor M/2 \rfloor -1} \mathrm{K}^{s*}(n,M-m) \frac{\mathrm{K}(N-n, m)}{\mathrm{K}(N,M)} \tag{4.47}$$

$$S_4 = \sum_{n=1}^{\lfloor N/2 \rfloor -1} \sum_{m=1}^{\lfloor M/2 \rfloor} \mathrm{K}^{s*}(N-n,m) \frac{\mathrm{K}(n, M-m)}{\mathrm{K}(N,M)}. \tag{4.48}$$

For any fixed n and m and as $N, M \to \infty$, the two ratios in S_1 and S_2 converge respectively to 1 and to s (recall that $\mathrm{K}(n,m) = K(n+m)$) and the two ratios are uniformly bounded (from the uniform convergence property of the slowly varying functions (A.1) for the ratio in S_1 and from (4.35) for the ratio in S_2). Then by (DOM), we obtain $S_1 + S_2 \to 1 + s$ as $N, M \to \infty$.

Since S_3 and S_4 are essentially the same quantity when we exchange M and N, we just focus on S_3. If we first assume that $M \geq N$ (hence $M + N \in [M, 2M]$), from (4.35) and (A.1) we obtain that

$$\begin{aligned} S_3 &\leq s^c \sum_{n=1}^{\lfloor N/2 \rfloor} \sum_{m=1}^{\lfloor M/2 \rfloor -1} \mathrm{K}(n, M-m) \frac{\mathrm{K}(N-n, m)}{\mathrm{K}(N,M)} \\ &\leq c_5 s^c \sum_{n=1}^{\lfloor N/2 \rfloor} \sum_{m=1}^{\lfloor M/2 \rfloor -1} \frac{L(M+n)}{(M+n)^{2+\alpha}} \frac{L(N+m)}{(N+m)^{2+\alpha}} \frac{(N+M)^{2+\alpha}}{L(N+M)} \\ &\leq c_6 s^c \sum_{n=1}^{\lfloor N/2 \rfloor} \sum_{m=1}^{\lfloor M/2 \rfloor -1} \frac{L(N+m)}{(N+m)^{2+\alpha}} \leq \frac{c_6 s^c}{2} N \sum_{n>N} \frac{L(n)}{n^{2+\alpha}} \leq c_7 L(N) N^{-\alpha}. \end{aligned} \tag{4.49}$$

By repeating the argument for $N \geq M$ we obtain that $S_3 = O(L(M)M^{-\alpha})$ in this case. Therefore, since $\alpha > 0$, $\lim_{N,M\to\infty} S_3 = 0$ by the basic properties of slowly varying functions and an elementary argument. □

To prove part (2) in Theorem 1.6, we need to know the sharp estimates in the constrained case for $h < 0$:

Proposition 4.4. *If $h < 0$, then there exists $c_h > 0$ such that for every (N, M)*

$$Z^c_{N,M,h} \leq c_h \mathrm{K}(N,M) \tag{4.50}$$

where $c_h = \sum_{j=0}^{\infty} j^c \exp(jh)$. Moreover

$$Z^c_{N,M,h} \overset{N,M\to\infty}{\sim} \frac{\exp(h)}{(1-\exp(h))^2} \mathrm{K}(N,M). \tag{4.51}$$

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

Proof. From (4.34), we have

$$\frac{Z^c_{N,M,h}}{\mathrm{K}(N,M)} = \frac{\mathbf{P}\left((N,M)\in\widetilde{\tau}_h\right)}{\mathrm{K}(N,M)} = \sum_{j=0}^{\infty}\exp(jh)\frac{\mathrm{K}^{j*}(N,M)}{\mathrm{K}(N,M)}, \tag{4.52}$$

and using (4.35), for fixed j, we see that the ratio is bounded above by j^c, therefore we obtain (4.50).

For (4.51), the ratio in (4.52) converges to j as $N, M \to \infty$ from (4.36) and bounded from (4.35), then by (DOM) we get

$$\lim_{N,M\to\infty}\frac{Z^c_{N,M,h}}{\mathrm{K}(N,M)} = \sum_{j=0}^{\infty} j\exp(jh) = \frac{\exp(h)}{(1-\exp(h))^2}. \quad \square \tag{4.53}$$

We are now ready to prove the sharp estimate of $Z^f_{N,M,h}$:

Proposition 4.5. *Suppose that $M \sim \gamma N$. For $h < 0$, as $N \to \infty$*

- *If $\overline{\alpha} < 1 + \alpha/2$, we have*

$$Z^f_{N,M,h} \sim \frac{K_f(N)K_f(M)}{1-\exp(h)}. \tag{4.54}$$

- *If $\overline{\alpha} > 1 + \alpha/2$, we have*

$$Z^f_{N,M,h} \sim \frac{\exp(h)\left(\sum_{n\geq 0} K_f(n)\right)^2}{(1-\exp(h))^2}\mathrm{K}(N,M). \tag{4.55}$$

Proof. Let us write

$$\frac{Z^f_{N,M,h}}{K_f(N)K_f(M)} = \sum_{n=0}^{N}\sum_{m=0}^{M} Z^c_{n,m,h}\frac{K_f(N-n)K_f(M-m)}{K_f(N)K_f(M)}. \tag{4.56}$$

We split the last sum into

$$\begin{aligned} &T_1+T_2+T_3+T_4 \\ &= \left(\sum_{n=0}^{\lfloor N/2\rfloor}\sum_{m=0}^{\lfloor M/2\rfloor} + \sum_{n=\lfloor N/2\rfloor+1}^{N}\sum_{m=\lfloor M/2\rfloor+1}^{M} + \sum_{n=0}^{\lfloor N/2\rfloor}\sum_{m=\lfloor M/2\rfloor+1}^{M} + \sum_{n=\lfloor N/2\rfloor+1}^{N}\sum_{m=0}^{\lfloor M/2\rfloor}\right) \\ &\quad \times Z^c_{n,m,h}\frac{K_f(N-n)K_f(M-m)}{K_f(N)K_f(M)}. \end{aligned} \tag{4.57}$$

For T_1, for fixed n and m, the ratio in (4.57) converges to 1 and by (A.1), this ratio is bounded. Then by (DOM), Fubini–Tonelli Theorem and the fact that $\mathrm{K}(\cdot,\cdot)$ is a discrete probability density we obtain (recall (1.17) and in this case $\mathrm{G}(h) = 0$)

$$\lim_{N,M\to\infty} T_1 = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} Z^c_{n,m,h} = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\sum_{j=0}^{\infty} \widetilde{\mathrm{K}}_h^{j*}(n,m)$$
$$= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\sum_{j=0}^{\infty} \exp(jh)\mathrm{K}^{j*}(n,m) = \frac{1}{1-\exp(h)}. \tag{4.58}$$

For T_2, if $\overline{\alpha} \in [1, 1+\alpha/2)$ and $\sum_{n\geq 0} K_f(n) < \infty$, using (4.51), we obtain (recall that $M \sim \gamma N$)

$$T_2 \leq c_8 \sum_{n=\lfloor N/2\rfloor+1}^{N} \sum_{m=\lfloor M/2\rfloor+1}^{M} \mathrm{K}(n,m) \frac{K_f(N-n)K_f(M-m)}{K_f(N)K_f(M)}$$
$$= O\left(N^{2\overline{\alpha}-\alpha-2} \frac{L(N)}{(\overline{L}(N))^2}\right), \tag{4.59}$$

so $T_2 \to 0$ as $N \to \infty$.

If $\overline{\alpha} \leq 1$ and $\sum_{n\geq 0} K_f(n) = \infty$, using once again (4.51) with (A.2), we get

$$T_2 \leq c_8 \sum_{n=\lfloor N/2\rfloor+1}^{N} \sum_{m=\lfloor M/2\rfloor+1}^{M} \mathrm{K}(n,m) \frac{K_f(N-n)K_f(M-m)}{K_f(N)K_f(M)}$$
$$\leq c_9 \mathrm{K}(N,M) \frac{\sum_{n=\lfloor N/2\rfloor+1}^{N} K_f(N-n) \sum_{m=\lfloor M/2\rfloor+1}^{M} K_f(M-m)}{K_f(N)K_f(M)} = O\left(N^{-\alpha}L(N)\right), \tag{4.60}$$

and $T_2 \to 0$ as $N \to \infty$ in this case too.

Let us look at T_3 (the argument for T_4 is identical). We have

$$T_3 = \sum_{n=0}^{\lfloor N/2\rfloor} \sum_{m=\lfloor M/2\rfloor+1}^{M} Z^c_{n,m,h} \frac{K_f(N-n)}{K_f(N)} \frac{K_f(M-m)}{K_f(M)}$$
$$\leq c_{10} \sum_{n=0}^{\lfloor N/2\rfloor} K(M+n) \sum_{m=\lfloor M/2\rfloor+1}^{M} \frac{K_f(M-m)}{K_f(M)} \leq c_{11} \frac{L(M)}{M^{1+\alpha}} \frac{\sum_{m=0}^{M} K_f(m)}{K_f(M)}. \tag{4.61}$$

If $\overline{\alpha} \in [1, 1+\alpha/2)$ and $\sum_m K_f(m) < \infty$ then $T_3 = O(N^{\overline{\alpha}-\alpha-1}L(N)/\overline{L}(N))$ and it tends to zero because $\alpha > 0$ implies $1+\alpha > \overline{\alpha}$. If instead $\overline{\alpha} \leq 1$ and $\sum_m K_f(m) = \infty$ the last ratio in the rightmost term in (4.61) is $O(M) = O(N)$. Hence $T_3 = O(L(N)N^{-\alpha})$ and T_3 tends to zero also in this case. The proof of (4.54) is therefore complete.

Now for $\overline{\alpha} > 1 + \alpha/2$ (which implies that $\overline{\alpha} > 1$ since $\alpha > 0$), we write

$$\frac{Z^f_{N,M,h}}{\mathrm{K}(N,M)} = \sum_{n=0}^{N}\sum_{m=0}^{M} Z^c_{n,m,h} \frac{K_f(N-n)K_f(M-m)}{\mathrm{K}(N,M)}. \tag{4.62}$$

We split the last sum to $U_1 + U_2 + U_3 + U_4$ as in (4.57). Since $\sum_{i\ge 0}\sum_{j\ge 0} Z^c_{i,j,h} = 1/(1-\exp(h)) < \infty$ (see (4.58)) we have

$$U_1 = O\left(\frac{K_f(N)K_f(M)}{\mathrm{K}(N,M)}\right) = O\left(N^{2+\alpha-2\overline{\alpha}}\frac{\left(\overline{L}(N)\right)^2}{L(N)}\right) = o(1). \tag{4.63}$$

Also the terms U_3 and U_4 give a vanishing contribution. Let us see it for U_3 (the computation is identical for U_4):

$$\begin{aligned} U_3 &= \sum_{n=0}^{\lfloor N/2\rfloor}\sum_{m=\lfloor M/2\rfloor+1}^{M}\frac{Z^c_{n,m,h}}{\mathrm{K}(N,M)}K_f(N-n)K_f(M-m) \\ &\le c_{12}\sum_{n=0}^{\lfloor N/2\rfloor}K_f(N-n)\sum_{m=\lfloor M/2\rfloor+1}^{M}K_f(M-m) \le c_{13}\sum_{n\ge N/2}K_f(n) \\ &= O(\overline{L}(N)N^{1-\overline{\alpha}}). \end{aligned} \tag{4.64}$$

The relevant contribution comes from U_2:

$$\begin{aligned} U_2 &= \sum_{n=\lfloor N/2\rfloor+1}^{N}\sum_{m=\lfloor M/2\rfloor+1}^{M}\frac{Z^c_{n,m,h}}{\mathrm{K}(N,M)}K_f(N-n)K_f(M-m) \\ &= \sum_{n=0}^{N-\lfloor N/2\rfloor-1}\sum_{m=0}^{M-\lfloor M/2\rfloor-1}\frac{Z^c_{N-n,M-m,h}}{\mathrm{K}(N,M)}K_f(n)K_f(m). \end{aligned} \tag{4.65}$$

But the ratio in the last term is bounded, cf. (4.51), and in fact (4.51) tells us also that for every m and n

$$\lim_{N,M\to\infty}\frac{Z^c_{N-n,M-m,h}}{\mathrm{K}(N,M)} = \frac{\exp(h)}{(1-\exp(h))^2}, \tag{4.66}$$

which, by applying (DOM), implies

$$\lim_{N,M\to\infty} U_2 = \frac{\exp(h)}{(1-\exp(h))^2}\left(\sum_{n\ge 0}K_f(n)\right)^2, \tag{4.67}$$

and completes the proof of (4.55) and, in turn, the proof of Proposition 4.4. □

4.3. Path properties: proof of Theorem 1.7

In this section, we suppose that $M \sim \gamma N$ and $\alpha > 0$.

Proof of (1) in Theorem 1.7. We first consider the case $h < 0$. Recall that $(\mathcal{F}_1, \mathcal{F}_2)$ is the last renewal epoch in $[0, N] \times [0, M]$. If $\overline{\alpha} < 1 + \alpha/2$, for fixed i and j (so we can assume $i < N$ and $j < M$) we have

$$\begin{aligned} \mathbf{P}^f_{N,M,h}((\mathcal{F}_1,\mathcal{F}_2) = (i,j)) &= \frac{Z^c_{i,j,h}K_f(N-i)K_f(M-j)}{Z^f_{N,M,h}} \\ &\overset{N\to\infty}{\sim} (1-\exp(h))Z^c_{i,j,h}\frac{K_f(N-i)K_f(M-j)}{K_f(N)K_f(M)}, \end{aligned} \tag{4.68}$$

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

where the estimation follows from (4.54). Since i and j are $O(1)$ and by (A.1) the ratio in the rightmost term in (4.68) converges to one. Hence it suffices to prove that $(1-\exp(h))\sum_{i,j} Z^c_{i,j,h} = 1$, but this is done in (4.58). We then recall that $Z_{i,j,h} = \mathbf{P}((i,j) \in \widetilde{\tau}_h)$ and (1.33) is proven.

Now recall that $(\mathcal{L}_1, \mathcal{L}_2) := (N - \mathcal{F}_1, M - \mathcal{F}_2)$. If $\overline{\alpha} > 1 + \alpha/2$, for fixed i and j by using (4.51) and (4.55) we see that

$$\mathbf{P}^f_{N,M,h}((\mathcal{L}_1, \mathcal{L}_2) = (i,j)) = \frac{Z^c_{N-i,M-j,h} K_f(i) K_f(j)}{Z^f_{N,M,h}} \overset{N\to\infty}{\sim} \frac{\mathrm{K}(N-i, M-j)}{\mathrm{K}(N,M)} \frac{K_f(i)K_f(j)}{\left(\sum_{n\geq 0} K_f(n)\right)^2}. \tag{4.69}$$

The proof of (1.34) is therefore complete by observing that by (A.1) the first ratio in (4.69) converges to one.

We are then left (for $h < 0$) with (1.35). Here we prove more: consider $(\mathcal{E}_1, \mathcal{E}_2) := \max\{\tau \cap [0, \lfloor N/2 \rfloor] \times [0, \lfloor M/2 \rfloor]\}$ under $\mathbf{P}^f_{N,M,h}$ for $\overline{\alpha} > 1 + \alpha/2$. If i and j are fixed, by (4.55) we have

$$\mathbf{P}^f_{N,M,h}((\mathcal{E}_1, \mathcal{E}_2) = (i,j)) = \exp(h) \sum_{s \geq \lfloor N/2 \rfloor} \sum_{t \geq \lfloor M/2 \rfloor} \frac{Z^c_{i,j,h} \mathrm{K}(s-i, t-j) Z^f_{N-s,M-t,h}}{Z^f_{N,M,h}}$$
$$\overset{N\to\infty}{\sim} \frac{(1-\exp(h))^2 Z^c_{i,j,h}}{\left(\sum_{n\geq 0} K_f(n)\right)^2} \sum_{s \geq \lfloor N/2 \rfloor} \sum_{t \geq \lfloor M/2 \rfloor} Z^f_{N-s,M-t,h} \frac{\mathrm{K}(s-i, t-j)}{\mathrm{K}(N,M)}. \tag{4.70}$$

By making the change of variable $(s,t) \to (N-s, M-t)$ we see that for s and t of $O(1)$, the very last ratio in (4.70) converges to one and the same ratio is bounded by (A.1) in all the range of the sum. Hence, by the (DOM), the expression in (4.70) converges to

$$\frac{(1-\exp(h))^2 Z^c_{i,j,h}}{\left(\sum_{n\geq 0} K_f(n)\right)^2} \sum_{s,t\geq 0} Z^f_{s,t,h}, \tag{4.71}$$

and observe that

$$\sum_{s,t\geq 0} Z^f_{s,t,h} = \sum_{s,t \geq 0} \sum_{n=0}^{s} \sum_{m=0}^{t} Z^c_{n,m,h} K_f(s-n) K_f(t-m)$$
$$= \sum_{n,m\geq 0} Z^c_{n,m,h} \sum_{s\geq n} \sum_{m \geq t} K_f(s-n) K_f(t-m) = \frac{(\sum_{n\geq 0} K_f(n))^2}{1-\exp(h)}, \tag{4.72}$$

where the last equality follows from (4.58). Therefore the law of $(\mathcal{E}_1, \mathcal{E}_2)$ under $\mathbf{P}^f_{N,M,h}$ converges for $N \to \infty$ to the probability distribution that assigns to (i,j) probability $(1-\exp(h))Z^c_{i,j,h}$ (which is correctly normalized, like (1.33), by (4.58)).

Let $(\mathcal{C}_1, \mathcal{C}_2) := \min\{\tau \cap [\lfloor N/2 \rfloor, N] \times [\lfloor M/2 \rfloor, M]\}$ be the first renewal epoch in $[\lfloor N/2 \rfloor, N] \times [\lfloor M/2 \rfloor, M]$ and set $(\mathcal{H}_1, \mathcal{H}_2) := (N - \mathcal{C}_1, M - \mathcal{C}_2)$. For fixed i and j by (4.55) we have

$$\mathbf{P}^f_{N,M,h}\left((\mathcal{H}_1, \mathcal{H}_2) = (i, j)\right) = \exp(h) \sum_{s \geq \lfloor N/2 \rfloor} \sum_{t \geq \lfloor M/2 \rfloor} \frac{Z^f_{i,j,h} \mathrm{K}(s-i, t-j) Z^c_{N-s,M-t,h}}{Z^f_{N,M,h}}$$

$$\underset{N\to\infty}{\sim} \frac{(1-\exp(h))^2 Z^f_{i,j,h}}{\left(\sum_{n \geq 0} K_f(n)\right)^2} \sum_{s \geq \lfloor N/2 \rfloor} \sum_{t \geq \lfloor M/2 \rfloor} Z^c_{N-s,M-t,h} \frac{\mathrm{K}(s-i, t-j)}{\mathrm{K}(N, M)}. \tag{4.73}$$

The second ratio converges to 1 (same argument as above) and therefore (4.73) converges to

$$\frac{(1-\exp(h)) Z^f_{i,j,h}}{\left(\sum_{n \geq 0} K_f(n)\right)^2}, \tag{4.74}$$

and from (4.72), we see that this expression adds up ($i, j \geq 0$) to one. The law of $(\mathcal{H}_1, \mathcal{H}_2)$ converges as $N \to \infty$ to the probability distribution that assigns to (i, j) probability (4.74). We conclude that the contacts are either close to $(0, 0)$ or to the last renewal epoch: therefore we get (1.35). □

Proof of (2) in Theorem 1.7. In this case $h > 0$ and $\gamma \in (1/\gamma_c(h), \gamma_c(h))$: the positivity of $\mathrm{F}_\gamma(h) - \gamma \partial_\gamma \mathrm{F}_\gamma(h)$ and $\partial_\gamma \mathrm{F}_\gamma(h)$ is a direct consequence of the strict concavity of $\mathrm{F}_\cdot(h)$ in the Cramér region (see caption of Fig. 7). Then choose (i, j) with non-negative integer entries. By (2.12) and (4.32) we have

$$\mathbf{P}^f_{N,M,h}\left((\mathcal{L}_1, \mathcal{L}_2) = (i, j)\right) = \frac{Z^c_{N-i,M-j,h} K_f(i) K_f(j)}{Z^f_{N,M,h}}$$

$$\underset{N\to\infty}{\sim} \frac{A\left((M-j)/(N-i)\right)\sqrt{N}}{C_{\gamma,h} A(\gamma) \sqrt{N-i}} \exp\left(\mathrm{F}_{(M-j)/(N-i)}(h)(N-i) - \mathrm{F}_{M/N}(h) N\right)$$

$$\times K_f(i) K_f(j). \tag{4.75}$$

Observe now that the ratio in the rightmost term of (4.75) converges as $N \to \infty$ to $1/C_{\gamma,h}$ (defined in (4.31)) and using (4.30) and the fact that $(M-j)/(N-i)$ is close to γ, (4.75) converges to

$$\frac{1}{C_{\gamma,h}} \exp\left(-j \partial_\gamma \mathrm{F}_\gamma(h) - i\left(\mathrm{F}_\gamma(h) - \gamma \partial_\gamma \mathrm{F}_\gamma(h)\right)\right) K_f(i) K_f(j), \tag{4.76}$$

where $C_{\gamma,h}$ is defined in (4.31). Hence (1.36) is proven.

To complete the proof of Theorem 1.7, we need to show that for $h > 0$ and if $M \sim \gamma N$ such that $\gamma \in (1/\gamma_c(h), \gamma_c(h))$, $\mathbf{P}^f_{N,M,h}$ converges, for $N \to \infty$, to the law of a bivariate renewal with the inter-arrival probability given in (1.37). For this, fix a $k \in \mathbb{N}$ and for every $\{(i_n, j_n)\}_{n=0,1,\ldots,k}$

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

with $(i_0, j_0) = (0, 0)$, $i_1 < i_2 < \cdots < i_k < N$ and $j_1 < j_2 < \cdots < j_k < M$, we have

$$\begin{aligned}&\mathbf{P}^f_{N,M,h}\left(\tau_n = (i_n, j_n) \text{ for } n = 1, 2, \ldots, k\right)\\ &\quad = \left(\prod_{n=1}^{k} K(i_n - i_{n-1} + j_n - j_{n-1}) \exp(h)\right) \frac{Z^f_{N-i_k, M-j_k, h}}{Z^f_{N,M,h}}.\end{aligned} \tag{4.77}$$

Since $\gamma_k = (M - j_k)/(N - i_k)$ is close to γ (in the Cramér region) as $N \to \infty$ for fixed i_k and j_k, by (4.32), we see that the ratio in (4.77) is equal to

$$1 + o(1) \exp\left(-N\left(\mathrm{F}_{\gamma_N}(h) - \mathrm{F}_{\gamma_k}(h) + \frac{i_k}{N}\mathrm{F}_{\gamma_k}(h)\right)\right). \tag{4.78}$$

Following the same procedure used for (4.30), we see that the exponent converges to

$$-\, i_k(\mathrm{F}_\gamma(h) - \gamma\partial_\gamma \mathrm{F}_\gamma(h)) - j_k\partial_\gamma \mathrm{F}_\gamma(h), \tag{4.79}$$

and therefore the left-hand side of (4.77) converges to

$$\begin{aligned}&\prod_{n=1}^{k} K(i_n - i_{n-1} + j_n - j_{n-1})\\ &\quad \times \exp\left(h - (i_n - i_{n-1})(\mathrm{F}_\gamma(h) - \gamma\partial_\gamma \mathrm{F}_\gamma(h)) - (j_n - j_{n-1})\partial_\gamma \mathrm{F}_\gamma(h)\right).\end{aligned} \tag{4.80}$$

We are left with proving that (1.37) is a probability distribution. Recall from (3.3) and (3.4) that

$$D_h(1, \gamma) = \max_{(\lambda_1,\lambda_2)\in B_h} (\lambda_1 + \gamma\lambda_2) = \hat{\lambda}_1(\gamma) + \gamma\hat{\lambda}_2(\gamma), \tag{4.81}$$

with of course $(\hat{\lambda}_1(\gamma), \hat{\lambda}_2(\gamma)) \in B_h$ (keep in mind that they depend also on h) i.e.

$$\sum_{n,m} K(n + m) \exp\left(h - n(\mathrm{G}(h) - \hat{\lambda}_1(\gamma)) - m(\mathrm{G}(h) - \hat{\lambda}_2(\gamma))\right) = 1, \tag{4.82}$$

therefore $\mathrm{F}_\gamma(h) = (\mathrm{G}(h) - \hat{\lambda}_1(\gamma)) + \gamma(\mathrm{G}(h) - \hat{\lambda}_2(\gamma))$. Replace $\mathrm{G}(h) - \hat{\lambda}_1(\gamma)$ in (4.82) by $\mathrm{F}_\gamma(h) - \gamma(\mathrm{G}(h) - \hat{\lambda}_2(\gamma))$, we obtain

$$\sum_{n,m} K(n + m) \exp\left(h - n\mathrm{F}_\gamma(h) - (m - \gamma n)(\mathrm{G}(h) - \hat{\lambda}_2(\gamma))\right) = 1, \tag{4.83}$$

and recall that $\hat{c}_\gamma(h) = \mathrm{F}_\gamma(h)$ in the Cramér regime, then by (3.17) one obtains

$$\sum_{n,m}(m - \gamma n)K(n + m) \exp\left(h - n\mathrm{F}_\gamma(h) - (m - \gamma n)(\mathrm{G}(h) - \hat{\lambda}_2(\gamma))\right) = 0. \tag{4.84}$$

Differentiating (4.83) with respect to γ and using (4.84), we obtain that

$$\begin{aligned}&\left(\partial_\gamma \mathrm{F}_\gamma(h) - (\mathrm{G}(h) - \hat{\lambda}_2(\gamma))\right)\\ &\quad \times \sum_{n,m} nK(n + m) \exp\left(-n\mathrm{F}_\gamma(h) - (m - \gamma n)(\mathrm{G}(h) - \hat{\lambda}_2(\gamma))\right),\end{aligned} \tag{4.85}$$

is zero, and this implies directly that

$$\partial_\gamma \mathrm{F}_\gamma(h) = \mathrm{G}(h) - \hat{\lambda}_2(\gamma), \tag{4.86}$$

$$\mathrm{F}_\gamma(h) - \gamma \partial_\gamma \mathrm{F}_\gamma(h) = \mathrm{G}(h) - \hat{\lambda}_1(\gamma). \tag{4.87}$$

Therefore from (4.82), we get that (1.37) is a probability distribution. □

Appendix. Slowly and regularly varying functions

A function $L : [0, \infty) \to (0, \infty)$ is a *slowly varying* function at ∞ if it is measurable and if $\lim_{x\to\infty} \frac{L(ux)}{L(x)} = 1$ for every $u > 0$. The function $x \mapsto L(1/x)$ is slowly varying at zero if $L(\cdot)$ is slowly varying at ∞. It can be shown that this convergence holds uniformly in u [2, Theorem 1.2.1]: for every $0 < c_1 < c_2 < \infty$

$$\lim_{x\to\infty} \sup_{u\in[c_1,c_2]} \left| \frac{L(ux)}{L(x)} - 1 \right| = 0. \tag{A.1}$$

A function of the form $x \mapsto x^a L(x)$, $a \in \mathbb{R}$, is said to be *regularly varying (at ∞) of exponent a* with an analogous definition for regularly varying at zero. Examples of slowly varying functions include logarithmic functions (of course the trivial example is the constant) like $a(\log(x))^b$ as $x \to \infty$ with $a > 0$ and $b \in \mathbb{R}$. We refer to [2] for the full theory of slowly and regularly varying functions: we just recall some basic important facts. First of all, both $L(x)$ and $1/L(x)$ are $o(x^\varepsilon)$ for every $\varepsilon > 0$, which directly implies that if $f(\cdot)$ is regularly varying with exponent a and $g(\cdot)$ is regularly varying with exponent $b < a$, then $g(x) = o(f(x))$.

We will often use that for $\beta > 0$ [2, Sec. 1.5.6]

$$\sum_{n\geq N} \frac{L(n)}{n^{1+\beta}} \overset{N\to\infty}{\sim} \frac{L(N)}{\beta N^\beta} \quad \text{and} \quad \sum_{n=1}^{N} \frac{L(n)}{n^{1-\beta}} \overset{N\to\infty}{\sim} \frac{L(N)}{\beta N^{-\beta}}, \tag{A.2}$$

which can be proven by Riemann sum approximation. We will often use Riemann sum approximations involving regularly varying functions also beyond (A.2) and the central tool to control these approximations is the so called *Potter bounds* [2, Th. 1.5.6].

Another important issue is about asymptotic invertibility of regular functions: a regular function of exponent $a > 0$ (respectively $a < 0$) is asymptotically equivalent to an increasing (respectively decreasing) function. Moreover the inverse of a monotonic regularly varying function of exponent $a \neq 0$ is a regularly varying function of exponent $1/a$. In different terms, if $f(\cdot)$ is regularly varying of exponent $a \neq 0$, then there exists $g(\cdot)$ regularly varying of exponent $1/a$ such that $f(g(x)) \sim g(f(x)) \sim x$ [2, Sec. 1.5.7]. Occasionally we use other properties of slowly varying functions and we refer directly to [2].

References

[1] Q. Berger, J. Poisat, On the critical curve of the pinning and copolymer models in correlated Gaussian environment, Electron. J. Probab. 20 (2015) 35.
[2] N.H. Bingham, C.M. Goldie, J.L. Teugels, Regular Variation, Cambridge University Press, Cambridge, 1987.
[3] A.A. Borovkov, A.A. Mogul'skiĭ, The second rate function and the asymptotic problems of renewal and hitting the boundary for multidimensional random walks, Sibirsk. Mat. Zh. 37 (4) (1996) 745–782.
[4] A.A. Borovkov, A.A. Mogul'skiĭ, Integro-local limit theorems including large deviations for sums of random vectors II, Theory Probab. Appl. 45 (2001) 3–22.

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017

[5] A.A. Borovkov, A.A. Mogul'skiĭ, On large deviations of sums of independent random vectors on the boundary and outside of the cramér zone I, Theory Probab. Appl. 53 (2) (2009) 301–311.

[6] R. Brak, A.L. Owczarek, T. Prellberg, A.J. Guttmann, Finite-length scaling of collapsing directed walks, Phys. Rev. E 48 (1993) 2386–2396.

[7] R. Bundschuh, T. Hwa, Statistical mechanics of secondary structures formed by random RNA sequences, Phys. Rev. E 65 (2002) 031903 (22 pages).

[8] F. Caravenna, G. Giacomin, L. Zambotti, Sharp asymptotic behavior for wetting models in (1+1)-dimension, Electron. J. Probab. 11 (2006) 345–362.

[9] P. Carmona, N. Pétrélis, G.B. Nguyen, Interacting partially self-avoiding walk. From phase transition to the geometry of the collapsed phase, Ann. Probab., in press, preprint (2013), http://arxiv.org/abs/1306.4887.

[10] F. den Hollander, Random polymers, in: Lectures from the 37th Probability Summer School Held in Saint-Flour, 2007, in: Lecture Notes in Mathematics 1974, Springer-Verlag, 2009.

[11] R.A. Doney, An analogue of the renewal theorem in higher dimensions, Proc. Lond. Math. Soc. (3) 16 (1966) 669–684.

[12] T.R. Einert, H. Orland, R.R. Netz, Secondary structure formation of homopolymeric single-stranded nucleic acids including force and loop entropy: implications for DNA hybridization, Eur. Phys. J. E 34 (2011) 55 (15 pages).

[13] M.E. Fisher, Walks, walls, wetting, and melting, J. Stat. Phys. 34 (1984) 667–729.

[14] T. Garel, H. Orland, Generalized Poland–Scheraga model for DNA hybridization, Biopolymers 75 (2004) 453–467.

[15] T. Garel, H. Orland, On the role of mismatches in DNA denaturation, arXiv:cond-mat/0304080.

[16] G. Giacomin, Random Polymer Models, Imperial College Press, World Scientific, 2007.

[17] G. Giacomin, Renewal convergence rates and correlation decay for homogeneous pinning models, Electron. J. Probab. 13 (2008) 513–529.

[18] G. Giacomin, Disorder and critical phenomena through basic probability models, in: Lecture Notes From the 40th Probability Summer School Held in Saint-Flour, 2010, in: Lecture Notes in Mathematics, vol. 2025, Springer, 2011.

[19] G. Giacomin, F.L. Toninelli, The localized phase of disordered copolymers with adsorption, ALEA Lat. Am. J. Probab. Math. Stat. 1 (2006) 149–180.

[20] A. Kabakçıŏglu, A. Bar, D. Mukamel, Macroscopic loop formation in circular DNA denaturation, Phys. Rev. E 85 (2012) 051919.

[21] Y. Kafri, D. Mukamel, L. Peliti, Why is the DNA denaturation transition first order? Phys. Rev. Lett. 85 (2000) 4988–4991.

[22] S.G. Krantz, H.R. Parks, A Primer of Real Analytic Functions, second ed., in: Birkhäuser Advanced Texts: Basel Textbooks, Birkhäuser Boston, Inc., Boston, MA, 2002.

[23] A. Litan, A statistical mechanical treatment of the open ends in a double-stranded polynucleotide molecule, Biopolymers 2 (1964) 279–282.

[24] R.A. Neher, U. Gerland, Intermediate phase in DNA melting, Phys. Rev. E 73 (2006) 030902R.

[25] G.B. Nguyen, N. Pétrélis, A variational formula for the free energy of the partially directed polymer collapse, J. Stat. Phys. 151 (2013) 1099–1120.

[26] D. Poland, H.A. Scheraga, Theory of Helix-Coil Transitions in Biopolymers: Statistical Mechanical Theory of Order-Disorder Transitions in Biological Macromolecules, Academic Press, 1970.

[27] C. Richard, A.J. Guttmann, Poland–Scheraga models and the DNA denaturation transition, J. Stat. Phys. 115 (2004) 925–947.

[28] K. Sture, J. Nordholm, S.A. Rice, Phase transitions and end effects in models of biopolymers, J. Chem. Phys. 59 (1973) 5605–5613.

[29] E. Yeramian, Complexity and tractability. statistical mechanics of Helix-Coil transitions in circular DNA as a model-problem, Europhys. Lett. EPL 25 (1964) 49–55.

Please cite this article in press as: G. Giacomin, M. Khatib, Generalized Poland–Scheraga denaturation model and two-dimensional renewal processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.06.017