Supplementary Materials to the article "Intrachain distances in a crumpled polymer with random loops"

Statistical weights of the diagrams

In order to compute the average over different classes of diagrams one needs to know their statistical contributions. These weights were originally derived by P.J. Pedler [1] for the two-state Markov process and were recently used by some of us for the computation of the contact probability function $P_c(s)$ in the presence of disorder of loops. Despite their ultimate expressions can be found [2] and [3] we provide their derivation here for completeness.

Let us treat the sizes of the loops and of the gaps as the independent random variables having exponential probability distributions with $\alpha_l = \lambda^{-1}$ and $\alpha_g = g^{-1}$. Then, in order to derive the statistical weights of the diagrams, it is convenient to introduce an auxiliary Markov jump process with two states, "Loop" and "Gap", in continuous time where time intervals are measured in the units of the polymer contour length. Clearly, the statistics of the random pattern of alternating loops and gaps is analogues to the joint statistics of the time intervals which the Markov process spends in the "Loop" and the "Gap" states in the course of its stochastic dynamics.

Based on this analogy, we can express the statistical weight of the diagram (a) as follows

$$\mathcal{W}_a(x|s) = p_g g(G, s|G, 0)\mathcal{F}(x|s), \tag{S1}$$

where $p_g = \frac{\alpha_l}{\alpha_g+\alpha_l}$ is the probability to find the Markov process in the state "Gap" at a random moment of time (i.e. the steady-state probability that a randomly chosen point of the polymer belongs to a gap); $g(G, s|G, 0) = \frac{1}{\alpha_g+\alpha_l}(\alpha_l + \alpha_g e^{-(\alpha_l+\alpha_g)s})$ is the probability to find the Markov jump process in the state "Gap" after time s under the condition that it starts in the state "Gap" (i.e. the probability that the second point, which is separated by contour distance s from the first point, also belongs to a gap region). and $\mathcal{F}(x|s)$ is the probability distribution of the random variable x representing the fraction of time that the Markov process spent in the state "Gap" under the condition that it starts in the state "Gap" and find itself in the same state after time s (i.e., x is the fraction of the contour length which is occupied by loops for a segment whose both ends belong to the gap regions), which is given by (see [1])

$$\mathcal{F}(x|s) = \frac{\alpha_g + \alpha_l}{\alpha_l + \alpha_g e^{-(\alpha_l+\alpha_g)s}} \left\{ e^{-\alpha_g s}\delta(x) + \left(\frac{\alpha_g\alpha_l(1-x)s^2}{x}\right)^{1/2} I_1\left[2\sqrt{\alpha_g\alpha_l x(1-x)s^2}\right] e^{-\alpha_g(1-x)s-\alpha_l xs} \right\}, \tag{S2}$$

with $I_1[x]$ denoting the modified Bessel function of the first kind [4]. In the polymer language x is the fraction of the loops contour length of a segment that has both ends in gaps.

Next, the statistical weight of configurations responding to the diagram (b) is

$$\mathcal{W}_b(l_1, l_2, x|s) = 2p_l\rho(l_1)\rho(l_2)g(G, s|G, l_2)\mathcal{F}(x|s - l_2), \tag{S3}$$

where $p_l = \frac{\alpha_g}{\alpha_g+\alpha_l}$ denotes the probability to find the statistically stationary Markov process in the state "Loop" if visited at a random time moment; the factors $\rho(l_1) = \alpha_l e^{-\alpha_l l_1}$ and $\rho(l_2) = \alpha_l e^{-\alpha_l l_2}$ represent, respectively, the probability densities of the random time l_1 elapsed since the last entrance to the state "Loop" and of the random time l_2 remaining before the next jump to the state "Gap", and $\mathcal{F}(x|s - l_2)$ is the probability distribution of the fraction of time x that the Markov process spent in the state "Gap" under the condition that it starts in the state "Gap" and find itself also in the state "Gap" after the time $s - l_2$. Note also that factor 2 in Eq. (S3) arises due to the left-right symmetry in the choice of point which is assumed to reside on the loop.

Similarly, the statistical weight of configurations responding to the diagram (c) is

$$\mathcal{W}_c(l_1, l_2|s) = p_l\rho(l_1)\rho(l_2). \tag{S4}$$

Finally, the statistical weight of subchains described by the diagram (d) is given by

$$\mathcal{W}_d(l_1, l_2, h, x, \tilde{L}|s) = p_l\rho(l_1)\rho(l_2)\rho(\tilde{L})\alpha_g g(G, l_2 + h|G, l_2)\mathcal{F}(x|h). \tag{S5}$$

Final expressions for the mean value

Substituting (Eq.S1)-(Eq.S5) in the equations of the averaging procedure from the main text (see Eq.(6)-(10) there), one can come to the following expressions for an arbitrary function $F(\mathbf{R})$ on the radius vector $\mathbf{R}$ between polymer sites:

$$\langle F_{(a)}(s \mid x)\rangle = \int_0^1 \mathrm{d}x \; \mathcal{W}_a(x \mid s) F_{(a)}(s \mid x) = \tag{S6}$$

$$= \frac{\alpha_l}{\alpha_l+\alpha_g} \cdot F_{(a)}(s \mid 0) + \int_0^1 \mathrm{d}x \; s \left(\frac{\alpha_l \alpha_g (1-x)}{x}\right)^{1/2} I_1\left[2s\sqrt{\alpha_l\alpha_g x(1-x)}\right] \times \tag{S7}$$

$$\times \, e^{-(\alpha_g(1-x)+x\alpha_l)s} \; F_{(a)}(s \mid x)\Big) , \tag{S8}$$

$$\langle F_{(b)}(s \mid l_1, l_2, x)\rangle = \int_0^{+\infty} \mathrm{d}l_1 \int_0^{s} \mathrm{d}l_2 \int_0^1 \mathrm{d}x \; \mathcal{W}_b(l_1, l_2, x \mid s) F_{(b)}(s \mid l_1, l_2, x) = \tag{S9}$$

$$= \frac{2\alpha_g\alpha_l^2}{\alpha_l+\alpha_g} \cdot \left(\int_0^{+\infty} \mathrm{d}l_1 \int_0^{s} \mathrm{d}l_2 \; e^{-\alpha_l(l_1+l_2)-\alpha_g(s-l_2)} F_{(b)}(s \mid l_1, l_2, 0) + \int_0^{+\infty} \mathrm{d}l_1 \int_0^{s} \mathrm{d}l_2 \int_0^1 \mathrm{d}x \; (s-l_2)\times \right. \tag{S10}$$

$$\times \left(\frac{\alpha_l\alpha_g(1-x)}{x}\right)^{1/2} I_1\left[2(s-l_2)\sqrt{\alpha_l\alpha_g x(1-x)}\right] e^{-\alpha_l(l_1+l_2)-(\alpha_g(1-x)+x\alpha_l)(s-l_2)} \cdot F_{(b)}(s \mid l_1, l_2, x)\Big) , \tag{S11}$$

$$\langle F_{(c)}(s \mid l_1, l_2)\rangle = \int_0^{+\infty} \mathrm{d}l_1 \int_s^{+\infty} \mathrm{d}l_2 \; \mathcal{W}_c(l_1, l_2 \mid s) F_{(c)}(s \mid l_1, l_2) = \tag{S12}$$

$$= \frac{\alpha_g\alpha_l^2}{\alpha_l+\alpha_g} \cdot \int_0^{+\infty} \mathrm{d}l_1 \int_s^{+\infty} \mathrm{d}l_2 \; e^{-\alpha_l(l_1+l_2)} F_{(c)}(s \mid l_1, l_2), \tag{S13}$$

$$\langle F_{(d)}(s \mid l_1, l_2, h, \tilde{L}, x)\rangle = \int_0^{+\infty} \mathrm{d}l_1 \int_0^{s} \mathrm{d}l_2 \int_0^{s-l_2} \mathrm{d}h \int_{s-l_2-h}^{+\infty} \mathrm{d}\tilde{L} \int_0^1 \mathrm{d}x \; \mathcal{W}_d(l_1, l_2, h, x, \tilde{L} \mid s) F_{(d)}(s \mid l_1, l_2, h, \tilde{L}, x) = \tag{S14}$$

$$= \frac{\alpha_g^2\alpha_l^3}{\alpha_l+\alpha_g} \cdot \left(\int_0^{+\infty} \mathrm{d}l_1 \int_0^{s} \mathrm{d}l_2 \int_0^{s-l_2} \mathrm{d}h \int_{s-l_2-h}^{+\infty} \mathrm{d}\tilde{L} \; e^{-\alpha_l(l_1+l_2)-\alpha_l\tilde{L}-\alpha_g h} F_{(d)}(s \mid l_1, l_2, h, \tilde{L}, 0) + \right. \tag{S15}$$

$$+ \int_0^{+\infty} \mathrm{d}l_1 \int_0^{s} \mathrm{d}l_2 \int_0^{s-l_2} \mathrm{d}h \int_{s-l_2-h}^{+\infty} \mathrm{d}\tilde{L} \int_0^1 \mathrm{d}x h \left(\frac{\alpha_g\alpha_l(1-x)}{x}\right)^{1/2} I_1\left[2h\sqrt{\alpha_g\alpha_l x(1-x)}\right] e^{-\alpha_l(l_1+l_2)-\alpha_l\tilde{L}-\alpha_g(1-x)h-\alpha_l x h} \cdot \tag{S16}$$

$$\cdot F_{(d)}(s \mid l_1, l_2, h, \tilde{L}, x)\Big) , \tag{S17}$$

where $F_{(i)}(s \mid \{A\}_i)$ is the mean value of a function $F(\mathbf{R})$ on the diagram (i), the variances for each diagram are expressed due to their independence

$$\begin{aligned}
\sigma^2_{\text{(a)}}\left[s, x\right] &= \sigma^2_{\text{free}}\left[s(1-x)\right] , \\
\sigma^2_{\text{(b)}}\left[s, l_1, l_2, x\right] &= \sigma^2_{\text{bridge}}\left[l_2, l_1+l_2\right] + \sigma^2_{\text{free}}\left[(s-l_2)(1-x)\right] , \\
\sigma^2_{\text{(c)}}\left[s, l_1, l_2\right] &= \sigma^2_{\text{bridge}}\left[s, l_1+l_2\right] , \\
\sigma^2_{\text{(d)}}\left[s, l_1, l_2, h, \tilde{L}, x\right] &= \sigma^2_{\text{bridge}}\left[l_2, l_1+l_2\right] + \sigma^2_{\text{free}}\left[h(1-x)\right] + \sigma^2_{\text{bridge}}\left[s-h-l_2, \tilde{L}\right]
\end{aligned} \tag{S18}$$

and

$$\sigma^2_{\text{free}}\left[s\right] = b^2 s^{2/d_f} \tag{S19}$$

for a fractal polymer with arbitrary fractal dimension $d_f \geq 2$.

2

REFERENCES

[1] P. Pedler, Occupation times for two-state markov chains, J.Appl.Prob. **8**, 381 (1971).

[2] K. Polovnikov, B. Slavov, S. Belan, M. Imakaev, H. B. Brandão, and L. A. Mirny, Crumpled polymer with loops recapitulates key features of chromosome organization, bioRxiv (2023).

[3] K. Polovnikov and B. Slavov, Topological and nontopological mechanisms of loop formation in chromosomes: Effects on the contact probability, Physical Review E **107**, 054135 (2023).

[4] M. Abramowitz, I. A. Stegun, and R. H. Romer, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, US Government printing office **55** (1964).