Magnetic Polymer Models for Epigenetics-Driven Chromosome Folding

Davide Colì* and Enzo Orlandini[†]

Dipartimento di Fisica and Sezione INFN, Università degli Studi di Padova, I-35131 Padova, Italy

Davide Michieletto* \ddagger

SUPA, School of Physics and Astronomy, University of Edinburgh, Edinburgh EH9 3FD, UK

MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine,

University of Edinburgh, Edinburgh EH4 2XU, UK and

Centre for Mathematical Biology, and Department of Mathematical Sciences,

University of Bath, North Rd, Bath BA2 7AY, UK

Davide Marenduzzo

SUPA, School of Physics and Astronomy, University of Edinburgh, Edinburgh EH9 3FD, UK

Epigenetics is a driving force of important and ubiquitous phenomena in nature such as cell differentiation or even metamorphosis. Oppositely to its widespread role, understanding the biophysical principles that allow epigenetics to control and rewire gene regulatory networks remains an open challenge. In this work we study the effects of epigenetic modifications on the spatial folding of chromosomes – and hence on the expression of the underlying genes – by mapping the problem to a class of models known as magnetic polymers. In this work we show that a first order phase transition underlies the simultaneous spreading of certain epigenetic marks and the conformational collapse of a chromosome. Further, we describe Brownian Dynamics simulations of the model in which the topology of the polymer and thermal fluctuations are fully taken into account and that confirm our mean field predictions. Extending our models to allow for non-equilibrium terms yields new stable phases which qualitatively agrees with observations in vivo. Our results show that statistical mechanics techniques applied to models of magnetic polymers can be successfully exploited to rationalize the outcomes of experiments designed to probe the interplay between a dynamic epigenetic landscape and chromatin organization.

I. INTRODUCTION

Epigenetic effects in biology are defined as inheritable changes to a phenotype that do not involve alterations in the underlying DNA sequence [1-3]. The best and perhaps most important example of this class of changes is that resulting in the difference between cell types in our body: while all cells possess the same DNA, they can differentiate into neurons, epithelium, retinal, etc. Other important examples in which epigenetic effects are at work are the reprogramming of germ [4] and pluripotent [5] cells, temperature-dependent sex determination in certain fish and reptiles [6], metamorphosis of certain animals (such as caterpillars into butterflies) and generic polyphenism in insects [7]. All these widespread phenomena share the fact that while the original gene regulatory networks are rewired to give rise to previously absent features, organs or appendices, the underlying DNA sequence is untouched.

It is becoming increasingly clear that epigenetic changes are orchestrated through certain biochemical – or epigenetic – marks that are deposited along the genome. How these marks affect gene expression is, however, still poorly understood. One widespread idea is that epigenetic marks change the spatial folding of the genome and hence its accessibility for proteins and transcription factors; in turn, this change in local folding affects the expression of the underlying genes [8]. Increasing amount of evidence shows that genome folding correlates with epigenetics [9] but the biophysical mechanisms controlling this folding are unclear, especially considering that epigenetic marks are transient biochemical modifications which are dynamically deposited to and removed from the substrate [3, 5].

First, several models have been proposed to explain how a chromosome may fold given a certain *fixed* pattern of epigenetic marks: among the most successful are the bridging-induced attraction [10–12], the strings-andbinders [13] and co-polymer [14] models (Fig. 1A). Second, a different class of models have been developed to understand how epigenetic marks may spread and self-organise into patterns on a *passive* substrate with a given average probability of looping [15–18] (Fig. 1B). Both classes of models lack an element that is natural to consider in a more complete and refined model: the interplay between the kinetics of the epigenetic marks and the dynamical spatial arrangements of the polymeric substrate (see Fig. 1C). Accounting for this interplay maps the problem to that of "magnetic polymer" or "annealed co-polymer" models [19]. So far few groups have developed polymer models apt to specifically address questions concerning chromatin folding [20–24] but there are still many open questions that

^{*} Joint first author

[†] orlandini@pd.infn.it

[‡] Joint first author; davide.michieletto@ed.ac.uk



FIG. 1. Classes of Models for Epigenetic Patterns and Chromosome Folding. A Co-polymer [13, 14] and bridginginduced attraction [10–12] models predict the folding of a chromosome given a certain fixed pattern of epigenetic marks. B Ising-like models predict the spreading and patterns of epigenetic marks given a passive substrate with a given average contact probability [15]. C In this work we study magnetic polymer models that couple the dynamics of epigenetic marks with that of the polymeric substrate.

can be addressed within a more physically-minded statistical mechanics framework.

In this, and the companion [25], work we focus on fieldtheoretic approaches to study analytically and numerically solvable models of magnetic polymers. This paper is organised as follows. In Section II we derive a free energy for a single chromosome with epigenetic and 3D dynamics from the partition function of a magnetic polymer. In Section III we derive the kinetics equations for the equilibrium model and generalise them to described non-equilibrium conditions. In Sections IVA and IVB we detail the computational strategy that we use to perform Brownian dynamics simulations of coarse-grained polymers with dynamic (epigenetic) "recolouring" and we explain how to break detailed balance and achieve non-equilibrium conditions (also explored in Ref. [20]). Finally, in Section IVC we discuss the Brownian dynamics simulations of melts of polymers with recolouring dynamics. By studying these models we show that equilibrium models are not compatible with experiments and a qualitative agreement is recovered by accounting for non-equilibrium processes. We show that non-equilibrium effects can either promote new phases or stabilise an arrested phase separated coexistence of epigenetic marks.

II. EQUILIBRIUM MODEL

We describe a single chromosome as an N-step selfavoiding walk (SAW) on a lattice with coordination number z and with each vertex *i* displaying an epigenetic state, or spin, q_i . A generic situation is one in which two classes of epigenetic states (plus one neutral) are present: those marking transcriptionally active (say q = 1) and inactive (say q = -1) chromatin [3, 26]. A neutral state (q = 0) may be associated to a region of chromatin that does not contain any clear epigenetic mark, such as gene deserts in Drosophila [27]. Non-neutral states are known to associate with specific protein complexes that can mediate chromosome bridging through "reading" domains (e.g. HP1 [28], PRC [29] or others [10, 13, 14]) and epigenetic spreading through "writing" domains (e.g. Suv39 [20, 30]).

Within this magnetic polymer framework, both processes (bridging and spreading) can be modelled by the same ferromagnetic interaction that tends to align 3D proximal spins, or tends to bring them together when already aligned [20–22]. Thus, any pair of 3D proximal monomers interact with each other via a contact potential that depends on the values of q in a pair. More precisely, if the *i*-th and the *j*-th monomers are nearest neighbours on the lattice, their contact energy $J(q_i, q_j)$ is

$$J(q_i, q_j) = \begin{cases} -\epsilon & \text{if } q_i = q_j = \pm 1\\ 0 & \text{otherwise} \end{cases}$$
, (1)

with $\epsilon > 0$. Note that the mark q = 0 does not contribute to this configurational energy and we will define it as a neutral mark. The equilibrium properties of this system are described by the partition function

$$\mathcal{Z} = \sum_{\text{SAW}} \sum_{\{q\}} \exp\left[-\frac{\beta}{2} \sum_{i,j=1}^{N} \Delta_{\boldsymbol{r}_i, \boldsymbol{r}_j} J(q_i, q_j)\right], \quad (2)$$

where $1/\beta = k_B T$. The sums $\sum_{\text{SAW}} \text{and } \sum_{\{q\}} \text{run over the}$ set of all *N*-steps SAWs and all the possible combinations of *N* epigenetic states, respectively. The matrix $\Delta_{\boldsymbol{r}_i \boldsymbol{r}_j}$ is the adjacency matrix associated to a given SAW and is given by

$$\Delta_{\boldsymbol{r}_i \boldsymbol{r}_j} = \begin{cases} 1 & \text{if } \boldsymbol{r}_i \text{ and } \boldsymbol{r}_j \text{ are nearest neighbours in 3D} \\ 0 & \text{otherwise} \end{cases}$$
(3)

Notice that the partition function in Eq. (2) presents a clear \mathbb{Z}_2 symmetry as $J(-q_i, -q_j) = J(q_i, q_j)$.

Since here we focus on the transitions between a disordered and an ordered phase, irrespective of the mark type, we follow the strategy, sometimes used in Potts models [31], to restrict the phase space of the epigenetic variables, q, to the case where the states q = 0 and q = -1 are equally populated. With this restriction the \mathbb{Z}_2 symmetry of Eq. (2) is violated but the system can be faithfully described by a two-valued spin variable $S = \{1, -\frac{1}{2}\}$ where S = 1 corresponds to the mark q = 1, while the values $S = -\frac{1}{2}$ has multiplicity 2 as it corresponds both to q = 0 and q = -1.

By using the spin variable S, the coupling in Eq. (1) becomes

$$J(S_i, S_j) = \begin{cases} -\epsilon & \text{if } S_i = S_j = 1\\ -\frac{1}{4}\epsilon & \text{if } S_i = S_j = -\frac{1}{2}\\ 0 & \text{otherwise} \end{cases}$$
(4)

which can be re-written as

$$J(S_i, S_j) = -\frac{5}{9}\epsilon \left(S_i + \frac{1}{5}\right) \left(S_j + \frac{1}{5}\right) - \frac{1}{5}\epsilon.$$
 (5)

The partition function in Eq. (2) can then be recast into the following form,

$$\mathcal{Z} = \sum_{\text{SAW}} \sum_{\{S\}} \exp\left[\frac{\epsilon\beta}{2} \sum_{i,j=1}^{N} \Delta_{r_i r_j} \left(\frac{5}{9} (S_i + \frac{1}{5})(S_j + \frac{1}{5}) + \frac{1}{5}\right)\right] \,.$$
(6)

We now perform a Hubbard-Stratonovich transformation to rewrite the term

$$\exp\left[\frac{5\epsilon\beta}{18}\sum_{i,j=1}^{N}(S_i+\frac{1}{5})\Delta_{\boldsymbol{r}_i\boldsymbol{r}_j}(S_j+\frac{1}{5})\right]$$
(7)

as

$$\int \mathcal{D}\boldsymbol{\phi} \exp\left[-\frac{9}{10\epsilon\beta}\sum_{i,j=1}^{N}\phi_i\Delta_{\boldsymbol{r}_i\boldsymbol{r}_j}^{-1}\phi_j + \sum_{i=1}^{N}\phi_i(S_i + \frac{1}{5})\right]$$

where $\mathcal{D}\phi = (2\pi)^{-N/2} (9/5\epsilon\beta)^{1/2} (\det \Delta^{-1})^{1/2} \prod_{i=1}^{N} \mathrm{d}\phi_i$. By summing over all possible spin configurations $\{S\}$ we get

$$\int \mathcal{D}\phi \exp\left[-\frac{9}{10\epsilon\beta}\sum_{i,j=1}^{N}\phi_i\Delta_{\boldsymbol{r}_i\boldsymbol{r}_j}^{-1}\phi_j\right]\prod_{i=1}^{N}\left(e^{\frac{6}{5}\phi_i}+2e^{-\frac{3}{10}\phi_i}\right)$$

which can be re-written as

$$\int \mathcal{D}\boldsymbol{\phi} \exp\left[-\frac{9}{10\epsilon\beta}\sum_{i,j=1}^{N}\phi_i\Delta_{\boldsymbol{r}_i\boldsymbol{r}_j}^{-1}\phi_j + \sum_{i=1}^{N}\log\left(e^{\frac{6}{5}\phi_i} + 2e^{-\frac{3}{10}\phi_i}\right)\right]$$
(8)

This integral can be evaluated through a homogeneous saddle point approximation, i.e. by replacing the integral with the value of the integrand at its maximum which we assume is attained for $\phi_i = \phi$. This approximation yields

$$\mathcal{Z} = \sum_{\text{SAW}} A \exp\left[-\frac{9}{10\epsilon\beta}\phi^2 \sum_{i,j=1}^N \Delta_{\boldsymbol{r}_i\boldsymbol{r}_j}^{-1} + N \log\left(e^{\frac{6}{5}\phi} + 2e^{-\frac{3}{10}\phi}\right)\right]$$
(9)

where $A = (2\pi)^{-N/2} (9/5\epsilon\beta)^{1/2} (\det \Delta^{-1})^{1/2}$. In general, the term $\sum_{i,j=1}^{N} \Delta_{r_i r_j}^{-1}$, depends on the given SAW and it is not easy to compute. However, it can be estimated if we restrict the set of SAWs to the ones that are almost space filling, such as for example the subset of Hamiltonian walks [19]. A Hamiltonian walk is a path that visits each vertex of a lattice embedded in a volume V exactly once and have been used to study equilibrium properties of highly compact polymers [32, 33]. For a Hamiltonian walk, the adjacency matrix of the SAW Δ takes the same form of the adjacency matrix of the underlying lattice and it is characterised by the coordination number z. Hence, $\sum_{i,j=1}^{N} \Delta_{\mathbf{r}_i \mathbf{r}_j}^{-1} = N/z$. Here, we consider N-steps configurations that, similarly to Hamiltonian walks, are contained in a volume V but may in principle display a lower mean number of nearest neighbours, i.e. ρz instead of z. One can thus approximate

$$\sum_{i,j=1}^{N} \Delta_{\boldsymbol{r}_i \boldsymbol{r}_j}^{-1} \approx \frac{N}{\rho z} \quad . \tag{10}$$



FIG. 2. Free Energy of Single Chromosome in Equilibrium. A-B Contour plots of the free energy in Eq. (13) for $\alpha = 2.5$ and $\alpha = 4$ respectively. C Plot of the minimised free energy $f(\phi^*, \rho; \alpha)$ as a function of ρ and for different values of α . One can notice that, at the critical value $\alpha_c \simeq 3.95$, there is a first order transition.

Note that for generic SAWs with low ρ , the right hand side in Eq. (10) is actually an upper bound. Finally, by following the approach described in Ref. [34] we evaluate the last term in Eq. (6), i.e.

$$e^{F_{\text{SAW}}} = \sum_{\gamma \in \text{SAW}} \exp\left[\frac{\beta \epsilon}{10} \sum_{i,j=1}^{N} \left(\Delta|_{\gamma}\right)_{ij}\right]$$
(11)

by approximating F_{SAW} as

$$\frac{F_{\rm SAW}}{TN} \approx -\log\left(\frac{z}{e}\right) + \frac{1-\rho}{\rho}\log(1-\rho) - \frac{\beta\epsilon z}{10}\rho \quad . \tag{12}$$

By collecting all the terms and taking $f = -\frac{T}{N} \log \mathcal{Z}$, we obtain the following mean-field free energy density

$$\frac{f}{T} = -\log\left(\frac{z}{e}\right) + \frac{1-\rho}{\rho}\log(1-\rho) - \frac{\alpha}{10}\rho + \frac{9}{10\alpha}\frac{\phi^2}{\rho} - \log\left(e^{\frac{6}{5}\phi} + 2e^{-\frac{3}{10}\phi}\right) - \frac{1}{N}\log A \quad , (13)$$

where $\alpha \equiv \beta \epsilon z$ and the term $\log A/N$ is negligible for $N \to \infty$. The equilibrium properties of the model are then obtained by finding the value of magnetisation and density that minimize Eq. (13). First, the values of ϕ^* such that $\partial f/\partial \phi|_{\phi^*} = 0$ are obtained by solving

$$\frac{9\phi}{5\alpha\rho} + \frac{-\frac{3}{5}e^{-\frac{3\phi}{10}} + \frac{6}{5}e^{\frac{5\phi}{5}}}{2e^{-\frac{3\phi}{10}} + e^{\frac{6\phi}{5}}} = 0.$$
(14)

This equation can be solved numerically to find $\phi^*(\rho; \alpha)$ which can then be put back into the free energy to express it as $f(\phi^*, \rho; \alpha) = f^*(\rho; \alpha)$. This function is plotted in Fig. 2 for different values of α showing the appearance of a local minimum at $\rho > 0$, which becomes a global minimum for $\alpha > \alpha_c$. The behaviour shown in Fig. 2 is characteristic of a first order transition: i.e., the value of ρ^* minimising $f(\phi^*, \rho; \alpha)$ is discontinuous in α and jumps from $\rho^* = 0$ to $\rho^* > 0$ at $\alpha = \alpha_c \simeq 3.95$. This transition in density also corresponds to a discontinuous jump in magnetisation $\phi^*(\rho^*; a)$ from $\phi^* = 0$ to $\phi^* \neq 0$.

Thus, this theory gives two possible equilibrium phases (see Fig. 3). For $\alpha < \alpha_c$ the chain is extended in space ($\rho \simeq 0$) and has low magnetisation which corresponds to heterogeneous epigenetic marks ($\phi^2 = 0$): we dub this phase swollen-disordered (SD). For values of $\alpha > \alpha_c \simeq 3.96$ the chain is collapsed ($\rho > 0$) and nearly uniformly coloured ($\phi^2 > 0$) and hence we dub this phase compact-ordered (CO). Even though the free energy in Eq. (13) is obtained via several approximations, our theory is in excellent agreement with Brownian Dynamics simulations [20] and it unambiguously predicts the presence of a discontinuous jump of the order parameters ρ and ϕ at the transition.

As mentioned in the introduction, one of the most puzzling aspects of epigenetics is that, while it allows plastic transformations of cells or even organisms, it can be inherited even across generations. It is thus intriguing that our model predicts a first order transition for epigeneticallydriven folding since its discontinuous nature provides a mechanism that endows memory to the system. For instance, once a chromosome has been taken over by a single epigenetic mark – e.g., in the case of the inactive-X in female mammalian cells [29, 35] – our model predicts that such a state would be robust even against extensive perturbations such as those occurring during mitosis. Accordingly, it is well known that, while the initial choice of which of the two X's to inactivate is stochastic [35], this choice is remarkably robust across cell division; this balance between stochasticity and robustness of epigenetic processes can be appreciated everyday as it is manifested in the patchy coat of calico cats [36].

Our findings do not exclude, but rather complement, other mechanisms that are known to play important roles to ensure the correct transmission of a certain transcriptional programme, e.g. non-coding nuclear RNA [37], genomic bookmarks [38, 39] or specialised transcription factors [40].

Finally, we highlight that the two phases predicted by the equilibrium model do not capture in full the epigeneticconformational states observed *in vivo*. In the next section we provide a non-equilibrium generalisation of our model in order to obtain a wider spectrum of possible phases.

III. NON-EQUILIBRIUM MODEL

From the free energy in Eq. (13), we can derive dynamical equations for ρ and ϕ . Since both order pa-



FIG. 3. Equilibrium Phase Diagram for Single Chromosomes. The phase diagram is obtained by numerical minimisation of Eq. (13) (red and blue symbols refer respectively to the order parameter and density profiles). The first order transition occurs at $\alpha = \epsilon/k_BT \simeq 3.95$ and separates a swollendisordered (SD) phase from a compact-ordered (CO) one. The insets report snapshots from BD simulations at $\alpha_{BD} = 0.6$ and $\alpha_{BD} = 1.1$ (the critical value of α for the model used in the BD simulations is $\alpha_{BD,c} \simeq 0.9$ [20]). See Sec. IV for details on the BD simulations.

rameters are non-conserved (here ρ should be understood as the density of beads within the smallest box containing the polymer chain), we can obtain them by computing the steepest descent to the free energy minimum, i.e. $\partial_t \phi \sim -\delta \mathcal{H}/\delta \phi$ and $\partial_t \rho \sim -\delta \mathcal{H}/\delta \rho$, where \mathcal{H} is the energy functional $\mathcal{H}[\rho, \phi] = \int d\mathbf{x} \left[f + \kappa_{\phi} (\nabla \phi)^2 + \kappa_{\rho} (\nabla \rho)^2 \right]$. This set of coupled equations is analogous to the equation defining "Model A" for the dynamics of non-conserved fields [41, 42], and for our free energy they read

$$\partial_t \phi(\mathbf{r}, t) = \Gamma_\phi \left(-\frac{9\phi}{5\alpha_\phi\rho} - \frac{3}{e^{3\phi/2} + 2} + \frac{6}{5} + \kappa_\phi \nabla^2 \phi \right), \quad (15)$$
$$\partial_t \rho(\mathbf{r}, t) = \Gamma_\rho \left(\frac{9\phi^2}{10\alpha_\rho\rho^2} + \frac{\alpha_\rho}{10} + \frac{\ln(1-\rho)}{\rho^2} + \frac{1}{\rho} + \kappa_\rho \nabla^2 \rho \right),$$

where $\Gamma_{\rho/\phi}$ and $\kappa_{\rho/\phi}$ are mobilities and surface tension-like coefficients, respectively. In Eqs. (15) we decouple α into two independent parameters affecting the dynamics of the polymer (α_{ρ}) and of the epigenetic field (α_{ϕ}) separately. The case $\alpha_{\rho} \neq \alpha_{\phi}$ leads to non-equilibrium dynamics as these equations can no longer be derived from a free energy. Since the dynamics of chromatin and that of the epigenetic marks are very different processes there is no *a priori* constraint for which α_{ϕ} should be equal to α_{ρ} .

We have numerically integrated Eqs. (15) using a standard Euler method with time step dt = 0.01 on a 50×50 grid and setting $\Gamma_{\phi} = 0.1$, $\Gamma_{\rho} = 0.001$, $\kappa_{\phi} = \kappa_{\rho} = 2$. The choice of these parameters only affect the kinetics of the process, not its long-time steady state. The initial condition was set so that each lattice site had a random density and magnetisation broadly distributed around the means



FIG. 4. Non-Equilibrium Phase Diagram for Single Chromosomes. Steady state values of A magnetization and B density obtained by evolving Eqs. (15) on a 50 × 50 grid with $\Gamma_{\phi} = 0.1$, $\Gamma_{\rho} = 0.001$ and $\kappa_{\phi} = \kappa_{\rho} = 2$. C By cross-checking the two phase diagrams, we discover three regions with different combinations of ρ and ϕ^2 . In addition to the Swollen-Disordered ($\rho = 0$, $\phi^2 = 0$) and Compact-Ordered ($\rho = 1$, $\phi^2 > 1$) phases found in equilibrium, we now observe a semi-Compact-Disordered (sCD) phase ($0 < \rho < 1$, $\phi^2 = 0$). The red circle denotes the first-order transition seen in equilibrium. (Inset) Snapshots of representative configurations from BD simulations of a magnetic polymer N = 200 beads long (see Sec. IV for details). Red and blue beads represent competing epigenetic states (or spins) whereas the grey ones are unmarked (neutral) states.

 $\bar{\rho} = 0.6$ and $\bar{m} = 1$. We note that by choosing extreme values of density and magnetisation for the initial state may drive the equations to a local minimum and that Eqs. (15) include terms which diverge for $\rho \to 0$, and which are therefore prone to cause numerical blow-up in the swollen phase. Finally, by plotting the long-time steady state value of ρ and ϕ for each combination of $(\alpha_{\rho}, \alpha_{\phi})$ we can compile a non-equilibrium phase diagram which is shown in Fig. 4.

Outside the equilibrium line $\alpha_{\rho} = \alpha_{\phi}$, a new phase characterized by a density greater than zero but a small total magnetisation ($0 < \rho < 1$, $\phi^2 = 0$) is observed. This phase transition is reminiscent of a collapsed conformation in which the epigenetic marks are heterogeneously distributed and we thus dub this phase semi-Compact-Disordered (CD). Biologically, this phase may be similar to the one identified *in vivo* with "gene deserts" in which no epigenetic mark is clearly dominating (also called "black" chromatin in Ref. [27, 43] where it was first discovered). These regions have been recently show to display compact conformations [44] but the mechanisms behind such collapse are still unclear. Our theory suggests that even in absence of a uniform epigenetic pattern, these regions may assume collapsed conformations in the limit in which the interaction between similar epigenetic marks is large enough to overcome the entropic penalty due to polymer folding.

Through BD simulations we also observe a fourth phase that is not obtained through the model A equations (see Sec. IV and Fig. 6): one with uniform epigenetic coloring but swollen conformations. The lack of this phase in our theoretical phase diagram in Fig. 4 may be due to the approximation of almost space-filling (Hamiltonian) walks that we used earlier to derive the free energy in Eq. (13). Additionally, it should be noted that our field theory cannot fully resolve the chain structure of the polymer backbone and thus is not expected to sustain a large magnetisation at low density. On the other hand, the polymer simulations presented in the next section fully account for chain topology; one may thus argue that this constraint is sufficient to create a local uniformity of epigenetic marks, even at low polymer density as in the swollen phase.

IV. BROWNIAN DYNAMICS SIMULATIONS OF MAGNETIC POLYMERS

A. Model

To validate our findings within a more realistic polymer framework, we perform Brownian Dynamics (BD) simulations of chromosomes with dynamic epigenetic marks. We model chromosomes as semi-flexible chains with beads of size σ [45] and mark each bead with a "spin" or epigenetic state $q = \{-1, 0, 1\}$. Such a co-polymer model for chromosomes has been successfully employed in the literature [11, 14]. We simulate the dynamics of each of the polymer segments via a Langevin equation at the temperature T_L , i.e.

$$m\frac{d^2\boldsymbol{r}}{dt^2} = -\nabla U - \gamma \frac{d\boldsymbol{r}}{dt} + \sqrt{2k_B T_L}\boldsymbol{\eta}(t), \qquad (16)$$

where \boldsymbol{r} is the position of the bead, γ describes the bead friction and $\boldsymbol{\eta}$ is a noise satisfying the fluctuationdissipation theorem [46]. The potential U is a contribution of three force fields which model steric interactions, chain connectivity and rigidity. More specifically we consider $U = U_{LJ} + U_H + U_K$, where

$$U_{\rm K} = \frac{k_B T_L \ell_P}{\sigma} \sum_{i=1}^{M-2} \left(1 - \frac{\boldsymbol{u}_i \cdot \boldsymbol{u}_{i+1}}{|\boldsymbol{u}_i| |\boldsymbol{u}_{i+1}|} \right) , \qquad (17)$$

is the bending potential with $u_j \equiv r_{j+1} - r_j$. In this equation, ℓ_P is identified with the persistence length of the chain, here set to $\ell_P = 3\sigma \simeq 90$ nm to roughly match that of chromatin [47]. Then to model the connectivity of the chain we set springs between consecutive beads

$$U_{\rm H} = \sum_{i=1}^{M-1} \frac{k_h}{2} \left(|\boldsymbol{r}_i - \boldsymbol{r}_{i+1}| - r_0 \right)^2 \,, \tag{18}$$

where k_h is a spring constant set to $k_H = 200\epsilon = 200k_BT_L$ and $r_0 = 1.1\sigma$. The key force field is the Lennard-Jones potential which modulates the repulsive/attractive interaction between beads. Importantly, we choose this pairwise force field to depend on the spins of the interacting pair as follows:

$$U_{\rm LJ}(r;q_i,q_j) = \frac{4\varepsilon(q_i,q_j)}{\mathcal{N}} \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 - U_0(r_c) \right],\tag{19}$$

for $r < r_c(q_i, q_j)$ and 0 otherwise; in this equation, U_0 is an auxiliary function which ensures that $U_{\rm LJ}(r_c; q_i, q_j) =$ 0 and \mathcal{N} is a normalisation so that the minimum of the potential is at $-\varepsilon$.

The essential parameter here is the cutoff $r_c(q_i, q_j)$ which is q-dependent and set so that beads with the same spin value, or epigenetic mark, attract one another whereas beads with different marks (i.e., different q) interact only repulsively. This is done by choosing:

$$r_c(q_i, q_j) = \begin{cases} 2^{1/6} \sigma \text{ if } q_i \neq q_j \text{ or } q_i = q_j = 0\\ 1.8\sigma \text{ if } q_i = q_j \neq 0. \end{cases}$$
(20)

With these choices the pair interaction between equally colored (but not neutral) beads displays a dip which models short range attraction, whereas it is fully repulsive (no dip) for differently colored (or neutral) beads. Finally, the free parameter $\varepsilon(q_i, q_j)$ is set to ε if $q_j = q_j \neq 0$ and 1 otherwise. These choices mimic a ferromagnetic interaction (with strength ε) which tend to align marks that are nearby or to attract beads with similar marks.

We then evolve the equations of motion for each bead in the system using fixed-volume and constant-temperature Brownian dynamics (BD) simulations (NVT ensemble). The simulations are run within the LAMMPS engine [48] and the equations of motion are integrated using a velocity Verlet algorithm, in which all beads are weakly coupled to a heat bath with friction $\gamma = \tau_B^{-1}$ where $\tau_B = 3\pi\eta\sigma^3/k_BT$ is the self-diffusion (Brownian) time of a bead of size σ moving in a solution with viscosity η . Finally, the integration time step is set to $\Delta \tau = 0.01 \tau_B$. The polymer dynamics subject to thermal fluctuations is then interleaved with the recolouring dynamics of the polymer beads which we now describe.

Each bead along the chain can change its epigenetic state, or spin, at rate $\kappa_s = \tau_R^{-1}$, or in other words, on average every τ_R steps each bead is picked once and an attemp is made to change its color into a different one. If the move lowers the energy of the system we accept it, otherwise we assign an acceptance probability $p = e^{-\Delta U/k_B T_R}$ where ΔU is the difference between the system energy after and before the move.

It is important to notice that the dynamics of the chain is subject to thermal fluctuations controlled by the temperature T_L , whereas the Metropolis algorithm on the beads recolouring is weighted by an effective temperature T_R with, a priori, $T_L \neq T_R$. One can show that this condition breaks detailed balance by constructing a Kolmogorov loop over some of the states of the system as shown in Fig. 5:



FIG. 5. Detailed Balance is Broken when $T_L \neq T_R$. This figure shows a Kolmogorov loop over some states of the polymer. The product of the transition probabilities in the clockwise loop gives $p(\Delta U, T_L)$ whereas the counter-clockwise loop gives $p(\Delta U, T_R)$. The Kolmogorov criterion for obeying detailed balance is that the two must be equal, which is true only when $T_L = T_R$.

the product of the transitions in the clockwise direction is equal to $p(\Delta U, T_R) = e^{-\Delta U/k_B T_R}$ whereas the one over the counter-clockwise loop to $p(\Delta U, T_L) = e^{-\Delta U/k_B T_L}$. The Kolmogorov criterion states that detailed balance is obeyed only if the two are equal, which is true only if $T_L = T_R$.

As mentioned above, there is no biological constraint for which T_L should be equal to T_R ; in fact, chromatin and epigenetic mark dynamics are very different biophysical processes. In our model, this is captured by the fact that their effective temperatures are independently tuned.

Another possible way to violate detailed balance in this model is by introducing a fourth bead type that cannot be magnetised (we call this the "off" state as opposed to the standard "on" state). By randomly switching off beads with any q and activating "off" beads to one particular state (say q = 0), one can set up a current in the epigenetic states and therefore break time-reversal symmetry (and detailed balance). Biologically, this switching between on and off states may correspond to chromatin regions whose assembly in nucleosomes is transiently disrupted and cannot bear epigenetic marks or to proteins that can switch between different conformations due to ATP-binding and hydrolysis or phosphorylation [37, 49] and it is explored in more detail in the companion paper [25].

Finally, it should be noted that the correspondence between PDEs (such as the ones derived and numerically solved above) and BD is qualitative and we do not infer BD parameters from PDEs or viceversa. For instance, the proposed PDEs do not take into account polymer backbone, self-avoidance and noise (which are instead fully accounted for in BD simulations). On the other hand, the



FIG. 6. **BD** simulations of single chromosomes out-of-equilibrium. A-D Kymographs showing the evolution of the epigenetic marks along the polymer in time (insets show typical conformations). In counterclockwise order: A $T_L = 0.5$, $T_R = 3$; B $T_L = 1.5$, $T_R = 1.5$; C $T_L = 0.9$, $T_R = 0.9$; D $T_L = 2$, $T_R = 0.1$. All temperatures are in units of the room temperature $T^* = 300K$. E-F Time evolution of magnetisation $\langle m^2 \rangle^{1/2}$ (E) and radius of gyration $\langle R_g^2 \rangle$ (F) averaged over several independent simulations. Note the, in the long time limit, both observables reach a well defined steady state whose value depend on the pair (T_L, T_R) . The boxes around (A-D) match the color scheme in panels E-F.

measured observables are the same: local concentration of chromatin and magnetisation, can be considered to define different phases and compile a phase diagram for BD simulations (see below) which can be qualitatively compared with the results from PDEs (see also Ref. [22]).

B. Results for Isolated Magnetic Polymers

In order to verify that the theories obtained in Eq. (13) and Eqs. (15) predict behaviours that are in agreement with more refined models accounting for chain connectivity and self-avoidance constraints, we first perform Brownian Dynamics simulations of the magnetic polymer model described in the previous section considering dilute conditions – i.e., one chain in free space.

We always initialise the system from a state in which a polymer is in a self-avoiding walk conformation and with random epigenetic marks. To obtain stable states of the system we monitor the evolution of the radius of gyration R_g of an N-beads long chain, defined as

$$R_g^2 = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} |\mathbf{r}_i - \mathbf{r}_j|^2$$
(21)

and of its epigenetic magnetisation m, defined as

$$m = \frac{n(1) - n(-1)}{N}, \qquad (22)$$

where n(q) is the number of beads with epigenetic state q. Finally, we assign a state based on the steady values attained by these two observables, which can be respectively related to ρ and ϕ^2 used in the theories above. Examples of BD simulations of single magnetic polymers are reported in Fig. 6 where kymographs show the evolution of q along the chain index as a function of time. We also report typical conformations and average over different replicas.

For the equilibrium case (Fig. 3) the key parameter is the depth of the Lennard-Jones attraction between equal non-neutral marks which is set to ε . As shown in Figure 3 and Fig. 6(B-C), as we vary ε (with $T_L = T_R = 1$ in units of the room temperature $T^* = 300K$) we find phases that are matching the ones expected from the free energy in Eq. (13), separated by a first order transition at the critical point $\varepsilon_c \simeq 0.9$ [20]. It should be noted that these findings (and the ones below) are robust with respect to the choice of recolouring rate τ_R^{-1} and initial conditions.

To verify the non-equilibrium case reported in Fig. 4, we fix $\varepsilon = 1$ and tune T_L and T_R independently. As mentioned above, we find the compact-disordered phase predicted by the theory but we also find a swollen-ordered phase as shown in Fig. 6 and Refs. [20, 22]. We argue that this is because, as mentioned above, Eqs. (15) are prone to numerical instabilities in the regime $\rho \to 0$ and are obtained assuming a space-filling curve (Hamiltonian walk).

Nevertheless, it is remarkable to notice how robust and generic our findings are, as we obtain very similar conformations in both the analytical model and BD simulations. Deriving a precise mapping between the parameters of our theory, such as α_{ϕ} and α_{ρ} and the ones of the BD simulations such as T_L and T_R or ϵ is complicated, as the underlying processes are very different. For instance, the former is



FIG. 7. **BD** simulations of magnetic Polymer Melts. Time evolution of a magnetic polymer melt with N = 50 chains each M = 256 beads long. Panel **A** refers to a system with monomer density $\rho = NM/V = 0.8\sigma^{-3}$ and attraction strength $\varepsilon = k_B T_L$, while panel **B** to one with $\rho = 0.1\sigma^{-3}$ and $\varepsilon = 0.86k_B T_L$. In both cases $T_R = T_L$.

based on continuum equations, whereas the latter on the diffusion of discrete beads that are linked together on a chain. Moreover, the Time-Dependent Landau-Ginzburg equations above are solved at zero temperature whereas the BD simulations are run through a Langevin equation accounting for thermal fluctuations and the recolouring dynamics through a Monte-Carlo scheme. In spite of all this, we observe similar regimes thus suggesting that the uncovered physics is not system-dependent and therefore universal.

C. Results for Magnetic Polymer Melts

In our companion paper [25] we present a theory for how epigenetic marks can drive the compartmentalisation of the genome in the nucleus. This theory can also be checked with more refined BD simulations of magnetic polymers. The difference with respect to the set up described above is that now the genome in the cell nucleus should be modelled as a melt of, rather than isolated, magnetic polymers.

To do this we prepare linear chains as random walks in a box with periodic boundary conditions and reduce the box size until the desired monomer density $\rho = NM/L^3$ is attained. We typically consider N = 50 polymers with M = 256 beads each and the range of parameters employed are $\rho = 0.1 - 0.8 \sigma^{-3}$ and $\varepsilon/k_BT_L = 0.75$ -1.1. Snapshots of this set up and its evolution in time for two choices of ρ and ε are reported in Fig. 7. In the first case ($\rho = 0.8\sigma^{-3}$, $\varepsilon = k_BT_L$) we observe bicontinuous spanning clusters that merge into one without instabilities in the density, as predicted by our mean-field theory (see Ref. [25]) for quenches outside the coexistence region. In the second, ($\rho = 0.1\sigma^{-3}$, $\varepsilon = 0.86k_BT_L$) we observe a phase separation of the system in dense epigenetically ordered regions surrounded by unoccupied/sparse regions again in line with our theory for quenches within the coexistence region. It should be noted that both cases are in equilibrium as $T_L = T_R$. More examples are given in the companion paper [25].

This magnetic polymer melt can be extended to nonequilibrium by adding a switching process as described above. Specifically, we can turn off (or inactivate) a bead at rate σ_i and re-activate any "off" bead at rate σ_a by changing its state to neutral (q = 0). This choice sets up a current in the epigenetic states which causes the system to violate details balance and time-reversal symmetry (see Ref. [25]). We have qualitatively explored this system for selected parameter choices in [25]. A more systematic analysis would require large-scale computer simulations and will be pursued elsewhere.

V. CONCLUSIONS

In this work we have proposed a field theoretical approach to study the interplay between the dynamics of epigenetic marks along chromosomes *in vivo* and their spatial 3D organization. By mapping the system onto a model of magnetic polymers whose spin variables describe the epigenetic marks, we have analytically established that the transition between the swollen epigenetically disordered phase and the compact epigenetically coherent one is first order. The discontinuous nature of the transition, confirmed by Brownian dynamics simulations of a more realistic model of magnetic polymers, is a genuine product of the competition between the collective dynamics of epigenetic marks and the chain conformation and provides a mechanism of bistable epigenetic switch which endows the system with memory. It is interesting to note that in our companion paper [25], where we study a model for the full nucleus, we obtain instead a continuous transition between uniform disordered and uniform ordered phases. We argue that this is due to the fact that, while the model analysed here for single chromosomes is globally non-conserved in both density and magnetisation, the model for the full (interphase) nucleus has to conserve the overall density of DNA. The biological implications for this qualitative difference is that there may be additional ingredients needed to retain memory of epigenetic and conformational states in a nucleus which is homogeneously filled with chromatin.

By allowing the time-dependent Ginzburg-Landau equations of the model to follow non-equilibrium dynamics, we also found a third (non-equilibrium) phase. This is characterized by compact states in which the epigenetic marks are incoherently distributed and is reminiscent of gene deserts (or "black" chromatin [27]) observed in *Drosophila*. Finally, by using the corresponding BD simulations (with broken detailed balance) we have been able to perform a wider exploration of the non-equilibrium phase diagram and to observe the predicted compact disordered phase as well as a new phase characterized by extended chain conformations that are coherently coloured.

These results and the ones obtained with similar techniques but at the scale of the full nucleus (see [25]) show how statistical mechanics and field theory approaches on models of magnetic polymers can contribute to pinpoint the general multiscale physical mechanisms that govern the interplay between epigenetic spreading and genomic organization in the nucleus. In particular, our strategy allows us to weakly push the system out of equilibrium and to study this problem by using a (semi-)analytical framework (Figs. 2-4); this would be impossible (or a lot harder) to do in a far-from-equilibrium model without an underlying free energy. In the future, it would be interesting to explore other models tailored to capture specific non-equilibrium processes such as the disruption of chromatin structure by polymerase, or remodelling during mitosis. The looping mediated by slip-link-like proteins such as cohesins and condensins might also be an important ingredient to include [50–52] while the introduction of quenched spins that seed domain formation - as in genomic bookmarking [22] - can be easily incorporated at a field-theoretic level.

We hope that technical advances will soon make it possible to perform experiments with epigenetic marks on reconstituted chromatin *in vitro* [53]; this should give more hints on how to extend polymer magnetic models to improve our understanding of both epigenetic spreading processes [21, 53] and the generic interplay between different read-write protein complexes.

ACKNOWLEDGEMENTS

We thank the European Research Council (ERC CoG 648050 THREEDCELLPHYSICS) for funding. DMi and EO would also like to acknowledge the networking support by EUTOPIA (CA17139).

- [1] C. H. Waddington, Nature **150**, 563 (1942).
- [2] B. Alberts, A. Johnson, J. Lewis, D. Morgan, and M. Raff, Molecular Biology of the Cell (Taylor & Francis, 2014) p. 1464.
- [3] R. Cortini, M. Barbi, B. R. Care, C. Lavelle, A. Lesne, J. Mozziconacci, and J.-M. Victor, Rev. Mod. Phys. 88, 1 (2015).
- [4] W. W. C. Tang, S. Dietmann, N. Irie, H. G. Leitch, V. I. Floros, C. R. Bradshaw, J. A. Hackett, P. F. Chinnery, and M. A. Surani, Cell **161**, 1453 (2015).
- [5] S. Rulands, H. J. Lee, S. J. Clark, O. Stegle, B. D. Simons, and W. R. Correspondence, , 63 (2018).
- [6] D. Crews, Evolution and Development 5, 50 (2003).
- [7] U. R. Ernst, M. B. Van Hiel, G. Depuydt, B. Boerjan, A. De Loof, and L. Schoofs, Journal of Experimental Biology 218, 88 (2015).
- [8] S. I. S. Grewal, S. I. S. Grewal, and D. Moazed, Science 798, 798 (2011).
- [9] A. N. Boettiger, B. Bintu, J. R. Moffitt, S. Wang, B. J. Beliveau, G. Fudenberg, M. Imakaev, L. A. Mirny, C.-t. Wu, and X. Zhuang, Nature **529**, 418 (2016).
- [10] C. A. Brackley, S. Taylor, A. Papantonis, P. R. Cook, and D. Marenduzzo, Proc. Natl. Acad. Sci. USA 110, E3605 (2013).
- [11] C. A. Brackley, J. Johnson, S. Kelly, P. R. Cook, and D. Marenduzzo, Nucleic Acids Res. 44, 3503 (2016).

- [12] C. A. Brackley, J. M. Brown, D. Waithe, C. Babbs, J. Davies, J. R. Hughes, V. J. Buckle, and D. Marenduzzo, Genome Biol. 17, 31 (2016).
- [13] M. Barbieri, M. Chotalia, J. Fraser, L.-M. Lavitas, J. Dostie, A. Pombo, and M. Nicodemi, Proc. Natl. Acad. Sci. USA 109, 16173 (2012).
- [14] D. Jost, P. Carrivain, G. Cavalli, and C. Vaillant, Nucleic Acids Res. 42, 1 (2014).
- [15] I. B. Dodd, M. a. Micheelsen, K. Sneppen, and G. Thon, Cell **129**, 813 (2007).
- [16] M. A. Micheelsen, N. Mitarai, K. Sneppen, and I. B. Dodd, Phys. Biol. 7, 026010 (2010).
- [17] D. Jost, Phys. Rev. E 89, 1 (2014).
- [18] S. Berry, C. Dean, and M. Howard, Cell Syst. 4, 445 (2017).
- [19] T. Garel, H. Orland, and E. Orlandini, EPJ B 268, 261 (1999).
- [20] D. Michieletto, E. Orlandini, and D. Marenduzzo, Phys. Rev. X 6, 041047 (2016).
- [21] D. Michieletto, E. Orlandini, and D. Marenduzzo, Sci. Rep. 7, 14642 (2017).
- [22] D. Michieletto, M. Chiang, D. Coli, A. Papantonis, E. Orlandini, P. R. Cook, and D. Marenduzzo, Nucleic Acids Res. 46, 83 (2018).
- [23] N. Haddad, D. Jost, and C. Vaillant, Chromosome Research 25, 35 (2017).

- $[24]\,$ D. Jost and C. Vaillant, Nucleic Acids Res. ${\bf 46},\,2252$ (2018).
- [25] D. Michieletto, D. Coli, D. Marenduzzo, and E. Orlandini, Phys. Rev. Lett. xx, xx (2019).
- [26] E. Lieberman-Aiden, N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. a. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. a. Mirny, E. S. Lander, and J. Dekker, Science (80-.). **326**, 289 (2009).
- [27] G. J. Filion, J. G. van Bemmel, U. Braunschweig, W. Talhout, J. Kind, L. D. Ward, W. Brugman, I. J. de Castro, R. M. Kerkhoven, H. J. Bussemaker, and B. van Steensel, Cell 143, 212 (2010).
- [28] A. G. Larson, D. Elnatan, M. M. Keenen, M. J. Trnka, J. B. Johnston, A. L. Burlingame, D. A. Agard, S. Redding, and G. J. Narlikar, Nature 547, 236 (2017).
- [29] S. F. Pinter, R. I. Sadreyev, E. Yildirim, Y. Jeon, T. K. Ohsumi, M. Borowsky, and J. T. Lee, Genome Res. 22, 1864 (2012).
- [30] N. A. Hathaway, O. Bell, C. Hodges, E. L. Miller, D. S. Neel, and G. R. Crabtree, Cell 149, 1447 (2012).
- [31] F. Y. Wu, Reviews of Modern Physics 54, 235 (1982).
- [32] B. Duplantier, Phys. Rev. B **35**, 5290 (1987).
- [33] H. Orland, C. Itzykson, and C. de Dominicis, Journal de Physique Lettres 46, 353 (1985).
- [34] A. M. Nemirovsky, J. Dudowicz, and K. F. Freed, Journal of Statistical Physics 67, 395 (1992).
- [35] M. Nicodemi and A. Prisco, Phys. Rev. Lett. 98, 108104 (2007).
- [36] S. Henikoff and J. M. Greally, Curr. Biol. 26, R644 (2016).
- [37] D. Michieletto and N. Gilbert, Curr. Opin. Cell Biol. 58, 120 (2019).
- [38] N. Festuccia, A. Dubois, S. Vandormael-Pournin, E. G. Tejeda, A. Mouren, S. Bessonnard, F. Mueller, C. Proux, M. Cohen-Tannoudji, and P. Navarro, Nat. Cell Biol. 18, 1139 (2016).

- [39] S. Wang, M. Liu, and Y. Dong, J. Phys. Condens. Matter 25, 184007 (2013).
- [40] D. Egli, G. Birkhoff, and K. Eggan, Nat. Rev. Mol. Cell. Biol. 9, 505 (2008).
- [41] P. Hohenberg and B. Halperin, Rev. Mod. Phys. 49, 435 (1977).
- [42] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics*, Vol. c (Cambridge University Press, 2007).
- [43] T. Sexton, E. Yaffe, E. Kenigsberg, F. Bantignies, B. Leblanc, M. Hoichman, H. Parrinello, A. Tanay, and G. Cavalli, Cell **148**, 458 (2012).
- [44] Q. Szabo, D. Jost, J. M. Chang, D. I. Cattoni, G. L. Papadopoulos, B. Bonev, T. Sexton, J. Gurgo, C. Jacquier, M. Nollmann, F. Bantignies, and G. Cavalli, Sci. Adv. 4, 1 (2018).
- [45] A. Rosa and R. Everaers, PLoS Comp. Biol. 4, 1 (2008).
- [46] K. Kremer and G. S. Grest, J. Chem. Phys. 92, 5057 (1990).
- [47] M. Socol, R. Wang, D. Jost, P. Carrivain, C. Vaillant, E. Le Cam, V. Dahirel, C. Normand, K. Bystricky, J.-M. Victor, O. Gadal, and A. Bancaud, Nucleic Acids Research 47, 6195 (2019).
- [48] S. Plimpton, J. Comp. Phys. 117, 1 (1995).
- [49] C. A. Brackley, B. Liebchen, D. Michieletto, F. Mouvet, P. R. Cook, and D. Marenduzzo, Biophys J. 112, 1085 (2017).
- [50] E. Orlandini, D. Marenduzzo, and D. Michieletto, Proc. Natl. Acad. Sci., 201815394 (2019).
- [51] G. Fudenberg, M. Imakaev, C. Lu, A. Goloborodko, N. Abdennur, and L. A. Mirny, Cell Rep. 15, 2038 (2016).
- [52] C. Brackley, J. Johnson, D. Michieletto, A. Morozov, M. Nicodemi, P. Cook, and D. Marenduzzo, Phys. Rev. Lett. **119**, 138101 (2017).
- [53] Y. Tauran, M. Kumemura, M. C. Tarhan, G. Perret, F. Perret, L. Jalabert, D. Collard, H. Fujita, and A. W. Coleman, Scientific Reports 9, 1 (2019).