Spatiotemporal Patterns of Active Epigenetic Turnover

Fabrizio Olmeda,[1,2,3,*] Misha Gupta,[1,4] Onurcan Bektas,[2] and Steffen Rulands[1,2,†]
[1]*Max-Planck-Institute for the Physics of Complex Systems, Noethnitzer Strasse 38, 01187 Dresden, Germany*
[2]*Ludwig-Maximilians-Universität München, Arnold-Sommerfeld-Center for Theoretical Physics, Theresienstrasse 37, 80333 München, Germany*
[3]*Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuberg, Austria*
[4]*Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA*

(Received 21 May 2025; accepted 19 December 2025; published 9 February 2026)

DNA methylation is a primary layer of epigenetic modification that plays a pivotal role in the regulation of development, aging, and cancer. The concurrent activity of opposing enzymes that mediate DNA methylation and demethylation gives rise to a biochemical cycle and active turnover of DNA methylation. While the ensuing biochemical oscillations have been implicated in the regulation of cell differentiation, their functional role and spatiotemporal dynamics are unknown. In this work, we demonstrate that chromatin-mediated coupling between these local biochemical cycles can lead to the emergence of phase-locked domains, regions of locally synchronized turnover activity, whose coarsening is arrested by genomic heterogeneity. We introduce a minimal model based on stochastic oscillators with constrained long-range and nonreciprocal interactions, shaped by the local chromatin organization. Through a combination of analytical theory and stochastic simulations, we predict both the degree of synchronization and the typical size of emergent phase-locked domains. We qualitatively test these predictions using single-cell sequencing data. Our results show that DNA methylation turnover exhibits surprisingly rich spatiotemporal patterns that may be used by cells to control cell differentiation.

DOI: 10.1103/89bj-79g5

I. INTRODUCTION

Epigenetics collectively describes biomolecular processes that regulate cell behavior beyond what is encoded in the DNA sequence. These processes involve the folding of the DNA and chemical modifications of the DNA and histone tails. DNA methylation primarily affects cytosines next to guanines, termed CpG pairs. The positioning of DNA methylation marks in the genome has been shown to be relevant for the assignment of cell types during embryonic development [1,2]. It is also closely associated with aging [3], and alterations in related enzymes are one of the hallmarks of blood cancer [4]. DNA methylation marks are established by a family of DNMT3 enzymes, and they are maintained during DNA replication by DNMT1 enzymes, among others. DNA methylation marks can also be actively removed through a chain of biochemical reactions involving TET enzymes. TET enzymes modify the methylated cytosines (5mC) by oxidizing 5mC to hydroxymethyl-cytosine (5hmC), then to formyl-cytosine (5fC), and finally to carboxyl-cytosine (5caC) [1,5,6] (Fig. 1).

Paradoxically, in several biologically relevant contexts, cells express antagonistic enzymes that methylate and demethylate the DNA, namely members of the protein families DNMT3 and TET [1,7]. These contexts include the priming of pluripotent cells for differentiation in early development [5,8] and the differentiation of hematopoietic stem cells [9]. This coexpression leads to a biochemical and largely irreversible cycle, in which a CpG undergoes cyclic chemical turnover involving the conversion of cytosines (C) to methylated cytosines (5mC), then to hydroxymethylated cytosines (5hmC), and via multiple intermediary steps back to the unmodified cytosine. The periodicity of this oscillation is approximately two hours, and the oscillation phase is coupled genome-wide to a limited degree [8]. Oscillatory DNA methylation turnover has been specifically linked to enhancer regions [5,8] and to inducible genes targeted by the estrogen receptor α [10,11]. Active turnover of DNA methylation seems to be necessary for the regulation of cell differentiation in development [5].

Experimental work on active DNA methylation turnover has mainly focused on DNA sequencing technologies such as bisulfite sequencing [12]. State-of-the-art technologies cover around one in five CpGs in a given cell and cannot distinguish between more than two chemical modifications. Methods that can distinguish between C, 5mC, and 5hmC in single cells have orders of magnitude lower resolution [13]. Because sequencing also relies on destructive sampling of the cells and DNA, dynamic information about DNA methylation turnover can only be inferred indirectly. Therefore, despite active turnover of DNA methylation being associated with relevant biological scenarios, its spatiotemporal dynamics and biological function are unknown. Theoretical predictions are hence pivotal for understanding the spatiotemporal dynamics

*Contact author: fabrizio.olmeda@ista.ac.at
†Contact author: rulands@lmu.de

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

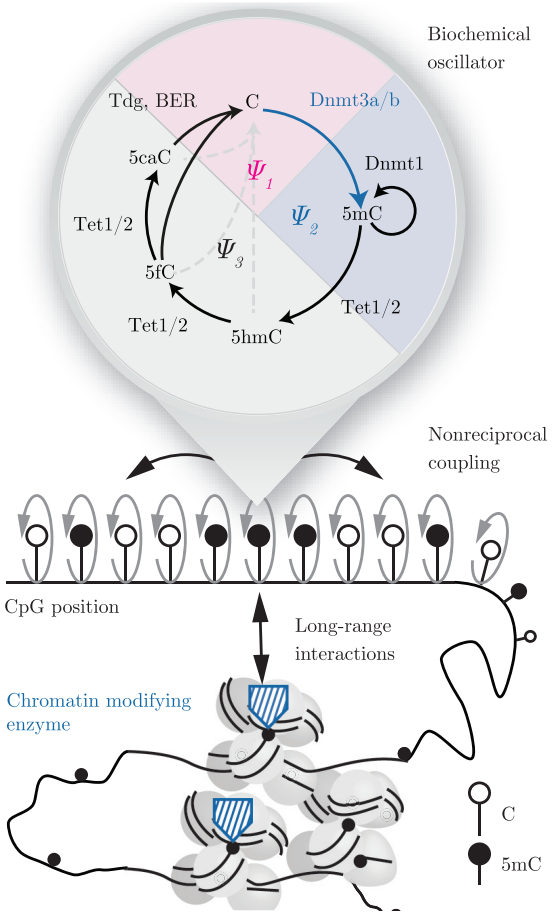

FIG. 1. The biochemistry of DNA methylation turnover involves a DNA methylation cycle that is divided into three phase intervals, ψ_1, ψ_2, ψ_3. CpGs in the first phase interact via restricted long-range couplings with CpGs that are in the second phase.

of DNA methylation turnover and inform targeted experimental approaches.

Here, we investigate the range of spatiotemporal behaviors that DNA methylation turnover can exhibit. We map the interplay between DNA methylation turnover and the geometrical configuration of chromatin onto a driven oscillator model with nonreciprocal and nonlocal interactions. We show that DNA methylation turnover leads to phase-locked domains that are globally weakly coupled. Using analytical and numerical calculations, we demonstrate that the degree of global synchronization increases with the coupling strength and then saturates, reaching a partially synchronized regime in the limit of infinite coupling strengths. Phase-locked domains coarsen over time much larger than a characteristic timescale, which is much larger than timescales of active epigenetic turnover for weak disorder in the frequencies of individual oscillators. We further show that genomic disorder stemming from the DNA sequence enhances local synchronization and leads to an arrest of coarsening and a constant size of phase-locked domains. Finally, we use sequencing data that simultaneously identify C, 5mC, and 5hmC states to qualitatively test key predictions of our theory.

II. RESULTS

We first derive a stochastic lattice model of DNA methylation turnover to investigate the potential spatiotemporal patterns emerging from these interactions. To this end, we consider a lattice of N sites describing CpG positions along a single chromosome. The state of the DNA methylation turnover cycle at each position, i, is described by a phase variable ϕ_i, which can take values in $\{1, \ldots, \Omega\}$, Fig. 1. The phase variable describes different configurations of the cytosine, such as the unmodified cytosine, the cytosine bound by a DNMT3 enzyme, the bound cytosine after the recruitment of the methyl group, and so on. The number of states, Ω, is given by the number of cytosine configurations with comparable lifetimes. The state of the system is then described by a vector containing the phase values at each site, $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_N)$. The phase at site i is cyclically advanced by 1 with a rate $\tilde{\omega}_i = \omega_i + k_i(\boldsymbol{\phi})$. This rate has a site-dependent, but constant in time, contribution ω_i, denoting the intrinsic rate of chemical conversions at that position i. The function $k_i(\boldsymbol{\phi})$ captures the contribution of the phases at other sites.

These interactions with other sites are local in physical space. The folding of the DNA and chromatin leads to long-range interactions along the one-dimensional DNA sequence. On the scale of thousands of base pairs, these interactions have been shown to correlate with DNA methylation [14]. Longer scale interactions through DNA loop formation or contacts between chromosomes do contribute, but with comparatively small statistical weight. The conformation of DNA and chromatin is reflected in enzyme binding rates, as the establishment of DNA methylation marks is associated with the local compaction of chromatin [15–17]. We describe the local structure of chromatin in terms of an exponent λ defining the inner contact probability, such that $k_i(\boldsymbol{\phi})$ decays like $|i-j|^{-\lambda}$. This exponent λ takes values like 2.1 for a self-avoiding random walk and 1/3 for space-filling chromatin. On the nanoscale, chromatin is organized into heterogeneous structures ("clutches") [18]. On the associated lengthscale, we therefore assume that long-range contacts are exponentially suppressed and $k_i(\boldsymbol{\phi})$ is exponentially cut off [19]. This cutoff also ensures the existence of the thermodynamic limit.

The local compaction of chromatin influences the chemical conversion rates in earlier steps of the cycle, such that interactions are nonreciprocal. For example, methylated sites recruit more DNMT3 enzymes and thereby increase conversion rates of unmethylated sites in the vicinity [15]. The reverse is generally not true. To explicitly capture changes in the biochemical cycle driven by enzyme kinetics, we define the set of chemical states associated with compacted chromatin as ψ_2. Biologically, these states correspond to the binding of DNMT3 enzymes and the subsequent methylation of CpG sites. Chemical states in ψ_2 induce an increase in the rate of state transitions in a set of states preceding ψ_2. We denote these preceding states by ψ_1 (Fig. 1). In the mean-field limit, i.e., neglecting spatial correlations, the interaction leads to a contribution to the transition rate of the form [15,20,21]

$$k_i(\boldsymbol{\phi}) = J \sum_{k=1,k\neq i}^{N} \mathcal{I}_{\phi_i,\psi_1}\mathcal{I}_{\phi_k,\psi_2}\frac{e^{-m|k-i|}}{|k-i|^{\lambda}}, \tag{1}$$

where J denotes the coupling strength between sites. The indicator function $\mathcal{I}_{A,B}$ is 1 if $A \in B$ and 0 otherwise. $m = N^{-1}\sum_{j=1}^{N}\mathcal{I}_{\phi_j,\psi_2}$ is then the fraction of sites in phase interval ψ_2. The DNA methylation dependence of the exponential cutoff in Eq. (1) is a consequence of screening due to a feedback between DNA methylation and chromatin [15,19]. The time evolution of the joint probability of finding a given lattice configuration $\boldsymbol{\phi}$ at time t follows the master equation (A1). Here, we consider the case that the coupling J is positive. This implies positive feedback, such that progression through the biochemical cycle is not rate-limited by the availability of enzymes. While in the context of histone-modifying enzymes such competition has been discussed [22], in the case of DNA methylation, *in vitro* experiments indicate that enzyme availability is not a limiting factor [23].

A. A continuum theory of DNA methylation turnover

We now ask under which conditions DNA methylation turnover can synchronize, and we characterize potentially emerging spatiotemporal structures. To gain an initial understanding, we first neglect the fact that CpG positions are unevenly spaced in the genome. To begin, we derive a description in terms of continuous phase variables. To this end, we separate the phase variable ϕ_i into a deterministic part, which scales as the number of states in the cycle, Ω, and a stochastic component, which scales as the square root of Ω (system size expansion) [24,25]. In doing so, we promote ϕ_i to a continuous variable that satisfies $\phi_i = \Omega\Phi_i(t) + \Omega^{1/2}\xi_i(t)$.

After substituting this expression into the time evolution equation for the probability density and collecting terms of equal order in Ω, Appendix A, we obtain a Langevin equation describing the time evolution of the phase at each site,

$$\frac{\mathrm{d}\phi_i}{\mathrm{d}t} = w_i + f_1(\phi_i)\sum_{k=1,k\neq i}^{N}\frac{Je^{-m|k-i|}}{|k-i|^{\lambda}}f_2(\phi_k) + \sqrt{2\omega_i}\xi_i(t). \quad (2)$$

Here, $\xi_i(t)$ is Gaussian white noise with zero mean and unit variance, $\langle\xi_i(t)\xi_j(t')\rangle = \delta(t-t')\delta_{i,j}$. The functions $f_{1,2}(\phi)$ are the continuous versions of $\mathcal{I}_{\phi,\psi_{1,2}}$, respectively. Equation (2) is similar in form to a stochastic version of the Kuramoto model, which has been extensively studied to understand synchronization phenomena of coupled oscillators [26]. In the case of DNA methylation turnover, the oscillators are coupled via nonreciprocal, restricted long-range interactions. We can therefore apply known methods for oscillators with a continuous phase [26,27]. An additional advantage of the continuous formulation of our model is the flexibility to explore different functional forms of f, which may in turn reveal novel dynamical features of stochastic oscillators. In the following section, we will use an exact, coarse-grained description of Eq. (2) (Appendix B) in order to investigate the emergence of locally synchronized states.

As a first step, we define a quantity that describes the degree to which CpG states synchronize systemwide. This quantity is given by the Kuramoto order parameter $r(t)$ [26], which is defined as

$$r(t)e^{i\psi(t)} = \frac{1}{N}\sum_{i=1}^{N}e^{i\phi_i(t)}. \quad (3)$$

Mathematically, the order parameter $r(t)$ is the amplitude of a complex number that represents the average phase of oscillators. It has a range that goes from 0, where all CpGs oscillate statistically independently, to 1, where all oscillators are perfectly synchronized. Intermediate values of $r(t)$ indicate partial synchronization, where a subset of oscillators share the same phase. Figure 2(a) shows representative simulations corresponding to intermediate values of $r(t)$: while local groups of sites oscillate in a synchronized manner, the global coupling between these groups is limited. We now aim to relate the order parameter r to the kinetic parameters of the model. To this end, we define the density of oscillators with phase ϕ, position z, and frequency ω at time t as $\rho(\phi,\omega,z,t)$. The time evolution of ρ follows a continuity equation derived by coarse-graining Eq. (2) (Appendix B),

$$\frac{\partial\rho(\phi,\omega,z)}{\partial t} = -\frac{\partial}{\partial\phi}[(\omega+\tilde{\nu})\rho(\phi,\omega,z,t)]. \quad (4)$$

This equation is a transport equation, and it describes the translation of the density ρ along the phase cycle with a velocity given by the constant term ω, and an additional contribution stemming from long-range interactions. This contribution to the phase velocity is given by

$$\tilde{\nu} = Jf_1(\phi)\left[\int\frac{g(\omega')e^{-mz'}}{|z'|^{\lambda}}f_2(\phi')\rho(\phi',\omega',z-z')\right], \quad (5)$$

where the triple integral goes over z',ω',ϕ', Appendix B. The function $g(\omega)$ denotes the distribution of intrinsic frequencies of DNA methylation turnover. It is determined by enzyme binding and unbinding rates, and by the availability of metabolites and chemical reaction products necessary for the conversion between different chemical states of the cytosine. Spectral analysis of a sequencing-based time course revealed that oscillation frequencies are region-specific but constrained to an overall period between 1 and 2 h [8]. Nevertheless, because enzyme activities in DNA methylation turnover depend sensitively on the sequence context of a given CpG position, we expect significant variability in the intrinsic frequencies.

The phase cycle velocity in Eq. (5) depends only on the global fraction of sites that carry DNA methylation marks (sites in the phase interval ψ_2), m, and the order parameter, r. As the average DNA methylation m and the Kuramoto order parameter r depend on the coarse-grained density ρ and vice versa, they must be obtained self-consistently. In the next section, we use an analytical solution for the distribution of phases and report the full analytical calculations in Appendixes C and D.

B. Coarsening of phase-locked domains

To investigate whether the coupling between CpGs gives rise to spatiotemporal structures, we numerically solved the deterministic part of Eq. (2). We set the initial conditions such that the phase variables at each CpG were drawn from a uniform distribution. We sampled their intrinsic frequencies ω_i from a δ distribution, thus neglecting genomic variability in enzyme activities. Figure 2(a) shows typical illustrative realizations for different coupling strengths, J. DNA methylation turnover at short times synchronizes in finite-size domains, Fig. 2(a), inset, meaning that within these domains, phase

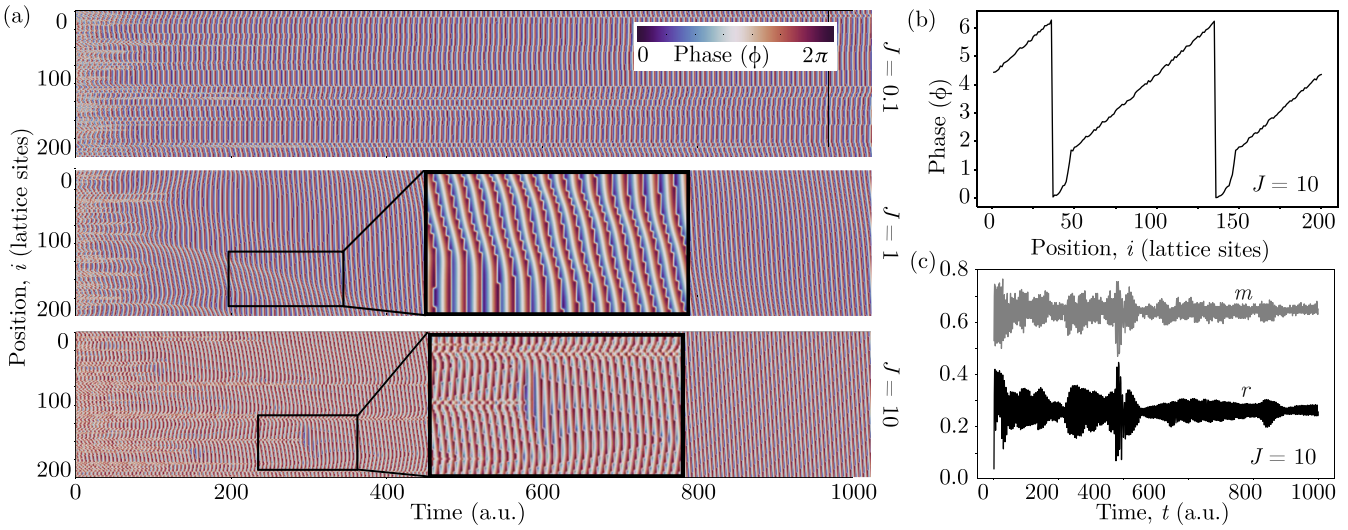

FIG. 2. (a) Heatmaps showing the spatiotemporal evolution of the deterministic part of Eq. (2) starting from random initial conditions and identical intrinsic frequencies for all sites, i.e., $\omega_i = 1$ for all i. The insets show enlarged heatmaps for a subset of sites. (b) Spatial distribution of phases after 10^6 time steps, highlighting the phase-locked structures for a high coupling strength ($J = 10$). (c) The Kuramoto order parameter r and average DNA methylation m exhibit fluctuations and oscillations before approaching a steady state ($J = 10$). Details of the numerical simulations are given in Appendix E.

differences between neighboring sites remain constant in time, which we refer to as phase locking. In this initial transient, biochemical oscillators are weakly synchronized systemwide.

After the initial transient, the long-term behavior is qualitatively independent of the coupling strength. Domains slowly coarsen over time, and at late times the system achieves systemwide phase-locking. This is apparent in the gradual disappearance of lines reflecting discrete phase differences in Fig. 2(a). Figure 2(b) shows the phase of each CpG site along the one-dimensional domain at the end of the simulations for coupling strength $J = 10$. It is evident that neighboring CpGs have a constant phase difference, and only a small number of sites deviate from this behavior. We explain this behavior later by analytical solutions of the model. To obtain a qualitative understanding of the global dynamics, we calculated the time evolution of the average DNA methylation and the degree of global coupling quantified by the Kuramoto order parameter, Fig. 2(c). Both quantities show strong fluctuations over time, and they ultimately converge. This is caused by the merging of phase-locked domains, resulting in abrupt changes in both quantities.

From Fig. 2(a) it is qualitatively clear that the coarsening dynamics of phase-locked domains depends nonmonotonically on the coupling strength: For small values of the coupling, $J = 0.1$, neighboring phase-locked domains coarsen slowly. However, for larger values of the coupling, $J = 1$, coarsening occurs faster. Interestingly, for $J = 10$, once larger domains are formed, their coarsening slows down again compared to the intermediate coupling strength, $J = 1$. These observations were consistent across 1000 simulations with randomly sampled initial conditions. In the following, we study how the specific couplings between CpGs shape the partial synchronization of DNA methylation domains. We then develop a method called strong-disorder renormalization that quantifies the coarsening time and its dependence on the interaction strength J.

We first investigate how the strength of the interaction J, which is controlled by the activity of enzymes, changes the average DNA methylation and global synchronization of the epigenetic states. To this end, we compute the stationary solution of Eq. (4) (Appendix C). We find that there are two possible solutions: Oscillators may be phase-locked at identical frequencies such that $\omega + \tilde{\nu} = 0$. Secondly, oscillators can rotate with frequencies given by $(\omega + \tilde{\nu})\rho = C(\omega)$, where $C(\omega)$ is a constant, which is determined by the normalization of the density ρ of oscillators [28]. In Appendix D, we solve for both branches of the solution and find an explicit integral solution for the fractions of sites carrying DNA methylation marks. In the limits of either strong or weak interactions, these solutions can be evaluated explicitly: for $J \to 0$ we obtain $m = 1/2$, and for $J \to \infty$ we obtain $m = 2/3$. Intuitively, in the strong interaction limit, $J \to \infty$, oscillators that are in the phase interval being affected by interactions, ψ_1, spend an infinitesimally small amount of time in that phase interval. The density ρ of oscillators is effectively restricted in the phase domain $[\pi/2, 2\pi]$. The oscillators in this interval do not interact. Here, on average, $m = \int d\phi f_2(\phi)/(3\pi/2)$, which for our specific choice of f_2 gives the value $2/3$. We attribute the small discrepancy between numerical simulations and theoretical results [Fig. 3(a), inset] to finite-size effects and the formation of spatial structures. The Kuramoto order parameter, r, can be computed numerically from Eq. (D2). As m converges to a constant value in the long-term limit, r must be strictly smaller than 1, and full synchronization cannot be achieved in the limit of infinitely many CpGs, $N \to \infty$.

For intermediate values of the coupling strength J, analytical solutions are not feasible. We therefore perform numerical simulations. In Fig. 3(a) we show the analytical prediction for both DNA methylation and the Kuramoto order parameter,

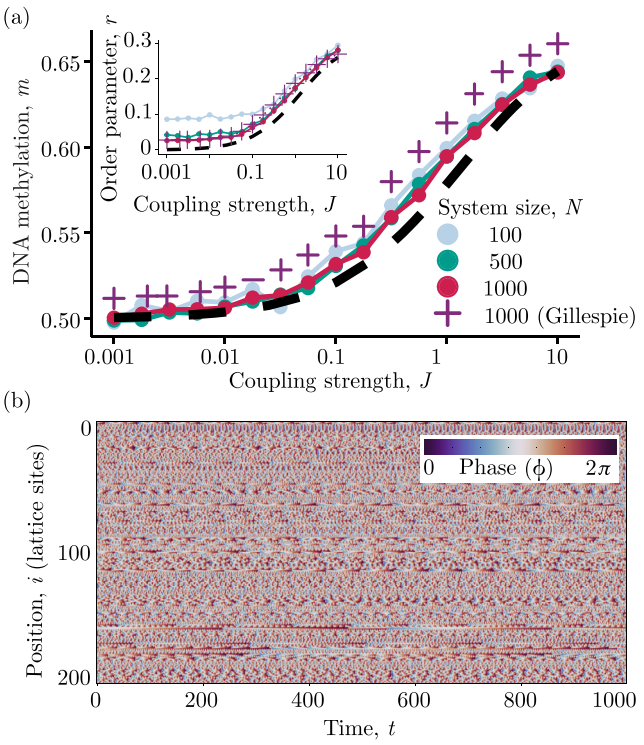

FIG. 3. (a) Kuramoto order parameter $\langle r \rangle$ and $\langle m \rangle$ (inset) as a function of the interaction strength J. The average is taken over 100 realizations of deterministic (points) and stochastic (crosses) numerical simulations of the model in Eq. (2). Different colors represent different values of the lattice size, N. Theoretical predictions from Eqs. (D5) and (4) are shown as a black dashed line. For the stochastic simulations, we set the number of states to 100. Here, the frequencies ω_i are random numbers drawn from an exponential distribution with unitary mean. (b) Long-term behavior of oscillators with exponential distribution of frequencies with intermediate-coupling strength, $J = 1$. Details of the numerical simulations are given in Appendix E.

compared to stochastic and mean-field lattice simulations for different lattice sizes (Appendix A). For the stochastic simulations, we varied the number of phase states between 10 and 100 with consistent results. For a very low number of states, we do not find any synchronization, which is analogous to the findings in Ref. [25]. This is due to the fact that there are not enough states that act as a source of interaction, such that the model is identical to a noninteracting system. In this case, a continuum description as in Eq. (2) is not valid. While the effective, empirical number of phase states depends on the specific values of the kinetic parameters of enzyme binding events and chemical reactions, the empirical observation of local synchronization in Ref. [8] suggests that the number of states is indeed high enough to facilitate effective interactions between oscillators.

As a next step, we now investigate the role of heterogeneity for synchronization. In DNA methylation turnover, there are two predominant sources of heterogeneity: First, the intrinsic frequency may differ between CpGs due to the dependence of enzyme activities on the genomic context [29–31]. Secondly, the strength of the couplings between pairs of neighboring sites depends on their genomic distance, which is strongly heterogeneous. We begin by focusing on understanding the effect of variability in the intrinsic frequencies, ω_i. Figure 3(b) shows simulations where the distribution of intrinsic frequencies is sampled from an exponential distribution. Unlike the case of equal frequencies, this heterogeneity prevents the formation of large phase-locked domains, but it yields smaller synchronized domains. However, these smaller domains do not coarsen over the same temporal scale as oscillators with identical frequencies. Taken together, our calculations and simulations reveal that DNA methylation turnover does not exhibit a phase transition from an asynchronous to a synchronous state, which is common in typical driven oscillator models [28]. Rather, we find that the degree of synchronization and DNA methylation increase continuously with the interaction strength, but with partial synchronization for all positive values of the coupling strength $J > 0$.

C. Effect of genomic disorder on phase locking

In the genome, the positions of CpGs are heterogeneously distributed, with many of them clustered in so-called CpG islands. Because the effective strength of interactions decays with genomic distance, Eq. (1), the coupling strength varies randomly with the genomic position i (quenched disorder). To investigate the role of such disorder on the degree of phase locking, we performed numerical simulations of Eq. (2), with distances between sites equal to those of different regions of chromosome 1 of the mouse genome, Figs. 4(a) and 4(b). We rescaled distances so that the average distance between neighboring CpGs is equal to 1. With genomic disorder, the coarsening dynamics of phase-locked domains are slower, Fig. 4(a), compared to the regular, one-dimensional lattice, Fig. 2(a). After approximately 1e4 oscillation periods, Fig. 4(c), the coarsening is arrested, resulting in stable, phase-locked structures, Fig. 4(b). In the steady state, domain boundaries are determined by the distribution of distances between neighboring CpGs: they tend to coincide with genomic regions in which the density of CpGs changes strongly, Fig. 4(d).

To investigate the coarsening dynamics of phase-locked domains, we now employ *strong disorder renormalization* of Eq. (2) [32,33], Fig. 4(e). This iterative renormalization procedure defines a coarse-graining scheme on blocks of neighboring CpGs. Initially, each block contains exactly one CpG. At each renormalization step, we select the pair of neighboring blocks with the closest genomic distance, and hence the strongest coupling strength. We then define a new block that incorporates the pair of chosen CpGs. We define the phase of the new block, ϕ_i, as the average phase of the coarse-grained blocks, weighted by the number of CpGs. The size of the coarse-grained block, N_i, is given by the sum of the number of CpGs of the blocks that were averaged, $N_i = n_i + n_{i+1}$, and its intrinsic frequency by the average weighted by the number of CpGs in each block, $\Omega_i = (n_i\omega_i + n_{i+1}\omega_{i+1})/(n_i + n_{i+1})$. We stop the renormalization procedure when the typical distance between blocks is greater than the average distance between CpG sites. After these renormalization steps, Eq. (2) is effectively approximated by a model with interactions between neighboring blocks.

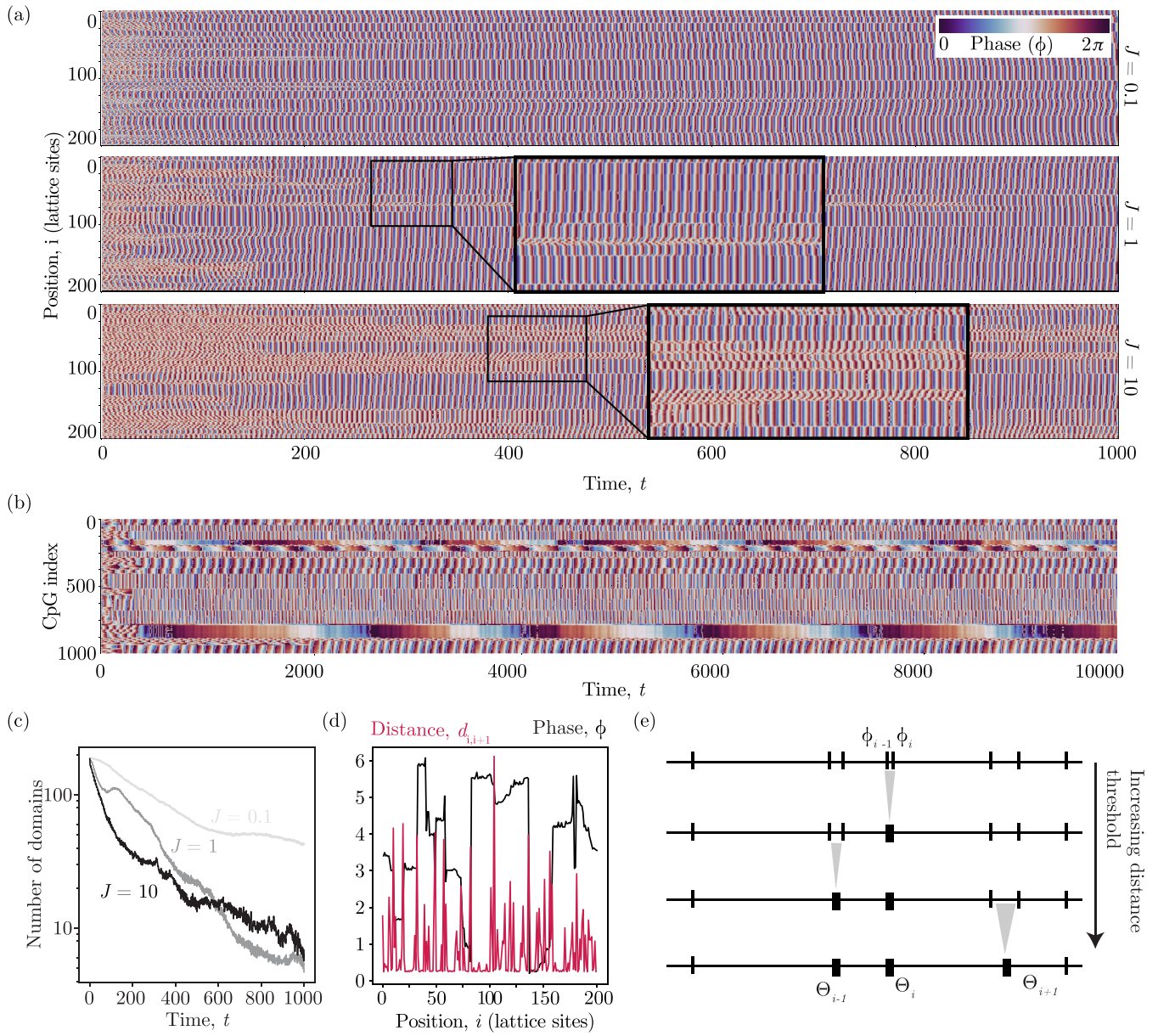

FIG. 4. (a) Numerical simulations of the deterministic part of Eq. (2) starting from random initial conditions and a δ distribution of intrinsic frequencies $\omega_i = 1$ for all i. The distances between neighboring sites are taken from the first 200 CpGs of Chromosome 1 of the mouse genome. For visualization, CpGs are indexed by their order in the genome, such that CpG index 1 refers to the first CpG, index 2 to the second, and so on. Color denotes the phase at a given CpG index and time. We tested different regions of the genome and found consistency in the results. (b) At times ten times larger than the one shown in (a), phase-locked domains have not yet coarsened. This is in accordance with our theoretical scaling result. Here we show a simulation with coupling strength $J = 10$, which would be expected to increase the coarsening dynamics compared to lower values of J. (c) Number of phase-locked domains over time for different coupling strengths, $J = 0.1, 1, 10$, for simulations with 1000 sites. Details on the inference of domain numbers are given in Appendix E. (d) Comparison of the distance to the nearest CpG and phase variable ϕ. Strong changes in the density of CpGs are associated with changes in the phase variable. Data points were taken at $1e4$ oscillation periods. (e) Schematic illustrating the strong-disorder renormalization-group scheme. At each step, we merge the closest neighboring blocks. We repeat this process until the distances between the chosen adjacent blocks are larger than the average distance between CpGs.

We now ask whether these distinct blocks coarsen over time, which would ultimately lead to systemwide synchronization. To estimate the typical time of coarsening, we evaluate how the phase difference between two adjacent coarse-grained blocks, $\phi_i - \phi_{i-1}$, evolves in time. If all oscillators have identical intrinsic frequencies, this coarse-grained phase difference decreases exponentially with a typical timescale given by

$$\tau_i \approx \frac{N_i}{Je^{-m}}\left(\frac{1}{N_i} + \frac{1}{N_{i-1}}\right)^{-1}. \tag{6}$$

This result shows that the coarsening time scales quadratically with the typical number of CpGs in a coarse-grained

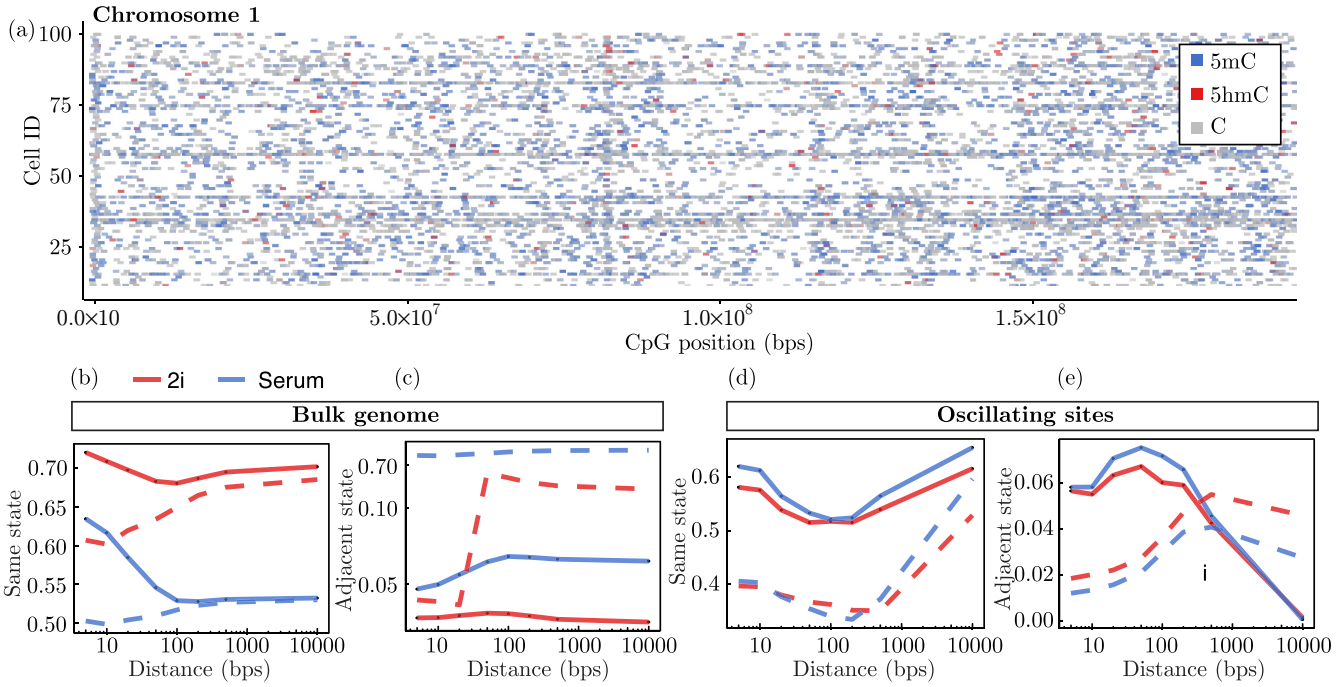

FIG. 5. (a) Representative experimental sequencing data, with individual CpG sites colored according to their measured state in the biochemical cycle. (b) Empirical probability of finding two consecutive CpGs with the same state in the biochemical cycle (solid line) compared to the case of statistical independence (dashed line). (c) Empirical probability of finding neighboring CpG states in adjacent states of the biochemical cycle (solid line, in this case C and 5mC). Dashed lines depict the prediction under the null hypothesis of statistical independence. Parts (d) and (e) show, respectively, the same analysis as in (b) and (c) but filtered for sites that putatively undergo DNA methylation turnover. The distance dependence of the probabilities is a consequence of systematic variations of the CpG density in the genome [34]. This also applies to the probability under the null hypothesis.

block. In particular, localized regions, such as CpG islands, which have approximately 100 CpGs, will be phase-locked for approximately $1e4$ oscillations for a value of Je^{-m} of order 1. This result is confirmed by our simulations (Fig. 4) for even smaller domains. For a typical oscillation period of 2 h, this translates to at least 27 months, which is longer than the typical lifetime of the model system that was used to experimentally determine the periodicity (mouse). Therefore, we expect that arrested phase-locked domains are stable across all biologically relevant timescales.

Our theoretical work shows that the dynamics of DNA methylation turnover exhibit behavior that is not predicted by driven oscillator models studied in the theoretical literature [28]. The Kuramoto model with long-range interactions [27] exhibits global synchronization if the coupling strength exceeds a threshold value. Here, nonreciprocal interactions lead to the formation of phase-locked domains which coarsen over time. This leads to global phase locking but not to global synchronization. Driven oscillator models have also been studied with disorder in the couplings. The Kuramoto model with disordered interactions leads to partial synchronization if the couplings can take negative values [35]. In the biochemistry of DNA methylation turnover, the coupling must be positive. The disorder in the coupling between oscillators induced by the DNA sequence qualitatively changes this behavior in the asymptotic limit: it leads to an arrest of coarsening such that, in this case, global phase locking is not achieved.

D. Validation of theoretical predictions using sequencing data

The precise quantitative validation of the spatiotemporal patterns predicted by the theory is highly challenging due to the lack of resolution in time, space, and biochemical phase of current sequencing technologies. To qualitatively test the existence of phase-locked domains, we use data recently generated using the SIMPLE-seq method that can distinguish between the states C, 5mC, and 5hmC, even with low coverage in single cells [13]. We performed the primary bioinformatics analysis comprising sequence alignment and quality control as described in Appendix F. We focused our analysis on mouse embryonic stem cells (mESCs) in 2i and serum conditions. These cells co-express the Dnmt1 or Dnmt3 and TET genes, and therefore fulfill the biochemical conditions for active turnover of DNA methylation. mESCs in serum conditions have been shown to exhibit oscillatory dynamics in DNA methylation [8] and globally in nascent transcription [36] with a similar oscillation period of 2 h. Figure 5(a) gives an overview of the spatial distribution of the three states measured in mouse embryonic stem cells. As has been previously reported, the 5hmC mark occurs with a lower probability compared to the C and 5mC states.

To test whether DNA methylation turnover leads to phase-locked domains, we aim to detect two signatures of phase-locked domains. First, we expect that CpGs in close proximity should show correlations in their DNA methylation state. Second, we expect that CpGs in close proximity should

also exhibit a higher than expected tendency to be in adjacent states of the biochemical cycle. To begin, we first asked whether CpGs in genomic proximity correlate in the stage of the DNA methylation cycle. To this end, we calculated the probability of finding two neighboring CpGs in the same DNA methylation state as a function of their distance. We compared this probability to the null hypothesis of statistical independence. Specifically, we sampled CpG states from a multinomial distribution using probabilities estimated as the empirical fractions of the different states for each genomic distance. Because of systematic variations in the density of CpGs across the genome, the probability corresponding to the null hypothesis changes with distance. Figure 5(b) shows that the probability that pairs of CpGs have the same DNA methylation mark is consistently higher than expected from statistical independence. This shows that the 5mC, 5hmC, and C states are locally correlated, as has been reported before in the cases of C and 5mC.

We also calculated the probability that neighboring CpGs are in consecutive stages of the biochemical cycle. Specifically, we computed the probability of finding pairs of CpGs in adjacent states of the cycle (5mC and C) and compared to, as in Fig. 5(b), the case of statistical independence. Because neighboring sites tend to have a small but constant phase difference [cf. Fig. 2(b)], we predict that these probabilities should also be enriched. In contrast to this prediction, we find that methylated CpGs are less likely to be in an adjacent state than expected by statistical independence.

The reason for this discrepancy is that active turnover of DNA methylation is restricted to accessible and typically active regions of the genome [1]. As these regions form only a small fraction of the entire genome, the analysis conducted is statistically dominated by "passive" sites not undergoing active turnover. Because this analysis reflects the strong and generic correlations of DNA methylation marks, it serves as a negative control for the analysis below.

To specifically test our predictions for genomic regions that undergo active turnover, we filtered for putatively oscillating sites. We defined these sites by the requirement that the 5hmC mark was detected in at least one of the sequenced cells. For these sites, we also found that nearby sites are more likely to be in the same state of the cycle compared to the null hypothesis. Significantly, unlike the negative control in Fig. 5(d), CpGs in close proximity are simultaneously enriched to be in adjacent stages of the cycle [Fig. 5(e)]. This effect is stronger in serum conditions than in 2i conditions. The relatively weaker effect for cells in 2i conditions is expected from the fact that these cells do not express the Dnmt3 genes, such that active turnover of DNA methylation is limited by the relatively inefficient de novo catalytic activity of DNMT1. Taken together, these empirical findings indicate signatures of phase locking. This phase locking is specific for putatively oscillating sites and, as expected, is not found in the rest of the genome.

Testing our predictions in more detail will require the profiling of several stages of the biochemical cycle with high spatial and temporal resolution in single cells. While such experimental technologies are not available, we propose an experimental approach that overcomes these limitations partially. As a first step, putative genomic regions undergoing active turnover could be identified using SIMPLE-seq experiments, as in Fig. 5. Temporal information could then be inferred by synchronizing cells and genomic loci with respect to their position along the biochemical cycle. In mouse embryonic stem cells this can be achieved by long-term culturing cells in 2i conditions [8]. By releasing cells to serum conditions, Dnmt3 genes are upregulated and active turnover of DNA methylation is facilitated. The final step would be performing a time-course of high-coverage sequencing experiments, such as amplicon-seq, of the genomic regions identified in the first step.

III. DISCUSSION AND CONCLUSION

DNA methylation turnover is an active, epigenetic phenomenon that has been associated with the control of key cell fate decisions. Yet, its genomic distribution, temporal evolution, and functional role remain unknown. This is due in part to limitations in experimental technologies, which put emphasis on approaches guided by theoretical predictions. In this work, we combined analytical calculations, simulations, and analysis of sequencing data to show that chromatin-mediated interactions between local biochemical cycles can give rise to emergent phase-locking dynamics. We found that phase-locked domains coarsen and that genomic disorder arrests the coarsening dynamics, resulting in the formation of stable phase-locked domains. DNA methylation interacts with modifications of histone tails [37,38], and both are coupled to chromatin structure [39]. While similar irreversible cycles have not been shown for histone tail modifications, their dynamics might follow that of DNA methylation and exhibit similar spatiotemporal dynamics.

Our theoretical predictions assume that the progression through the biochemical cycle describing DNA methylation turnover has a strong deterministic component. This would be the case if the number of effective cytosine configurations, Ω, is approximately larger than 10. Significantly smaller values of Ω do not give rise to any kind of order [25]. The biological value of Ω depends on the detailed kinetic rates describing enzyme binding, unbinding, and catalytic reactions. Because these rates are largely unknown, the biological value of Ω cannot be estimated. Experimental evidence [8] shows, however, local synchronization in small genomic domains, suggesting a largely deterministic dynamics of the phase variable.

We qualitatively tested these results using single-cell sequencing data. We detected qualitative signatures of phase-locking behavior that are specific to regions undergoing DNA methylation turnover. Such tests are currently limited by the constraints of sequencing technologies, which cannot capture dynamic information and have a tradeoff between the resolution of the biochemical cycle and the genomic coverage in individual cells. Our work suggests that understanding the role of active DNA methylation turnover may require novel experimental approaches that capture multiple states of DNA methylation turnover in single cells with high coverage, but they may be limited to small genomic regions, such as enhancers or inducible promoters.

Oscillators are a ubiquitous feature of biological systems. They are used to perform a variety of functions, including timing, as in the cell cycle or the circadian rhythm, and signal

processing, as in neuronal oscillators. In somitogenesis, coupled genetic oscillators lead to traveling waves that are used to pattern the early embryo. Our work demonstrates the emergence of surprisingly complex patterns arising from DNA methylation turnover. This raises the question of whether the functional significance of DNA methylation turnover might lie in these emergent patterns as opposed to the turnover of individual CpGs. Driven oscillator models have been extensively studied to understand synchronization phenomena. Our theoretical work highlights that chromatin-mediated nonreciprocal interactions give rise to phase-locking behavior that differs from models that have been studied so far in this field. It shows that active processes in a chromatin context provide an interesting testing ground for nonequilibrium physics.

ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 950349 and the Marie Skłodowska-Curie Grant Agreement No. 101034413. The computations in this paper were run in part on the the FASRC Cannon cluster supported by the FAS Division of Science Research Computing Group at Harvard University and the cluster of the Max Planck Institute for the Physics of Complex Systems.

Author contributions are as follows: Conceptualization, supervision and writing, F.O. and S.R.; numerical simulations, F.O. and M.G.; theoretical analysis, F.O.; experimental data processing, O.B.; experimental data analysis, F.O.

DATA AVAILABILITY

The data that support the findings of this article are openly available [40].

APPENDIX A: DISCRETE PHASE EXPANSION

In this appendix, we give details for the system size expansion of the master equation. From the definition of the rates [Eq. (1)], the master equation is

$$\partial P(\boldsymbol{\phi},t)\partial t = \sum_{i=1}^{N}(\mathrm{E}_i^{-1}-1)[\omega_i + k_i(\boldsymbol{\phi})]P(\boldsymbol{\phi},t). \tag{A1}$$

Here, we made use of the the shifting operator E_i, which raises or lowers the phase variable in a function G on a given site by 1, $\mathrm{E}_i^{\pm 1}G(\boldsymbol{\phi}) = G(\phi_i,\ldots,\phi_i \pm 1,\ldots,\phi_N)$ [24].

After expanding the variables with respect to the system size as $\phi_i = \Omega\Phi_i(t) + \Omega^{1/2}\xi_i(t)$, the left-hand side of Eq. (A1) is

$$\sum_i \frac{dP(\boldsymbol{\phi},t)}{d\phi_i} = \frac{d\Pi}{dt} - \sum_i\left[\Omega^{1/2}\frac{d\phi_i}{dt}\frac{d\Pi}{d\xi_i}\right]. \tag{A2}$$

The operators $\mathrm{E}_i^{\pm}$ in the system size expansion are approximated to the highest order in Ω as

$$\mathrm{E}_i^{\pm} \sim 1 \pm \Omega^{1/2}\frac{\partial}{\partial\xi_i} + \frac{1}{2}\Omega^{-1/2}\frac{\partial^2}{\partial\xi_i^2}. \tag{A3}$$

Upon using the expansion of the operators and collecting the highest order terms in Ω, we get

$$\begin{aligned}\frac{d\Pi}{dt} - \sum_i\left[\Omega^{1/2}\frac{d\Phi_i}{dt}\frac{\partial\Pi}{\partial\xi_i}\right] &= \sum_i -w_i\left[\Omega^{1/2}\frac{\partial}{\partial\xi_i} - \frac{1}{2}\frac{\partial^2}{\partial\xi_i^2}\right]\Pi \\ &\quad - \Omega^{1/2}\frac{\partial}{\partial\xi_i}\left[k(\Phi_i) + \Omega^{-1/2}\frac{\partial k}{\partial\Phi_i}\xi_i\right]\Pi.\end{aligned} \tag{A4}$$

Collecting terms in powers of $\Omega^{1/2}$, and using the chain rule, we arrive at

$$\frac{d\Phi_i}{dt} = w_i + f_1(\Phi_i)\sum_{k=1}^{N}\frac{J_1 e^{-m|k-i|}}{|k-i|^{\lambda}}f_2(\Phi_k), \tag{A5}$$

which is the mean-field equation and where we simply made $k(\Phi)$ explicit. The next lowest order in Ω expresses the dynamics of fluctuations as a Fokker-Planck equation,

$$\frac{\partial\Pi}{\partial t} = \sum_i\left[[w_i + k(\Phi)]\frac{\partial^2\Pi}{\partial\xi_i^2} - w_L\frac{\partial k(\Phi)}{\partial\Phi_i}\frac{\partial(\xi_i\Pi_i)}{\partial\xi_i}\right]. \tag{A6}$$

For small values of $k(\Phi)$, the noise is dominated by ω_i and it is a Gaussian white noise, Eq. (2).

APPENDIX B: DERIVATION OF THE CONTINUITY EQUATION

To quantify the degree of system-wide synchronization, we study the behavior of the Kuramoto order parameter $r(t)$ [26],

$$r(t)e^{i\psi(t)} = \frac{1}{N}\sum_{i=1}^{N}e^{i\phi_i(t)}. \tag{B1}$$

To find an analytical expression for $r(t)$, we first define the generator of moments [28,41,42],

$$H_{k,q}^{c} = \frac{1}{N}\sum_{j=1}^{N}\overline{\langle e^{ik\phi_j}e^{iqj}\rangle w_j^c}, \tag{B2}$$

where $\overline{(\cdot)}$ is the average over the distribution $g(\omega)$ of intrinsic frequency ω_i, and $\langle(\cdot)\rangle$ is an average over all possible realizations of noise. In Eq. (B2), we employed a Fourier transform of the fields. Hence, k is dual to ϕ, and q is dual to the position along the DNA sequence.

In Eq. (B2) we introduced the moment generating function for ω_i with respect to both stochastic noise and intrinsic noise given by the possible non-δ distribution of intrinsic frequencies. In Itô convention, the time evolution of an arbitrary function $F(\cdot)$ of a given stochastic process ϕ_i $(i = 1,\ldots,N)$ and with noise amplitude $\sqrt{2D_i}$ is

$$\partial_t F(\boldsymbol{\phi}) = \sum_j\left[\partial_{\phi_j}F(\phi)\partial_t\phi_j + \sum_k\frac{\partial^2 F(\boldsymbol{\phi})}{\partial\phi_j\partial\phi_k}\sqrt{D_jD_k}\right]. \tag{B3}$$

As the previous equation is valid for any function F, we compute the function $\partial_t\langle e^{ik\phi_j}e^{iqj}\rangle$. We substitute in Eq. (B3) the stochastic process where a trajectory is given by Eq. (2)

hence obtaining (after using the standard property of stochastic calculus)

$$\frac{\partial\langle e^{ik\phi_j}e^{iqj}\rangle}{\partial t} = ike^{ik\phi_j}e^{iqj}G(\phi_j) - k^2e^{ik\phi_j}e^{iqj}D_j. \tag{B4}$$

We then multiply the previous equation by ω_j^c, sum over all $j = 1, \ldots, N$, divide by N, and average over the frequency distribution,

$$\frac{1}{N}\sum_j \overline{\frac{\partial\langle e^{ik\phi_j}e^{iqj}\rangle\omega_j^c}{\partial t}} = \frac{1}{N}\sum_j \overline{\omega_j^c[ike^{ik\phi_j}e^{iqj}G(\phi_j) - k^2e^{ik\phi_j}e^{iqj}D_j]}, \tag{B5}$$

with $G(\phi_j) = \omega_j + f_1(\phi_j)\sum_{w=1}^{N}\frac{Je^{-m|w-j|}}{|w-j|^\lambda}f_2(\phi_w)$. The left-hand side of the previous equation is simply $\partial_t H_{k,q}^m$ as defined in Eq. (B2). To simplify the right-hand side, we introduce the Fourier representation of f_1,f_2 as

$$f_1(\phi_j) = \sum_{n=-\infty}^{\infty} a_n e^{in\phi_j},$$
$$f_2(\phi_w) = \sum_{l=-\infty}^{\infty} b_l e^{il\phi_w}. \tag{B6}$$

As $\frac{e^{-m|w-j|}}{|w-j|^\lambda}$ is just a function of the difference, we can define its Fourier transform as

$$\frac{e^{-m|w-j|}}{|w-j|^\lambda} = \sum_{s=-\infty}^{\infty} r_s e^{is(w-j)}. \tag{B7}$$

The first term on the right-hand side of Eq. (B5) is given by

$$\frac{J}{N}\sum_j\sum_{n,l,s}\sum_c a_n b_l r_s ike^{i(k+n)\phi_j}e^{i(q-s)j}e^{il\phi_w}e^{isw}\omega_j^c. \tag{B8}$$

As $\frac{1}{N}\sum_w e^{il\phi_w}e^{isw} = H_{l,s}^0$, the previous equation simplifies to

$$(ik)JN\sum_{n,l,s} a_n b_l r_s H_{k+n,q-s}^c H_{l,s}^0. \tag{B9}$$

The other terms can be computed in a similar way, and the resulting dynamical equation for the moments is

$$\partial_t H_{k,q}^c = (ik)JN\sum_{n,l,s} a_n b_l r_s H_{k+n,q-s}^m H_{l,s}^0 + (ik)H_{k,q}^{c+1} - k^2 H_{k,q}^{c+1}. \tag{B10}$$

To obtain a more compact expression, we define the generating function

$$\chi(\phi, y, z, t) = \sum_{k=-\infty}^{\infty}\sum_{c=\infty}^{\infty}\sum_{q=-\infty}^{\infty} e^{-ik\phi}e^{-iqz}\frac{y^c}{2\pi c!}H_{k,q}^c, \tag{B11}$$

and we show how all the terms of Eq. (B10) can be rewritten in terms of this function. We analyze, at first, the second term on the right-hand side of (B10), which can is simplified to (upon summing over m, k, q and multiplying by $\frac{y^c}{2\pi c!}$)

$$-\frac{\partial}{\partial\phi}\sum_{k,c,q} e^{-ik\phi}e^{-iqz}\frac{\partial}{\partial y}\frac{y^{c+1}}{2\pi c!}H_{k,q}^{c+1}. \tag{B12}$$

In terms of the function χ, the previous equation is

$$-\frac{\partial^2}{\partial\phi\partial y}\chi(\phi, y, z, t). \tag{B13}$$

Following the same procedure, the third term on the right-hand side of Eq. (B10) is rewritten as

$$\frac{1}{2}\frac{\partial}{\partial y}\frac{\partial}{\partial\phi}D(y)\frac{\partial}{\partial\phi}\chi(\phi, y, z, t). \tag{B14}$$

We proceed now to write the terms describing interactions in terms of the moments. Initially, we write them as

$$-JN\frac{\partial}{\partial\phi}\left[\sum_{n,l,s} a_n b_l r_s e^{inx}e^{isz}\chi(\phi, y, z, t)H_{l,s}^0\right], \tag{B15}$$

where we multiply and divide by $e^{inx}e^{isz}$ so that $H_{k,q} \to H_{k+n,q-s}$. We further define a term

$$\nu(\phi, y = 0, z, t) = JN\left[\sum_{n,l,s} a_n b_l r_s e^{inx}e^{isz}H_{l,s}^0\right], \tag{B16}$$

and writing Eq. (B6) inversely and going back to the space of function defined on ϕ, we obtain

$$\nu(\phi, y = 0, z, t) = JN\left[f_1(\phi)\sum_{l,s} b_l r_s e^{isz}H_{l,s}^0\right]. \tag{B17}$$

We show all the other terms in the same way as

$$\nu = J\left[f_1(\phi)\int d\hat{\phi}\int d\hat{z}\frac{e^{-m\hat{z}}}{|\hat{z}|^\lambda}f_2(\hat{\phi})\chi(\hat{\phi}, y = 0, z - \hat{z}, t)\right], \tag{B18}$$

with $z \in (0, 1)$. Altogether the time evolution of the moment generator χ is given by

$$\partial_t\chi(\phi, y, z, t) = -\frac{\partial}{\partial\phi}[\nu(\phi, y, z, t)\chi] + \frac{\partial}{\partial y}\frac{\partial}{\partial\phi} \times D(y)\frac{\partial}{\partial\phi}\chi(\phi, y, z, t) - \frac{\partial\chi}{\partial\phi\partial y}, \tag{B19}$$

where the drift term takes the form

$$\nu = y + J\left[f_1(\phi)\int d\hat{\phi}d\hat{z}\frac{e^{-m\hat{z}}}{|\hat{z}|^\lambda}f_2(\hat{\phi})\chi(\hat{x}, 0, z - \hat{z}, t)\right], \tag{B20}$$

where the diffusion term is $D(y) = 2y$. We define $\rho(\phi, \omega, z, t)$ as the density of oscillators with phase ϕ, position z, and frequency ω at time t. The continuity equation is given by exploiting the relationship $\chi(\phi, y, z, t) = \int d\omega g(\omega)e^{y\omega}\rho(\phi, \omega, z, t)$ and setting the diffusion to zero for the deterministic limit.

APPENDIX C: PHASE-LOCKED SOLUTION

Equation (4) does not admit asynchronous stationary solutions. The control functions f_1, f_2 break the rotational symmetry such that the equations are not invariant under a shift of the fields (oscillators). Indeed, the first integral in Eq. (4) will never vanish for an asynchronous solution as

long as $\int d\phi f_2(\phi) \neq 0$. We also have to consider consistency with the biological reality: the distribution of frequencies for each oscillator, $g(\omega)$, has positive support as, on average, the biochemical cycle advances in the positive direction. Typically one can shift the distribution by moving to a rotating reference frame with an angular velocity given by the median of $g(\omega)$ [28]. However, in our model, due to the lack of rotational symmetry, this transformation does not simplify the analytical calculations. For general functions $f_1(\phi_i)f_2(\phi_j)$, it is sometimes possible to change to a coordinate system that allows the asynchronous solution to be stable [43], as long as the product $f_1(\phi_i)f_2(\phi_j)$ can be written as a 2π-periodic function of the phase difference $\phi_i - \phi_j$, with g being a function that is 2π-periodic and well defined. As the $f_{1,2}(\phi)$ are step functions, it is not always possible to write the product in the rotationally symmetric form. Moreover, since the asynchronous solution is never stable, a stability analysis as in [44,45] or a power series expansion [46] around this state does not give any useful insight for this process; the stationary states are nontrivial even for a vanishing strength of interactions J. We thus look for an alternative order parameter, which still remains experimentally accessible. A natural order parameter is the fraction of sites for which $f_2(\phi) = 1$, namely m. Biologically, this is the average DNA methylation. For the biological control functions, $m = 1/2$ in a phase where the oscillators are uncoupled, and it is greater than $1/2$ whenever $J > J_c$. We are now interested in finding possible behaviors of the collective dynamics of oscillators. To this end, we search for a stationary solution of Eq. (4)

APPENDIX D: STATIONARY SOLUTION

A spatially homogeneous stationary solution of Eq. (4) is ($\omega \geqslant 0$ in the model definition)

$$\rho = \frac{C(\omega)}{|\omega + H(\pi/2 - \phi)A(\lambda)J|}, \tag{D1}$$

where

$$A(\lambda) = \left(\pi \int d\omega \frac{g(\omega)C(\omega)}{\omega}\right)^{\lambda} \int_0^{\pi \int d\omega \frac{g(\omega)C(\omega)}{\omega}} dy \frac{e^{-|y|}}{|y|^{\lambda}}, \tag{D2}$$

and $\Gamma(1-\lambda, 0, m) = \int_0^m dy\, e^{-|y|}/|y|^{\lambda}$ is the generalized incomplete gamma function. $H(x)$ denotes the Heaviside step function. $C(\omega)$ is set by the normalization of the probability density function ρ [28] ($\int_0^{2\pi} d\phi\rho = 1$),

$$C(\omega) = \frac{2|\omega||A(\lambda)J + \omega|}{\pi[|\omega| + 3|A(\lambda)J|]}. \tag{D3}$$

To understand how the degree of synchronization changes as a function of the coupling strength, J, we compute m using the stationary distribution of ρ [Eq. (D1)],

$$m = \int_0^{\infty} d\omega \frac{2(A(\lambda)J + \omega)}{3A(\lambda)J + 4\omega} g(\omega), \tag{D4}$$

where $A(\lambda) = A(\lambda) = m^{\lambda}\Gamma(1-\lambda, 0, m)$.

For an exponential distribution of intrinsic frequencies, $g(\omega) = \mu^{-1}\exp(\mu\omega)$, the average DNA methylation satisfies the self-consistent equation,

$$m = \frac{1}{2} - \frac{\mu JA(\lambda)}{8} e^{\frac{3\mu JA(\lambda)}{4}} \mathrm{Ei}\left(-\frac{3\mu JA(\lambda)}{4}\right), \tag{D5}$$

where $\mathrm{Ei}(x) = -\int_{-x}^{\infty} dy\, e^{-y}/y$ is the exponential integral function.

APPENDIX E: NUMERICAL SIMULATION METHODS

All mean-field simulations are run with a fourth-order Runge-Kutta method with a discrete time-step of $dt = 10^{-3}$ and $L = 10^3$ lattice sites, of which the first 200 are shown. The exponent of the long-range kernel [Eq. (1)] is $\lambda = 1/3$. Domain boundaries in Fig. 4(a) are defined whenever two neighboring sites have a phase difference greater than $dt \max_i(\omega_i + k_i)$.

For the stochastic simulations of Eq. (A1), we used the Gillespie algorithm [47]. Each point on the lattice, set equidistant from each other, mimics a single CpG site, with the genome size being 1000 sites under periodic boundary conditions. Each site $i \in [1, 2, \ldots, 1000]$ is defined by its position along the genome, i, its clock state, CS_i, and an intrinsic rate parameter, ω_i, drawn from an exponential distribution with the unitary mean. We tested a simulation with the number of clock states, N_{states}, ranging between 10 and 100 with qualitatively consistent results. We also tested the clock state at each site, which was was initialized randomly from a uniform distribution.

We defined the region ψ_2 to be between $\frac{\pi}{2} \to 25$ and $\frac{3\pi}{2} \to 75$. The value of m was defined as the fraction of sites in phase interval ψ_2. We also calculated the Kuramoto order parameter for each time step defined as $r = [\sum_i \cos(2\pi CS_i)]^2 + [\sum_i \sin(2\pi CS_i)^2]^{1/2}$, where $CS_i \in \{1, 2, \ldots, N_{\text{states}}\}$./ Each simulation was run for a maximum time of $T = L^2/N_{\text{states}}$ and we measured the average methylation once every 100 time steps.

APPENDIX F: DATA PROCESSING AND ANALYSIS

Raw sequencing used in this study is openly openly available from the Gene Expression Omnibus (GEO) [40] and sample number GSM5929312 (embryonic stem cells, mouse). The bioinformatics analysis is carried out as described in Ref. [13] with a slight modification. Briefly, we aligned the sequencing data to the mouse (GRCh38) genome reference and used default parameters for both alignment and quality control using a custom pipeline. Upon extracting methylation and hydroxymethylation states of cytosines in CpG contexts, we did not perform any binning. The pipeline used for the bioinformatics analysis is available in Ref. [48]. In Figs. 5(b) and 5(c) we use the data in bulk, with the only restriction being that we needed sites that were tested for both 5mC and 5hmC. In Figs. 5(b) and 5(c) we subset the data to sites where at least one site in all cells has positive reads for 5hmC and negative reads for 5mC.

[1] A. Parry, S. Rulands, and W. Reik, Active turnover of DNA methylation during cell fate decisions, Nat. Rev. Genet. **22**, 59 (2021).

[2] J. Agostinho de Sousa, C.-W. Wong, I. Dunkel, T. Owens, P. Voigt, A. Hodgson, D. Baker, E. G. Schulz, W. Reik, A. Smith, *et al.*, Epigenetic dynamics during capacitation of naïve human pluripotent stem cells, Sci. Adv. **9**, eadg1936 (2023).

[3] I. Hernando-Herraez, B. Evano, T. Stubbs, P.-H. Commere, M. Jan Bonder, S. Clark, S. Andrews, S. Tajbakhsh, and W. Reik, Ageing affects DNA methylation drift and transcriptional cell-to-cell variability in mouse muscle stem cells, Nat. Commun. **10**, 4361 (2019).

[4] G. Egger, G. Liang, A. Aparicio, and P. A. Jones, Epigenetics in human disease and prospects for epigenetic therapy, Nature (London) **429**, 457 (2004).

[5] A. J. Parry, C. Krueger, T. Lohoff, S. Wingett, S. Schoenfelder, and W. Reik, Dynamic DNA methylation turnover at the exit of pluripotency epigenetically primes gene regulatory elements for hematopoietic lineage specification, bioRxiv (2023).

[6] Y.-F. He, B.-Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, *et al.*, TET-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA, Science **333**, 1303 (2011).

[7] K. E. Szulwach, X. Li, Y. Li, C.-X. Song, H. Wu, Q. Dai, H. Irier, A. K. Upadhyay, M. Gearing, A. I. Levey, *et al.*, 5-hmC–mediated epigenetic dynamics during postnatal neurodevelopment and aging, Nat. Neurosci. **14**, 1607 (2011).

[8] S. Rulands, H. J. Lee, S. J. Clark, C. Angermueller, S. A. Smallwood, F. Krueger, H. Mohammed, W. Dean, J. Nichols, P. Rugg-Gunn, *et al.*, Genome-scale oscillations in DNA methylation during exit from pluripotency, Cell Syst. **7**, 63 (2018).

[9] L. Cimmino, O. Abdel-Wahab, R. L. Levine, and I. Aifantis, TET family proteins and their role in stem cell differentiation and transformation, Cell Stem Cell **9**, 193 (2011).

[10] S. Kangaspeska, B. Stride, R. Métivier, M. Polycarpou-Schwarz, D. Ibberson, R. P. Carmouche, V. Benes, F. Gannon, and G. Reid, Transient cyclical methylation of promoter DNA, Nature (London) **452**, 112 (2008).

[11] R. Métivier, R. Gallais, C. Tiffoche, C. Le Péron, R. Z. Jurkowska, R. P. Carmouche, D. Ibberson, P. Barath, F. Demay, G. Reid, *et al.*, Cyclical DNA methylation of a transcriptionally active promoter, Nature (London) **452**, 45 (2008).

[12] M. J. Booth, M. R. Branco, G. Ficz, D. Oxley, F. Krueger, W. Reik, and S. Balasubramanian, Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution, Science **336**, 934 (2012).

[13] D. Bai, *et al.*, Simultaneous single-cell analysis of 5mC and 5hmC with Simple-seq, Nat. Biotechnol. **43**, 85 (2024).

[14] H. Heyn and M. Esteller, DNA methylation profiling in the clinic: Applications and challenges, Nat. Rev. Genet. **13**, 679 (2012).

[15] F. Olmeda, T. Lohoff, S. J. Clark, L. Benson, F. Krüger, W. Reik, and S. Rulands, Inference of emergent spatio-temporal processes from single-cell sequencing reveals feedback between de novo DNA methylation and chromatin condensates, bioRxiv (2021).

[16] C. Lövkvist, I. B. Dodd, K. Sneppen, and J. O. Haerter, DNA methylation in human epigenomes depends on local topology of CpG sites, Nucl. Acids Res. **44**, 5123 (2016).

[17] S. Li, Y. Peng, and A. R. Panchenko, DNA methylation: Precise modulation of chromatin structure and dynamics, Curr. Opin. Struct. Biol. **75**, 102430 (2022).

[18] J. Otterstrom, A. Castells-Garcia, C. Vicario, P. A. Gomez-Garcia, M. P. Cosma, and M. Lakadamyali, Super-resolution microscopy reveals how histone tail acetylation affects DNA compaction within nucleosomes in vivo, Nucl. Acids Res. **47**, 8470 (2019).

[19] F. Olmeda and S. Rulands, Field theory of enzyme-substrate systems with restricted long-range interactions, Phys. Rev. E **110**, 024404 (2024).

[20] H. Hinrichsen, Non-equilibrium phase transitions, Physica A **369**, 1 (2006).

[21] F. Ginelli, H. Hinrichsen, R. Livi, D. Mukamel, and A. Torcini, Contact processes with long range interactions, J. Stat. Mech. (2006) P08008.

[22] M. A. Al-Radhawi, S. Tripathi, Y. Zhang, E. D. Sontag, and H. Levine, Epigenetic factor competition reshapes the EMT landscape, Proc. Natl. Acad. Sci. USA **119**, e2210844119 (2022).

[23] I. G. Lin, L. Han, A. Taghva, L. E. O'Brien, and C.-L. Hsieh, Murine de novo methyltransferase DNMT3A demonstrates strand asymmetry and site preference in the methylation of DNA in vitro, Mol. Cell. Biol. **22**, 704 (2002).

[24] N. V. Kampen, *Stochastic Processes in Physics and Chemistry* (North Holland, Amsterdam, 2007).

[25] D. J. Jörg, Stochastic Kuramoto oscillators with discrete phase states, Phys. Rev. E **96**, 032201 (2017).

[26] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence*, Dover Books on Chemistry Series (Dover, Mineola, NY, 2003).

[27] S. Gupta, A. Campa, and S. Ruffo, Kuramoto model of synchronization: Equilibrium and nonequilibrium aspects, J. Stat. Mech. (2014) R08001.

[28] J. A. Acebrón, L. L. Bonilla, C. J. Pérez Vicente, F. Ritort, and R. Spigler, The Kuramoto model: A simple paradigm for synchronization phenomena, Rev. Mod. Phys. **77**, 137 (2005).

[29] M. Dukatz, M. Dittrich, E. Stahl, S. Adam, A. De Mendoza, P. Bashtrykov, and A. Jeltsch, DNA methyltransferase DNMT3A forms interaction networks with the CpG site and flanking sequence elements for efficient methylation, J. Biol. Chem. **298**, 102462 (2022).

[30] P. A. Ginno, D. Gaidatzis, A. Feldmann, L. Hoerner, D. Imanci, L. Burger, F. Zilbermann, A. H. Peters, F. Edenhofer, S. A. Smallwood, *et al.*, A genome-scale map of DNA methylation turnover identifies site-specific dependencies of DNMT and TET activity, Nat. Commun. **11**, 2680 (2020).

[31] S. Adam, J. Bräcker, V. Klingel, B. Osteresch, N. E. Radde, J. Brockmeyer, P. Bashtrykov, and A. Jeltsch, Flanking sequences influence the activity of TET1 and TET2 methylcytosine dioxygenases and affect genomic 5hmC patterns, Commun. Biol. **5**, 92 (2022).

[32] T. E. Lee, G. Refael, M. C. Cross, O. Kogan, and J. L. Rogers, Universality in the one-dimensional chain of phase-coupled oscillators, Phys. Rev. E **80**, 046210 (2009).

[33] O. Kogan, J. L. Rogers, M. C. Cross, and G. Refael, Renormalization group approach to oscillator synchronization, Phys. Rev. E **80**, 036206 (2009).

[34] A. M. Deaton and A. Bird, CpG islands and the regulation of transcription, Genes Dev. **25**, 1010 (2011).

[35] L. Bonilla, C. Perez Vicente, and J. Rubi, Glassy synchronization in a population of coupled oscillators, J. Stat. Phys. **70**, 921 (1993).

[36] S. Shah, Y. Takei, W. Zhou, E. Lubeck, J. Yun, C.-H. L. Eng, N. Koulena, C. Cronin, C. Karp, E. J. Liaw, M. Amin, and L. Cai, Dynamics and spatial genomics of the nascent transcriptome by intron Seqfish, Cell **174**, 363 (2018).

[37] H. Zhang, X.-J. Tian, A. Mukhopadhyay, K. S. Kim, and J. Xing, Statistical mechanics model for the dynamics of collective epigenetic histone modification, Phys. Rev. Lett. **112**, 068101 (2014).

[38] K. S. Jani, S. U. Jain, E. J. Ge, K. L. Diehl, S. M. Lundgren, M. M. Müller, P. W. Lewis, and T. W. Muir, Histone H3 tail binds a unique sensing pocket in EZH2 to activate the PRC2 methyltransferase, Proc. Natl. Acad. Sci. USA **116**, 8295 (2019).

[39] T. Zhang, S. Cooper, and N. Brockdorff, The interplay of histone modifications–Writers that read, EMBO Rep. **16**, 1467 (2015).

[40] Gene Expression Omnibus (GEO) accession GSE197740, accession number, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE197740 GSE197740.

[41] C. J. Pérez and F. Ritort, A moment-based approach to the dynamical solution of the Kuramoto model, J. Phys. A **30**, 8095 (1997).

[42] L. L. Bonilla, C. J. Pérez Vicente, F. Ritort, and J. Soler, Exactly solvable phase oscillator models with synchronization dynamics, Phys. Rev. Lett. **81**, 3643 (1998).

[43] H. Daido, Order function and macroscopic mutual entrainment in uniformly coupled limit-cycle oscillators, Prog. Theor. Phys. **88**, 1213 (1992).

[44] S. H. Strogatz, R. E. Mirollo, and P. C. Matthews, Coupled nonlinear oscillators below the synchronization threshold: Relaxation by generalized Landau damping, Phys. Rev. Lett. **68**, 2730 (1992).

[45] S. H. Strogatz, From Kuramoto to Crawford: Exploring the onset of synchronization in populations of coupled oscillators, Physica D **143**, 1 (2000).

[46] E. Ott and T. M. Antonsen, Low dimensional behavior of large systems of globally coupled oscillators, Chaos **18**, 037113 (2008).

[47] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, J. Phys. Chem. **81**, 2340 (1977).

[48] O. Bektas, https://gitlab.com/onurcanbektas/simpleseq (2024).