Forces driving chromatin in the cell nucleus

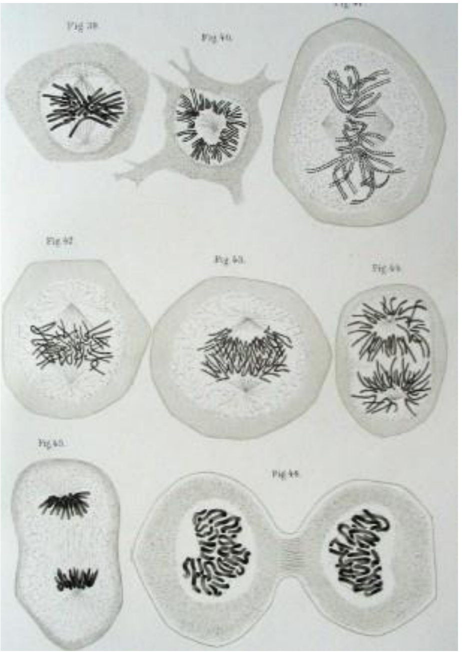
Vittore Scolari, CNRS, Institut Curie, Paris

CTMB3 meeting 2024, Sao Paulo, Brazil

The chromosomes as macro-molecules, a long story

Walther Flemming, Mitosis, 1882

Edward van Beneden, Meiosis, 1883



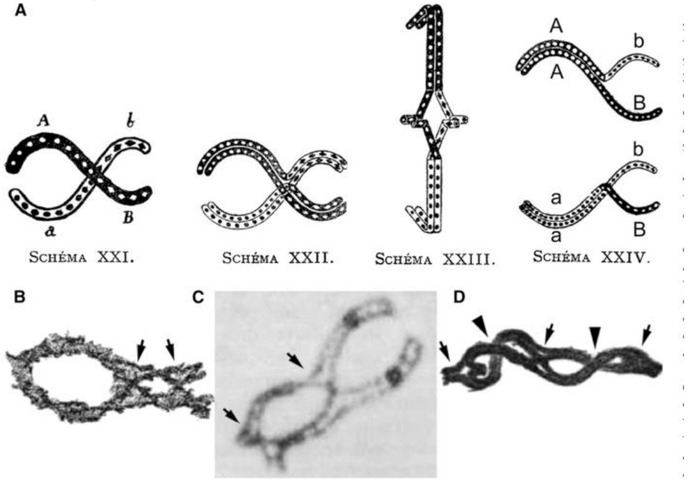
W. Flemming 1882

Boveri and Sutton formulated the chromosome theory of inheritance in terms of chromosome splitting at cell division (1902-1904):

- They observed that chromosomes exists in pairs

- They observed members of a chromosome pairs separate each other during gamete formation

A link with Mendel laws was made by Janssens and Morgan (1909-1911):



From Janssens 1909: exchange of chromosomal segments in chiasmata

50 yeast later, a couple of theoretical physicists discovered the double-helix

Erwin Schrodinger prediction: What is life, 1944.
Chromosomes: “the aperiodic crystal forming the hereditary substance, largely withdrawn from the disorder of heat motion.”

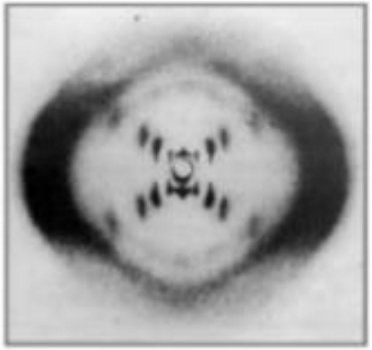
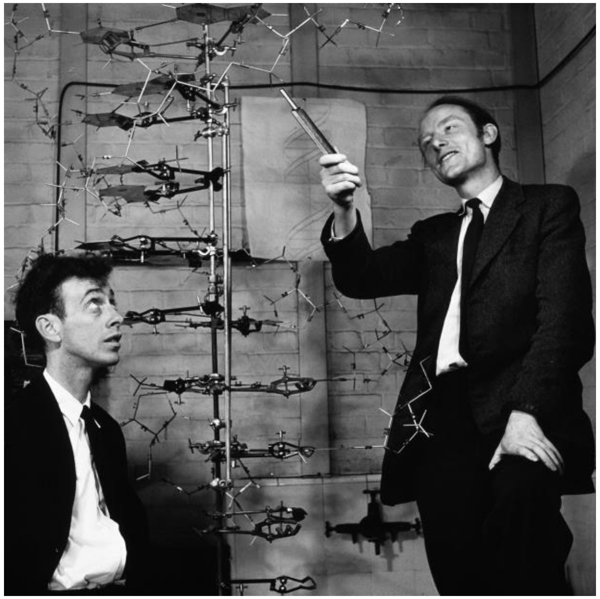


Fig (C):- X-ray diffraction photograph of the DNA double helix taken by Maurice Wilkins & Roslind Franklin published in 1953

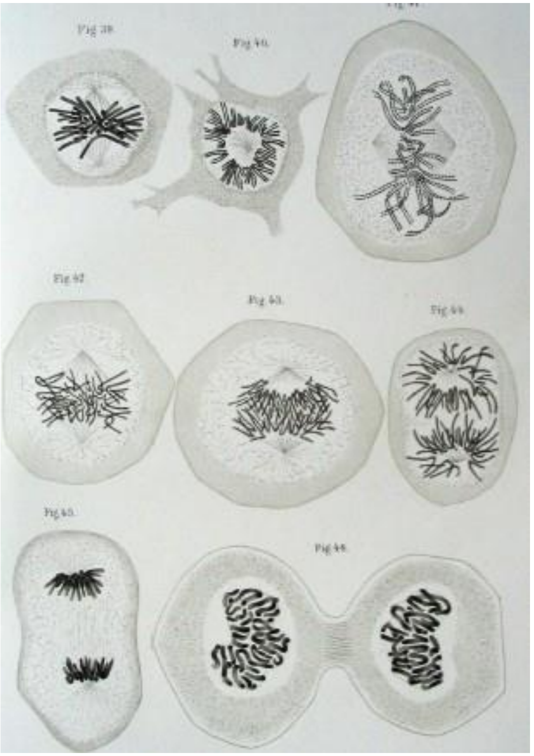


Watson and Crick, The structure of DNA, Nature, 1953

Over ~1 lifespan, the heat motion (in Schrodinger terms) corresponds to errors and variability

The chromosome conformation during Mitosis and Meiosis should be tightly regulated

Not only during mitosis and meiosis



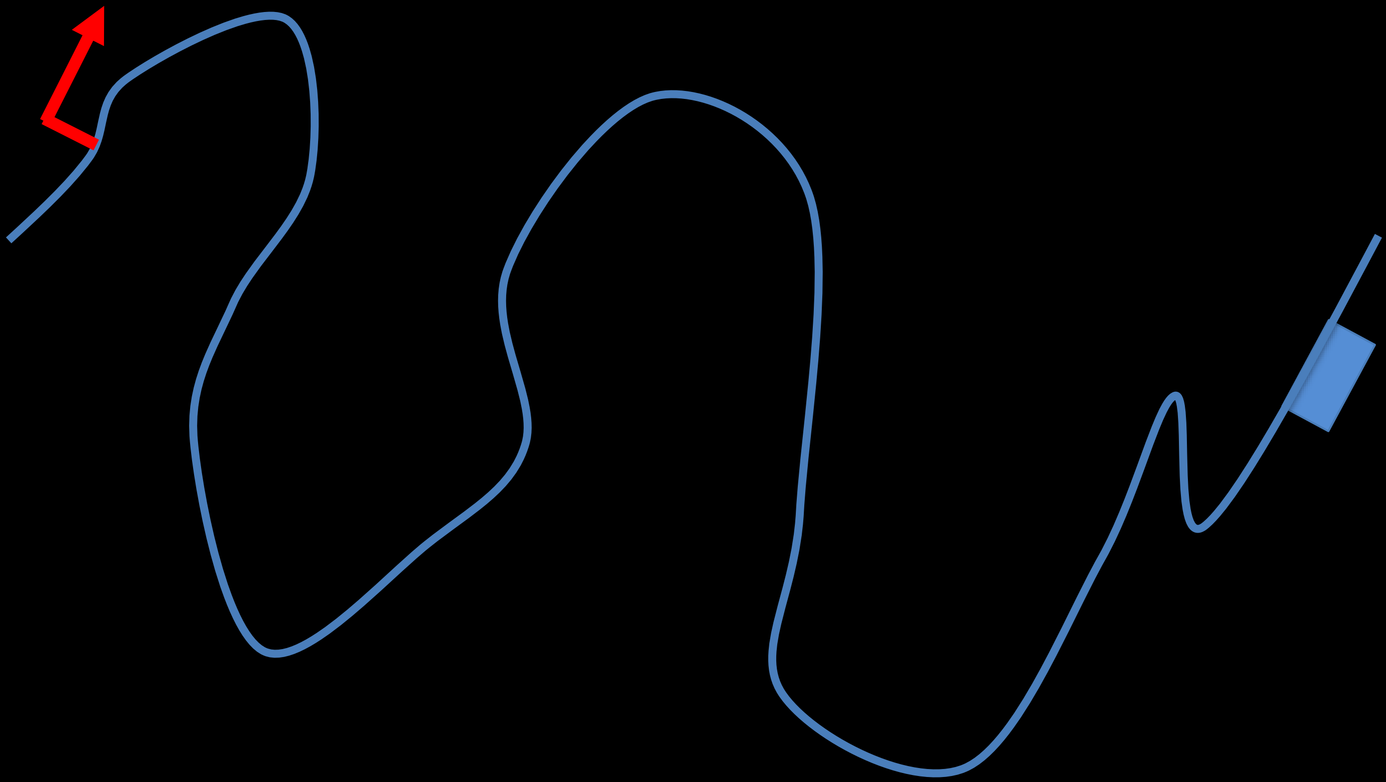
W. Flemming 1882

NOT TOO MUCH!!

Jacques Monod: « l’homme sait enfin qu’il est seul dans l’immensité indifférente de l’Univers d’ou il a émergé par hasard »

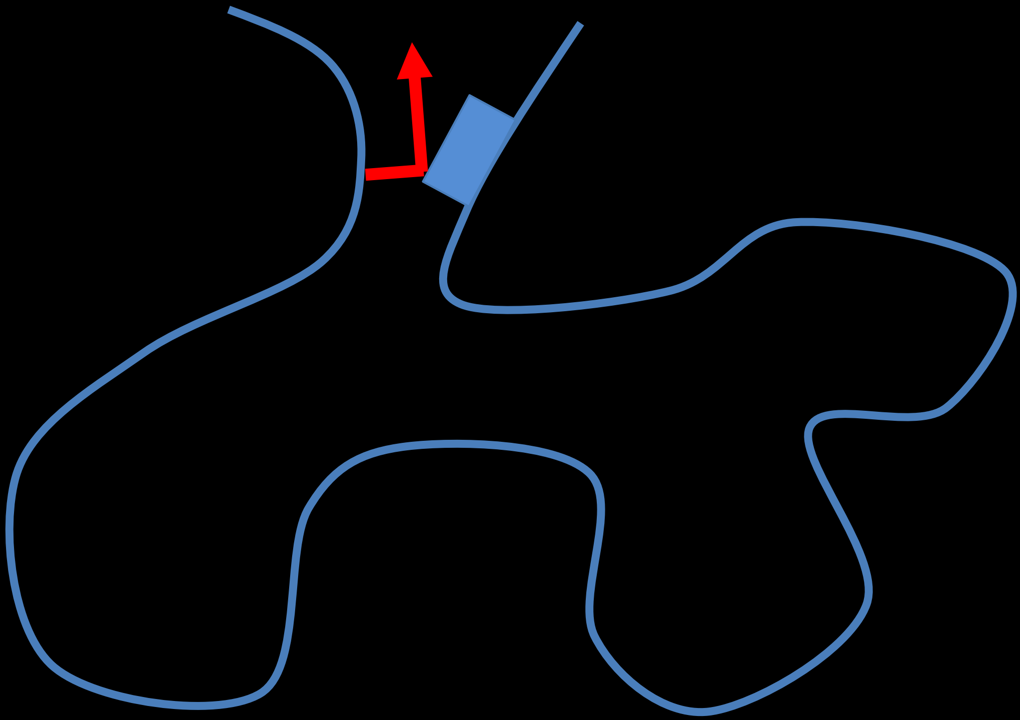
Scientific question:

How do cells bring distant genomic regions close in 3D space?



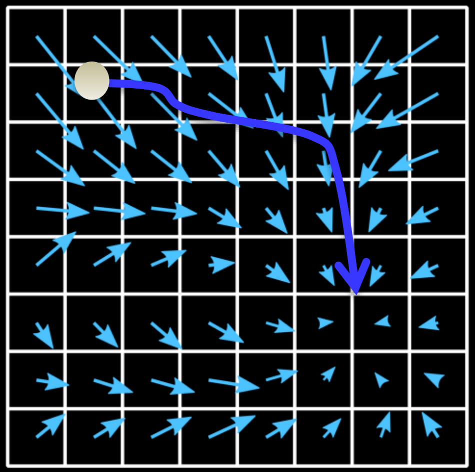
Scientific question:

How do cells bring distant genomic regions close in 3D space?



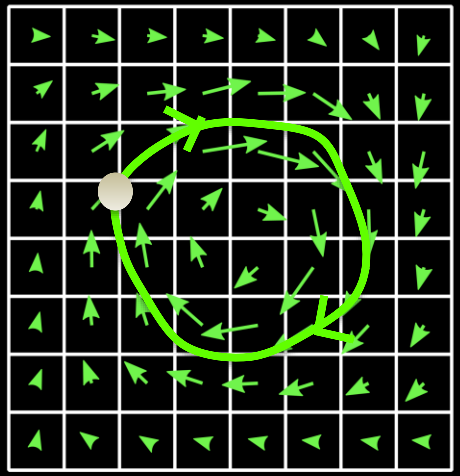
Which forces shape the 3D genome?

Relaxation toward equilibrium

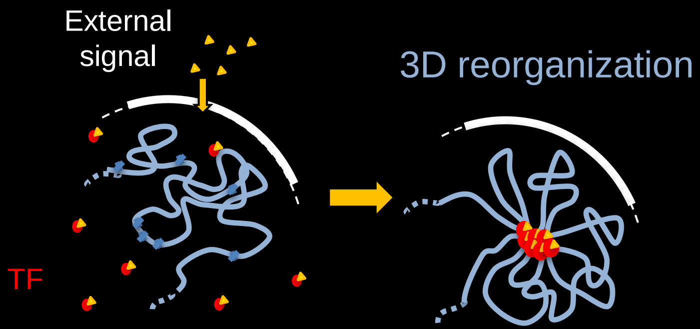


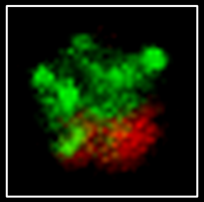
Configurational space

Out-of-equilibrium steady state



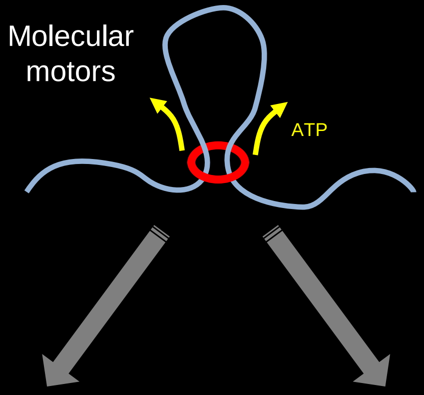
Configurational space





[Myriam Ruault, TaddeiLab]

Silencing epigenetic domains

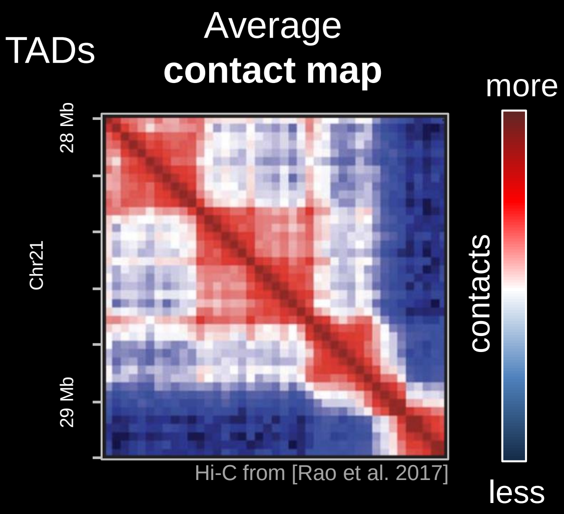


Evolution

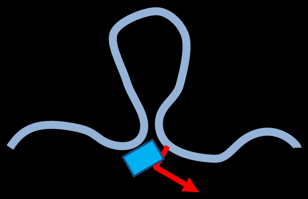
Transcription

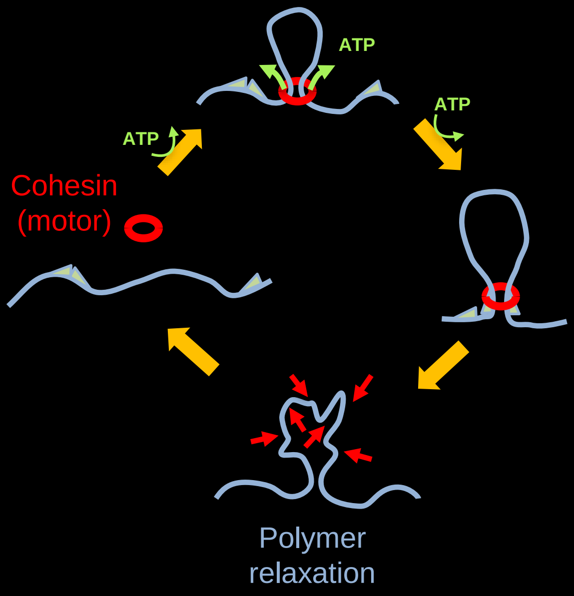
Ruault*, **Scolari*** et al, (2021) ***Genome res.***

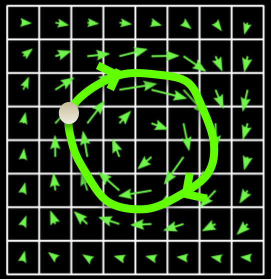
The role of conformational memory on loop interactions



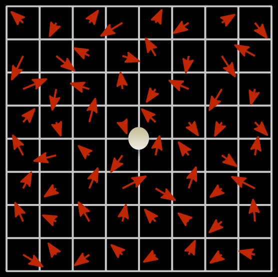
Hi-C from [Rao et al. 2017]







vs



How do
thermal fluctuations
mitigate energy
influxes?

How can
conformational
memory play a
functional role?

Explorative data-driven approach

↓

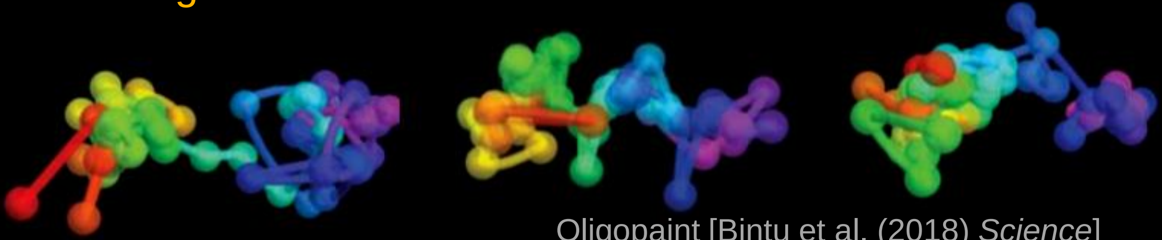
My original way of modelling loop extrusion

↓

Physical theory of chromatin conformation

Explorative “data-driven” modelling approach – static dataset

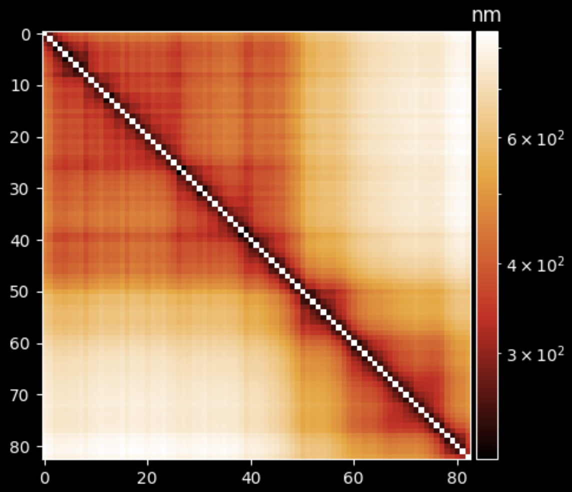
~ 10000 single-cell conformations available online



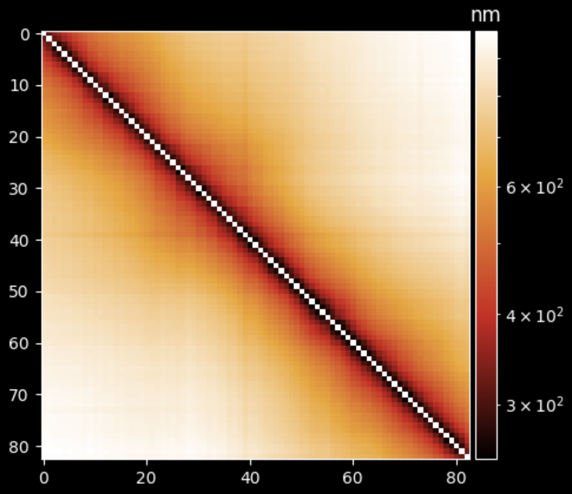
Oligopaint [Bintu et al. (2018) *Science*]

HCT116, chr21:34Mb-37Mb, 83 probes (30kbp mean probe size)

Wild-type



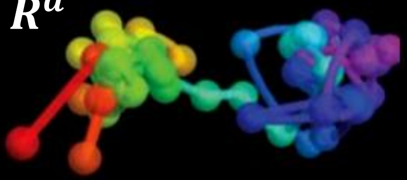
Cohesin degron induced



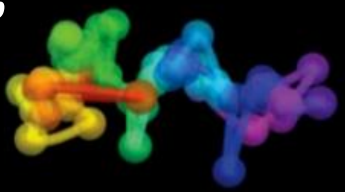
Gaussian model

Oligopaint [Bintu et al. (2018) *Science*]

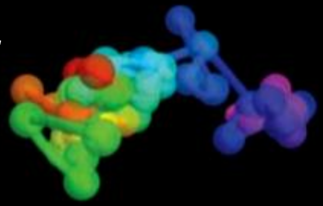
$\boldsymbol{R}^a$



$\boldsymbol{R}^b$



$\boldsymbol{R}^c$



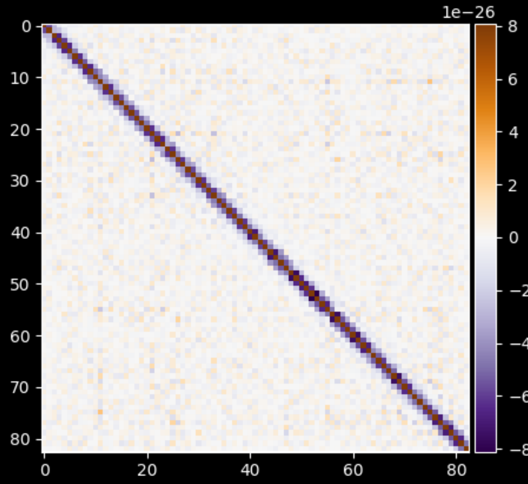
3x83 vectors x structure

$$\Xi_{ij} = \langle R_i R_j \rangle - \langle R_i \rangle \langle R_j \rangle$$

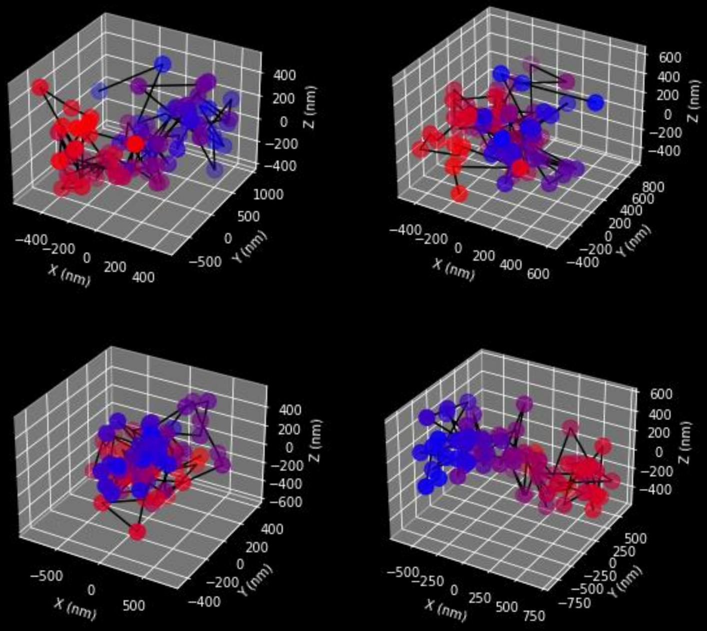
covariance matrix (83 by 83)
means are taken over all configurations

$$P(R) = \frac{1}{Z} \exp\left(-\frac{1}{2} \boldsymbol{R}^T \, \Xi^{-1} \boldsymbol{R} \right)$$

Ξ^{-1}



It is mostly close to zero

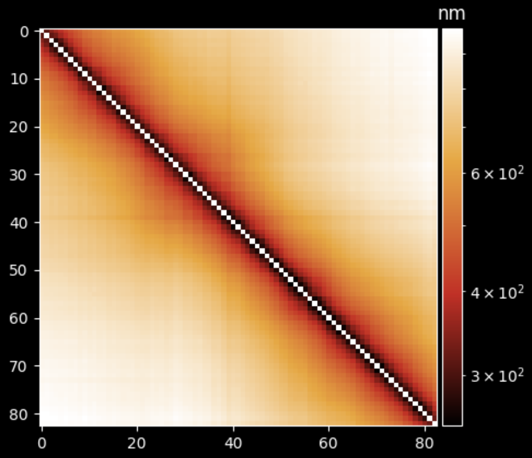


Generated data

Is the model missing anything?

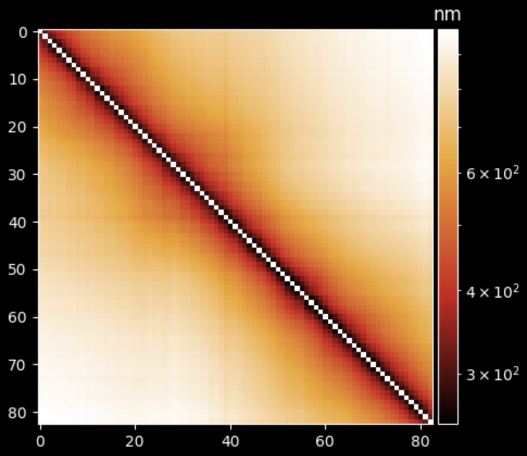
Original data

Distance matrix



Bintu et al. (2018)

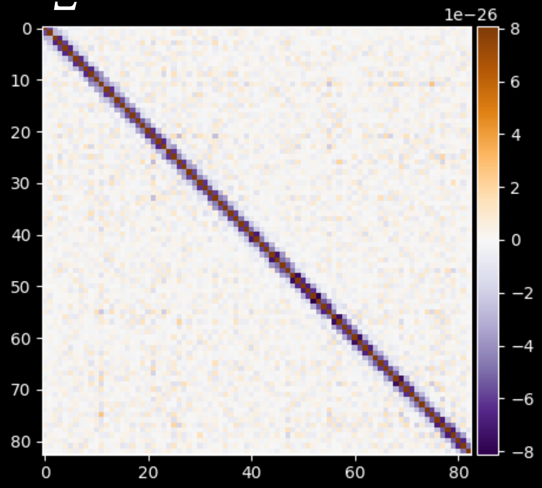
$$P(R) = \frac{1}{Z}\exp\left(-\frac{1}{2}\boldsymbol{R}^T\,\Xi^{-1}\boldsymbol{R}\right)$$



6889 parameters to “fit” 6889 mean observables

Principal component analysis

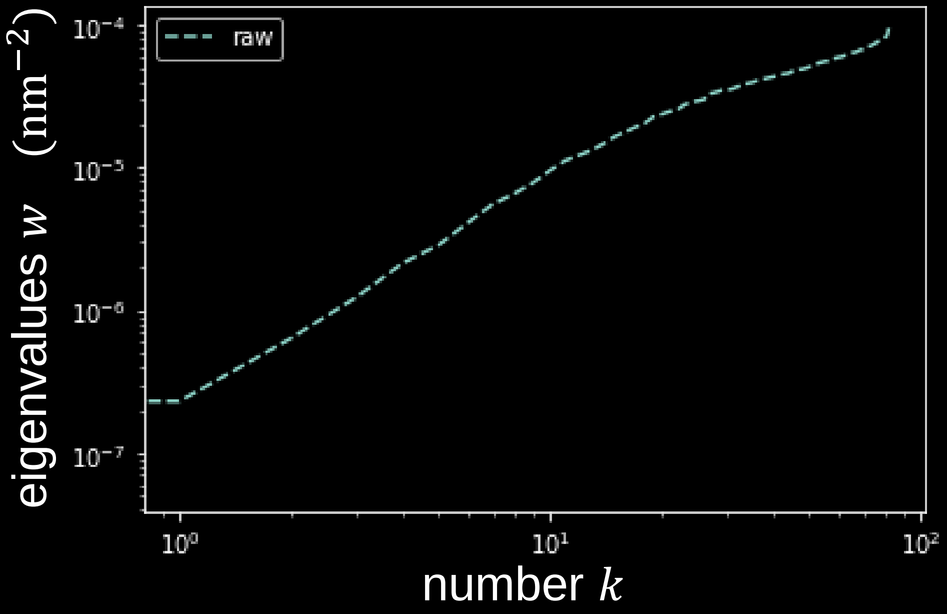
Ξ^{-1}

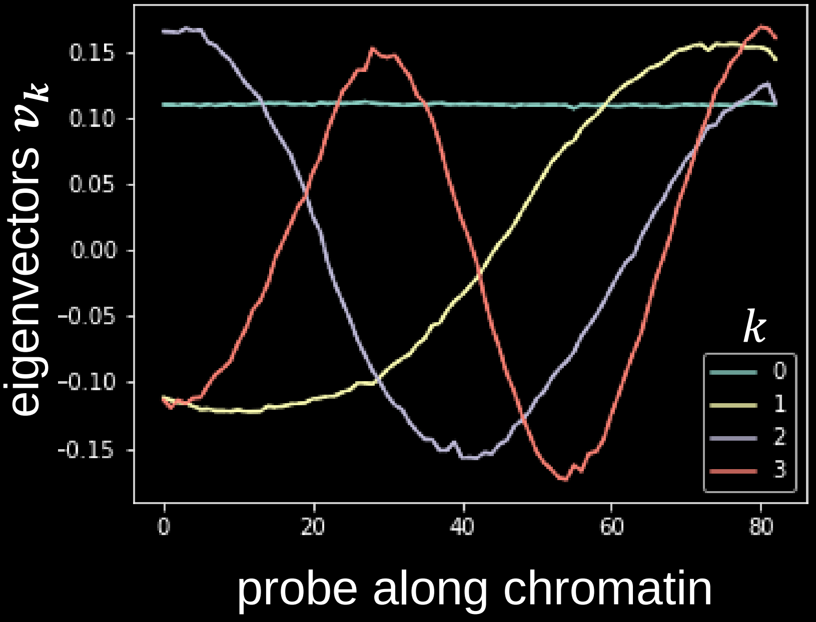


$$\Xi^{-1} = \sum_i w_i \, \boldsymbol{v}_i^T \boldsymbol{v}_i$$

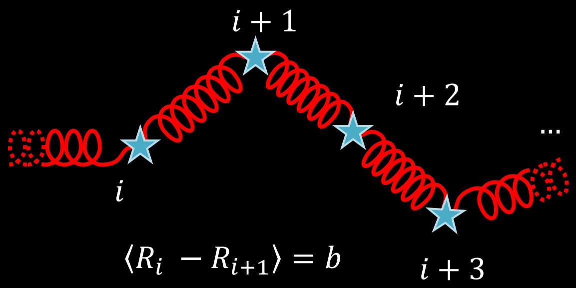
$\boldsymbol{v}_k$ eigenvectors of Ξ^{-1}

w_k eigenvalues (principal components^{-1})





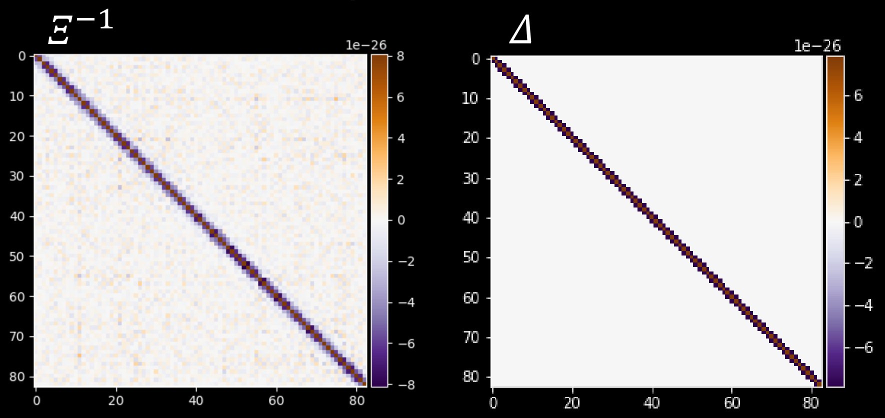
First static model: the random polymer chain

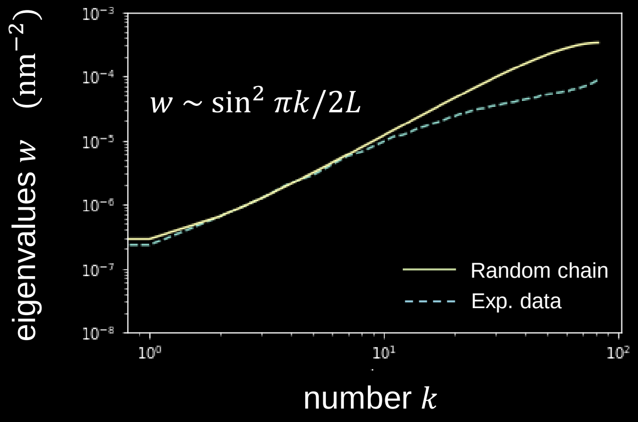


1 parameter

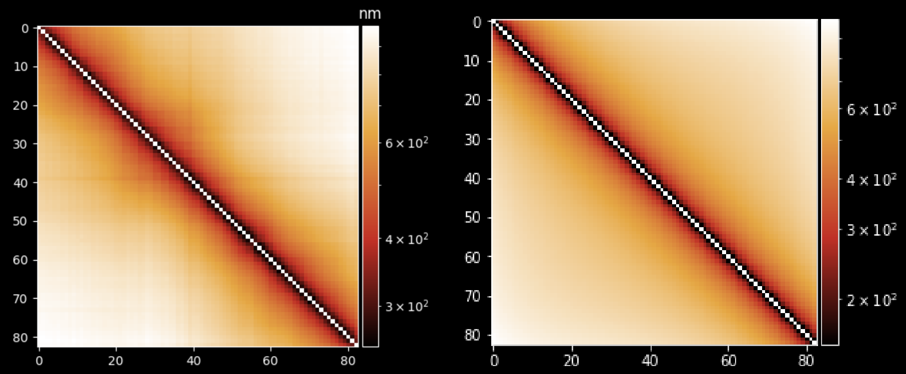
$$P(R) = \frac{1}{Z} \exp\left(-\frac{3}{b^2} \boldsymbol{R}^T \Delta\, \boldsymbol{R}\right)$$

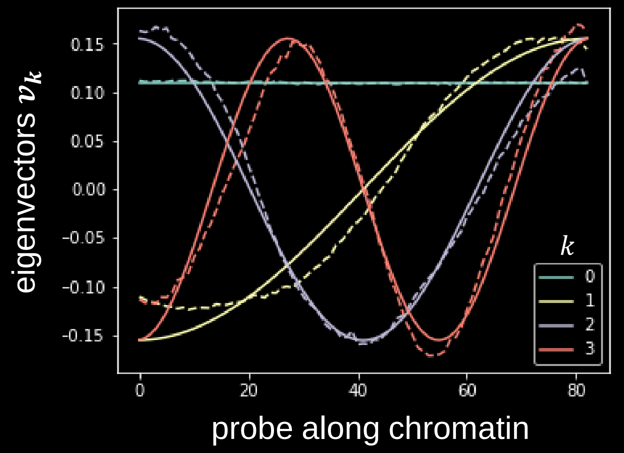
Operator





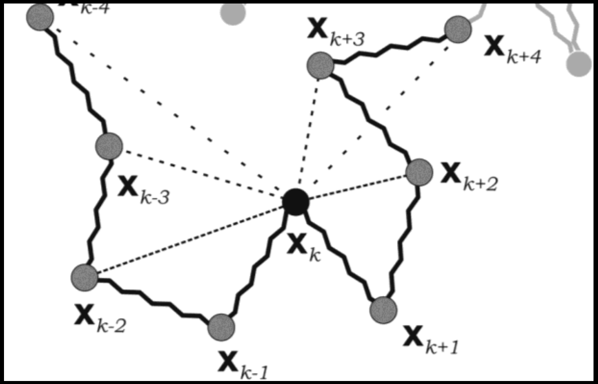
Mean distance matrices





Second static model: the β-chain

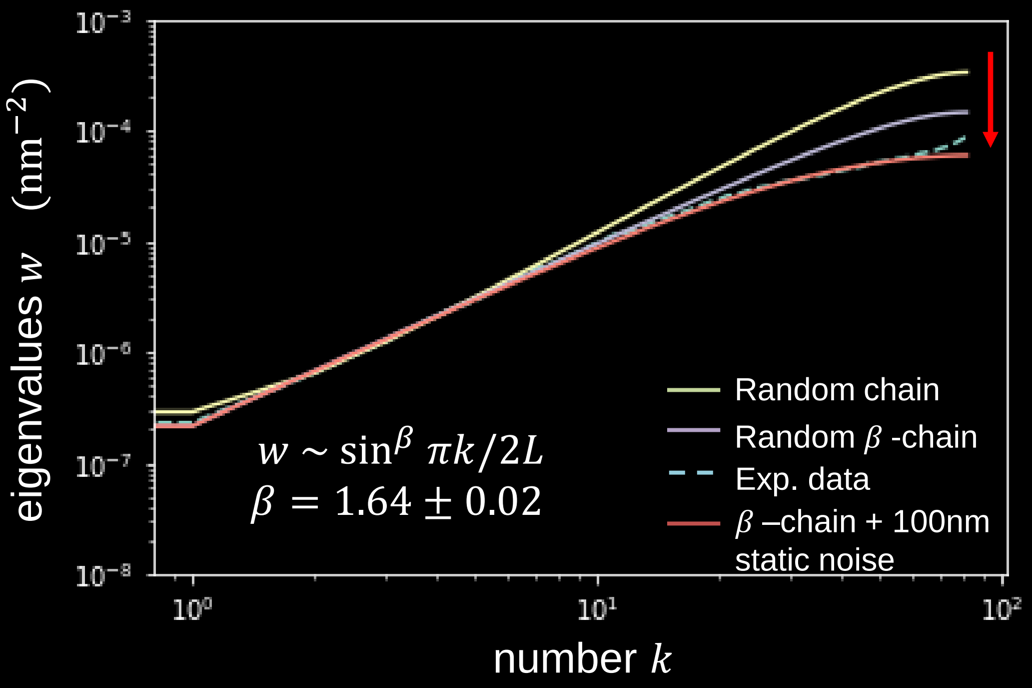
Amitai and Holcman (PRE 2013)



Adapted from *Polovnikov, Nechaev, Tamm (Soft Matter 2018)*

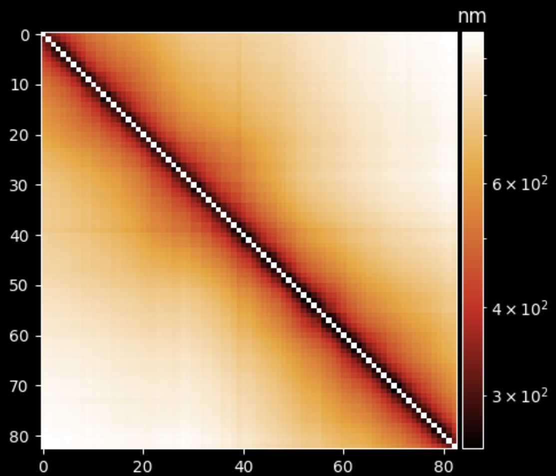
2 parameters

$$P(R) = \frac{1}{Z}\exp\left(-\frac{3}{k_B T}\boldsymbol{R}^T \mathrm{O}(\mathrm{c},\beta)\,\boldsymbol{R}\right)$$



Objective: introducing a dynamics to explain the wild type

Original data w/o cohesin



Bintu et al. (2018)

Static model

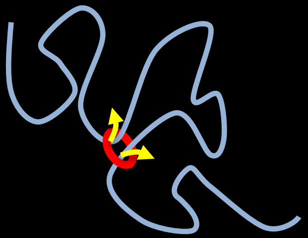
Simplified polymer model

$$P(R) = \frac{1}{Z}\exp\left(-\frac{3}{b^2}\boldsymbol{R}^T \Delta\, \boldsymbol{R}\right)$$

Simplified polymer model

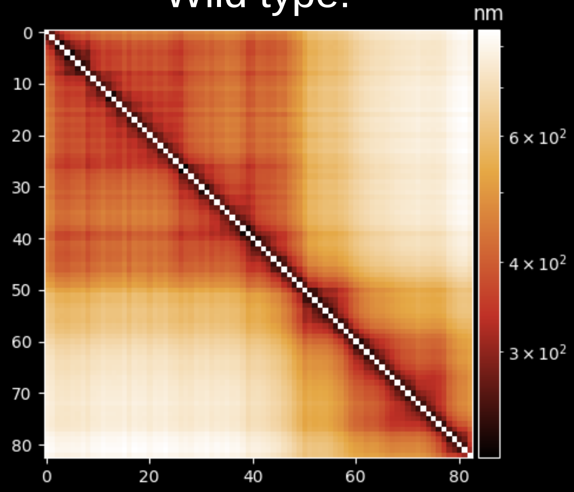
$$P(R) = \frac{1}{Z}\exp\left(-\frac{3}{b^2}\boldsymbol{R}^T \Delta\, \boldsymbol{R}\right)$$

Dynamic model



$$P(R,t)$$

Wild type:



Bintu et al. (2018)

out-of-equilibrium effects!

Explorative data-driven approach

↓

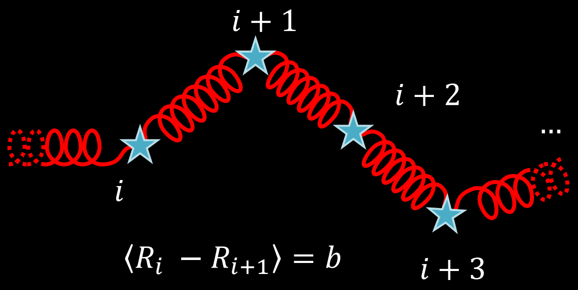
My original way of modelling loop extrusion

↓

Physical theory of chromatin conformation

Alternative approaches to model chromatin – minimal models

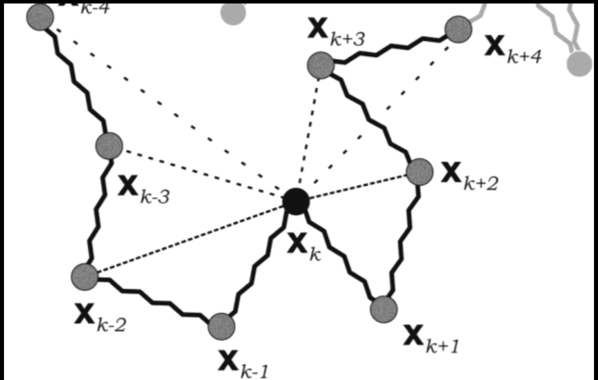
The random walk



1 parameter

$$P(R) = \frac{1}{Z} \exp\left(-\frac{3}{b^2} \boldsymbol{R}^T \Delta \, \boldsymbol{R} \right)$$

The beta chain (Amitai-Holcman 2013)



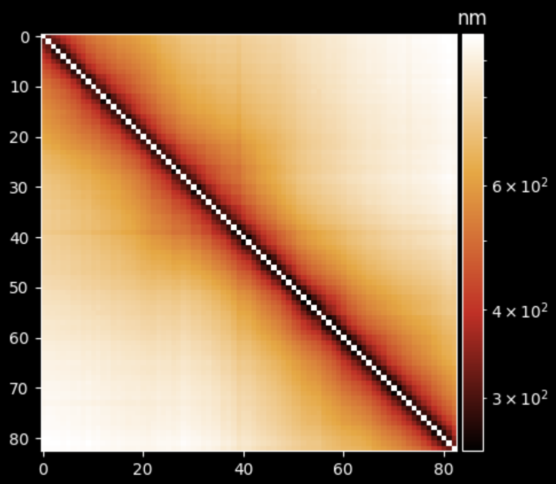
Adapted from *Polovnikov, Nechaev, Tamm (Soft Matter 2018)*

2 parameters

$$P(R) = \frac{1}{Z} \exp\left(-\frac{3}{k_B T} \boldsymbol{R}^T \mathrm{O}(\mathrm{c}, \beta) \, \boldsymbol{R} \right)$$

Objective: introducing a dynamics on top of static pictures

Original data w/o cohesin



Bintu et al. (2018)

Static model

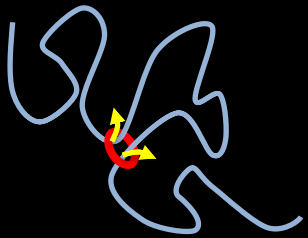
Simplified polymer model

$$P(R) = \frac{1}{Z}\exp\left(-\frac{3}{b^2}\boldsymbol{R}^T \Delta\, \boldsymbol{R}\right)$$

Simplified polymer model

$$P(R) = \frac{1}{Z}\exp\left(-\frac{3}{b^2}\boldsymbol{R}^T \Delta\, \boldsymbol{R}\right)$$

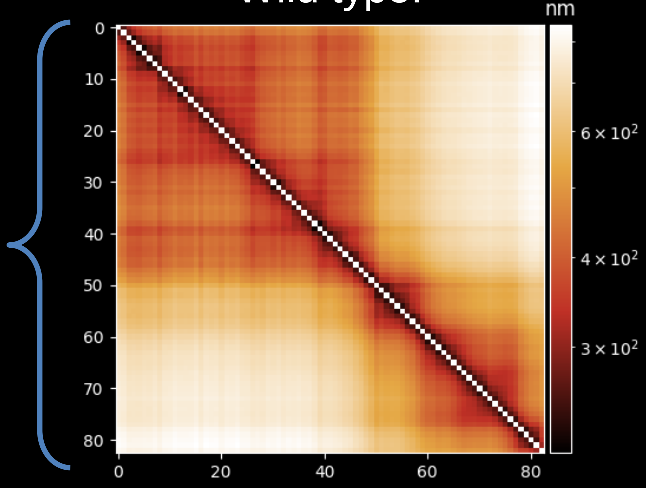
Dynamic model



out-of-equilibrium effects!

$$P(R,t)$$

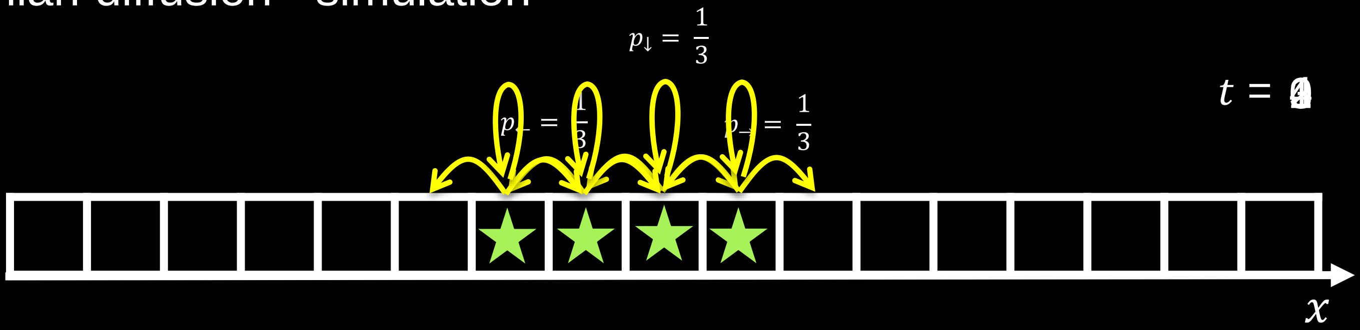
Wild type:



Bintu et al. (2018)

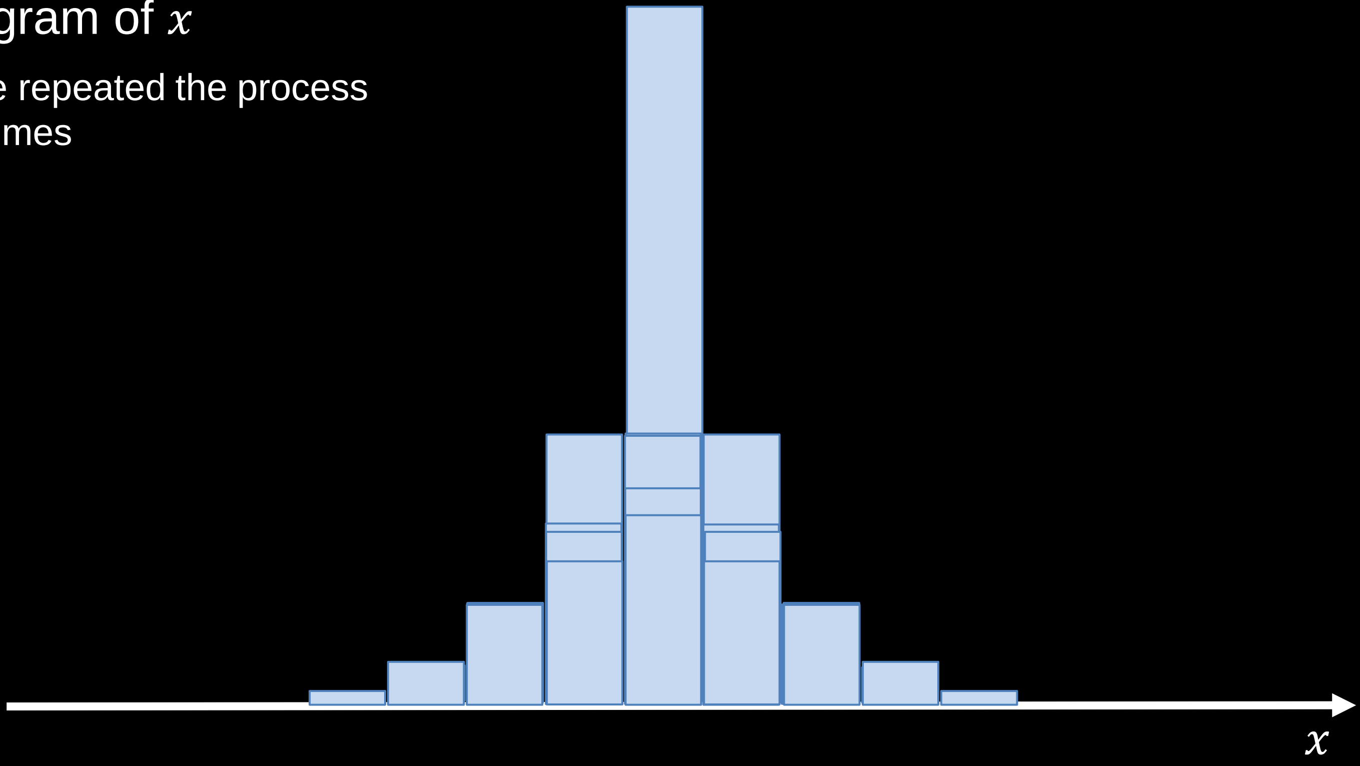
polyFlux: Original semi-analytical approach

Brownian diffusion - simulation



Histogram of x

as if we repeated the process many times



polyFlux: Semi-analytical approach

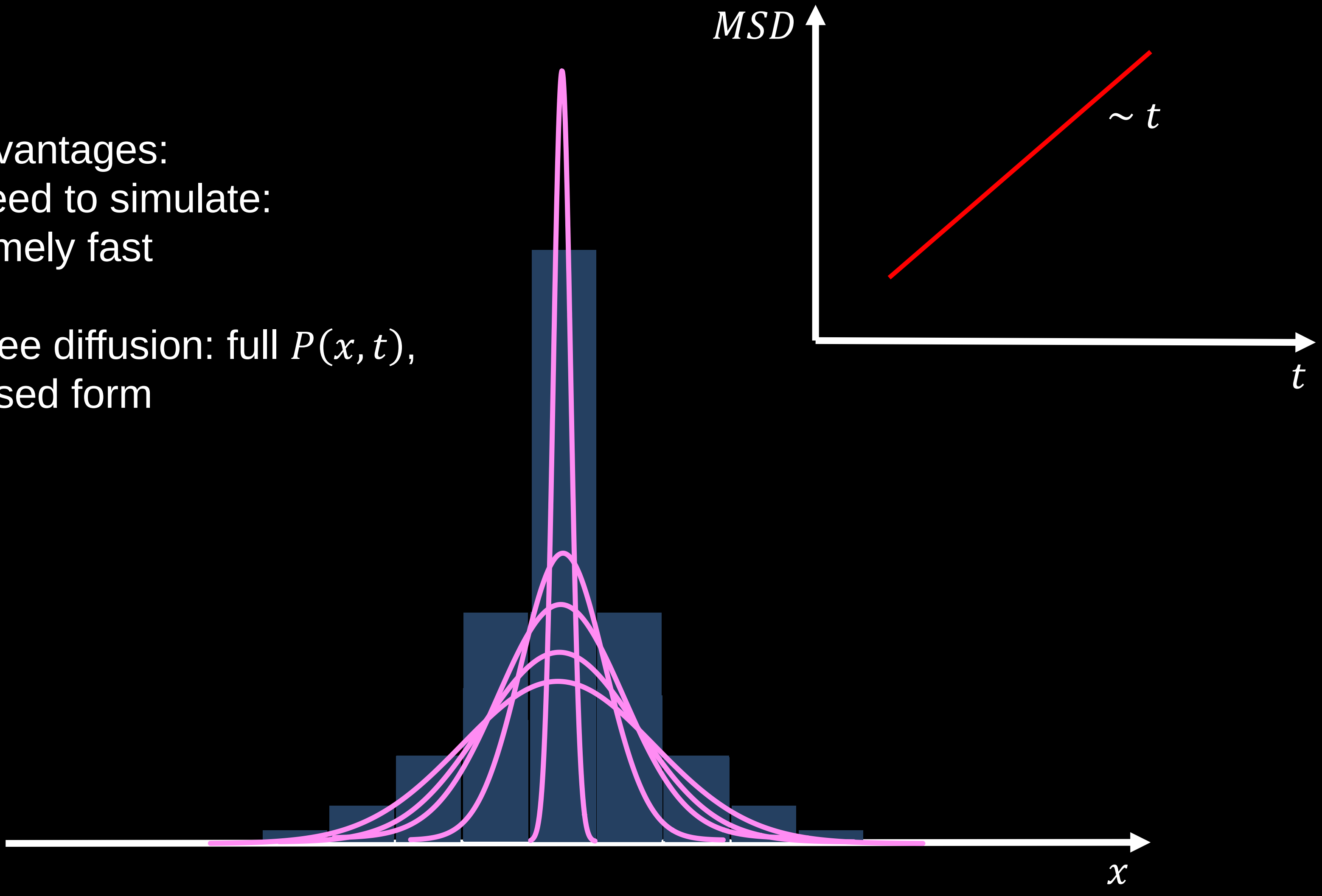
Analytical solution for x

Calculation of the Green function

Main advantages:

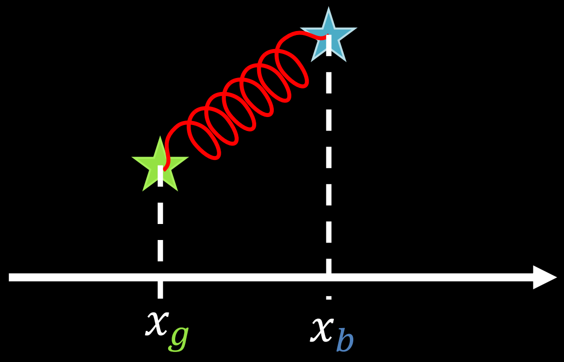
- No need to simulate: extremely fast

- For free diffusion: full $P(x,t)$, in closed form

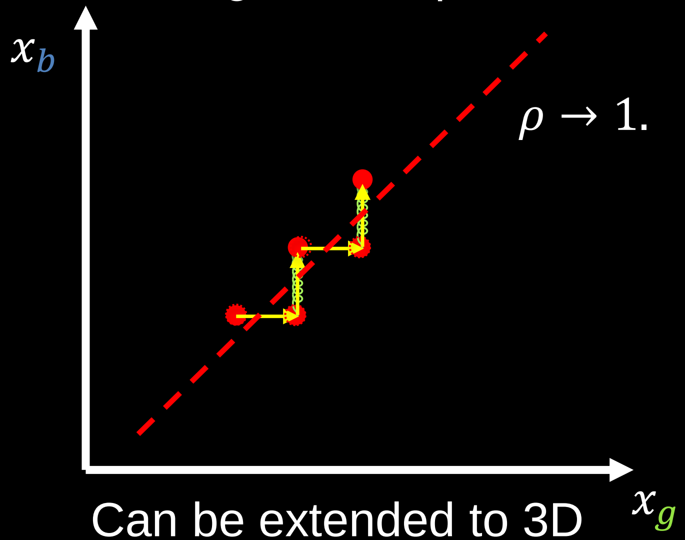


polyFlux: Original semi-analytical approach

Two linked beads, diffusion

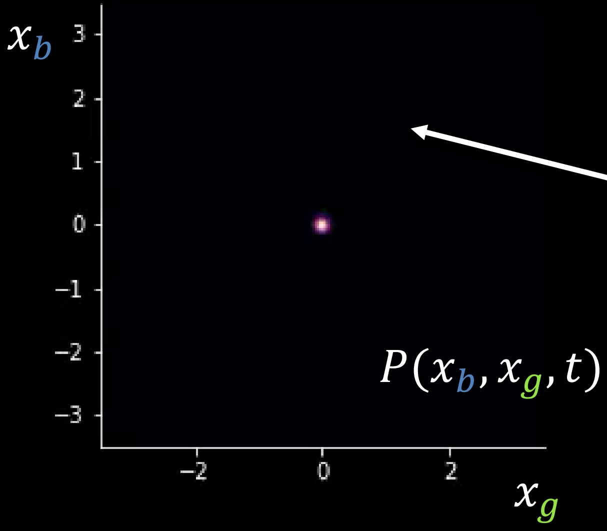


Configuration space

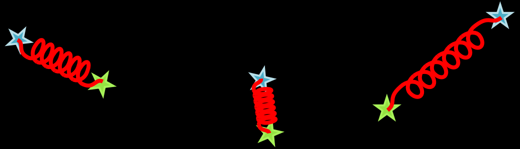


Can be extended to 3D

Analytical solution:

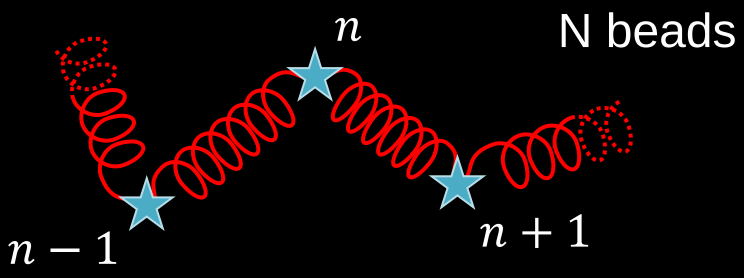


Each coordinate corresponds to a possible realization of the two linked beads



polyFlux: Original semi-analytical approach

The Rouse Model



$$\frac{\partial P(\mathbf{R},t)}{\partial t} = \sum_{i,j} \frac{\partial}{\partial R_i} L_{ij}\left(k_B T \frac{\partial P}{\partial R_j} - k\hat{O}_{ij} R_j P\right)$$

***Faster* relaxation**		***Slower* relaxation**	
$k \cdot \hat{O}_{ij}$ $k_B T$	Chromatin elasticity Thermal fluctuations	L_{ij}	Mobility (Nucleoplasm viscosity) normally $1/\zeta \cdot \delta_{ij}$

It is the simplest polymer model
1 extra parameter, viscosity ζ

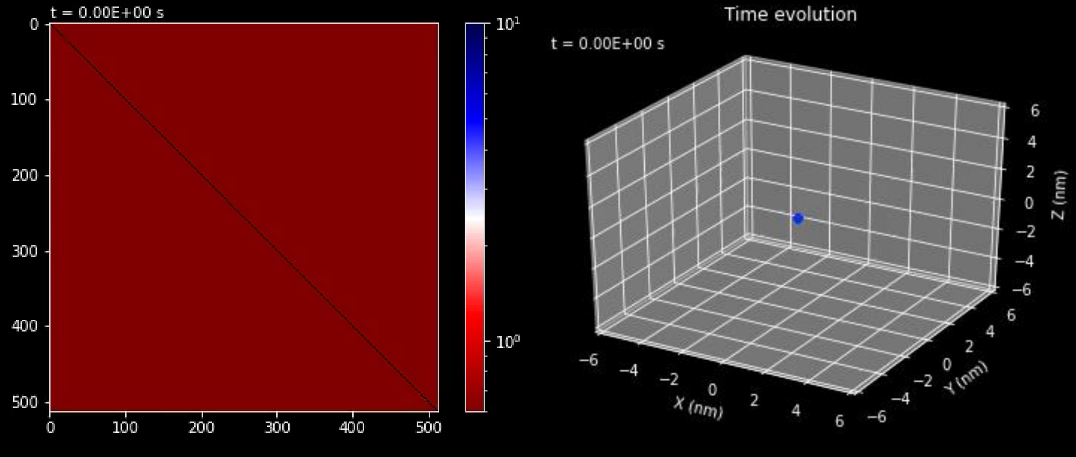
Provides:

- Connectivity
- Static and Dynamical aspects

Ignores:

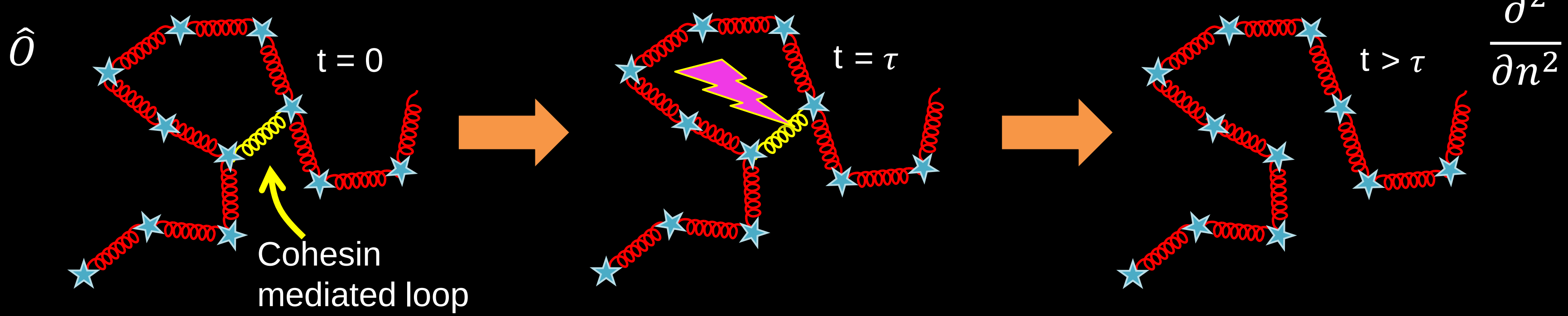
- Steric and volume excluded effects
- Topological constraints

Calculation of the Green function (propagator)


Predicted distances

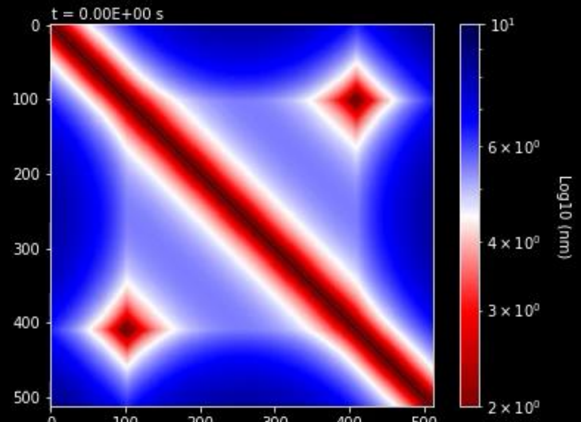
Simulation

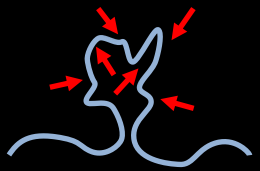
Loops in the analytical Rouse Model



Cohesin mediated loop

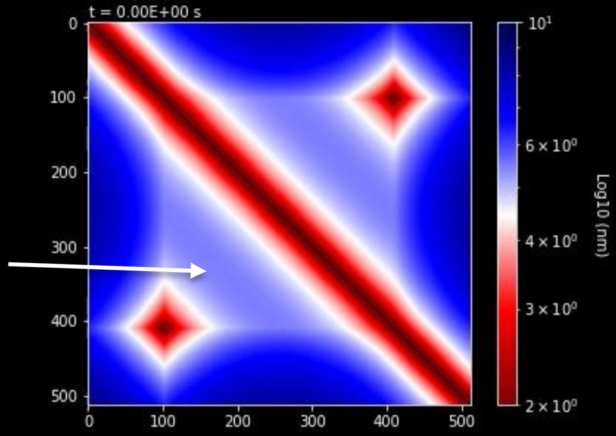
Fast relaxation



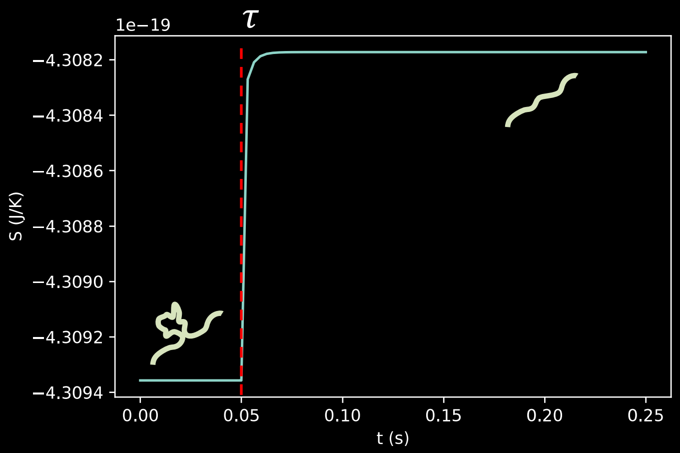


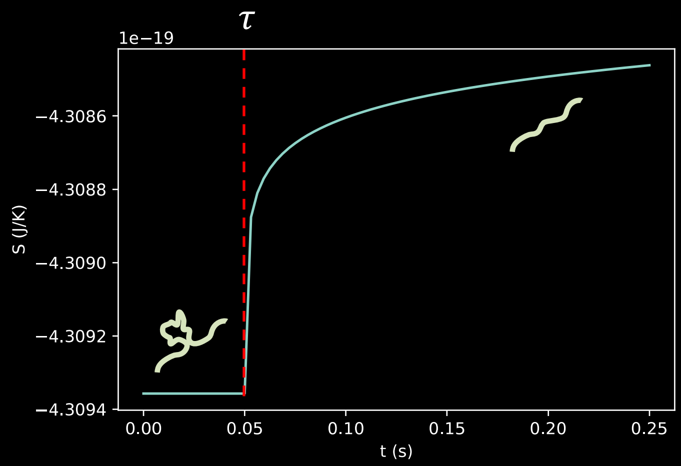
Conformational memory

Slow relaxation



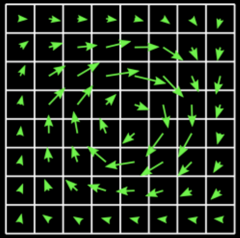
Entropy



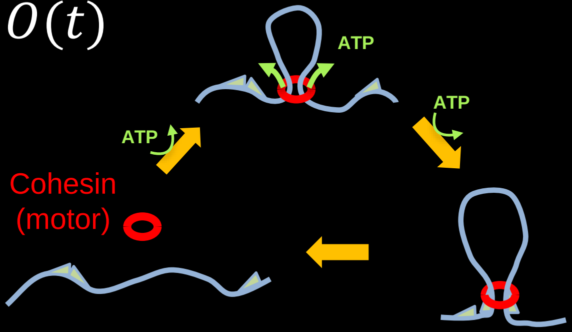


Changing network topology *in time*, we obtain loop extrusion (in a toy-model form)

Non equilibrium steady state

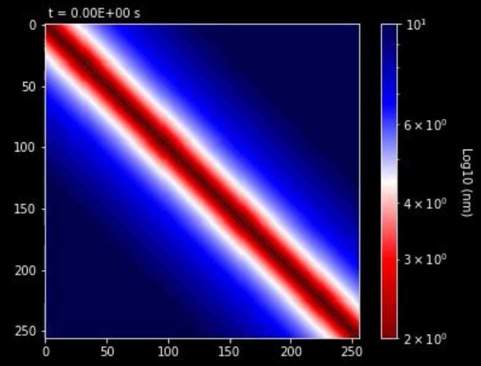


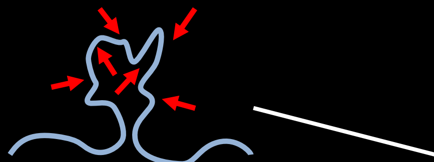
$\hat{O}(t)$



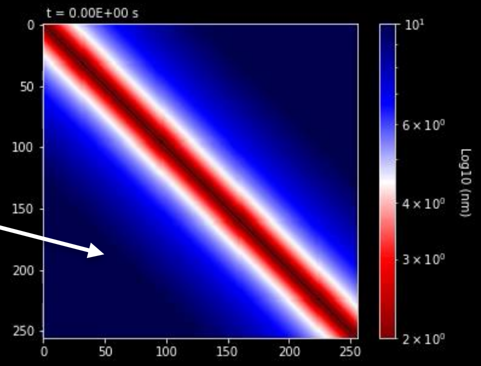
Slow extrusion

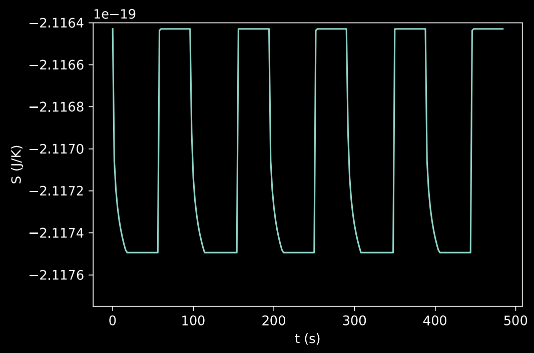
~ 100x Faster extrusion

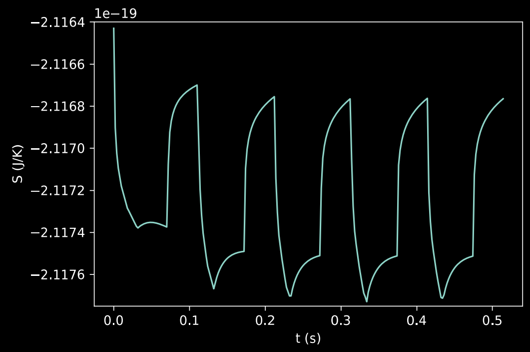




Conformational memory





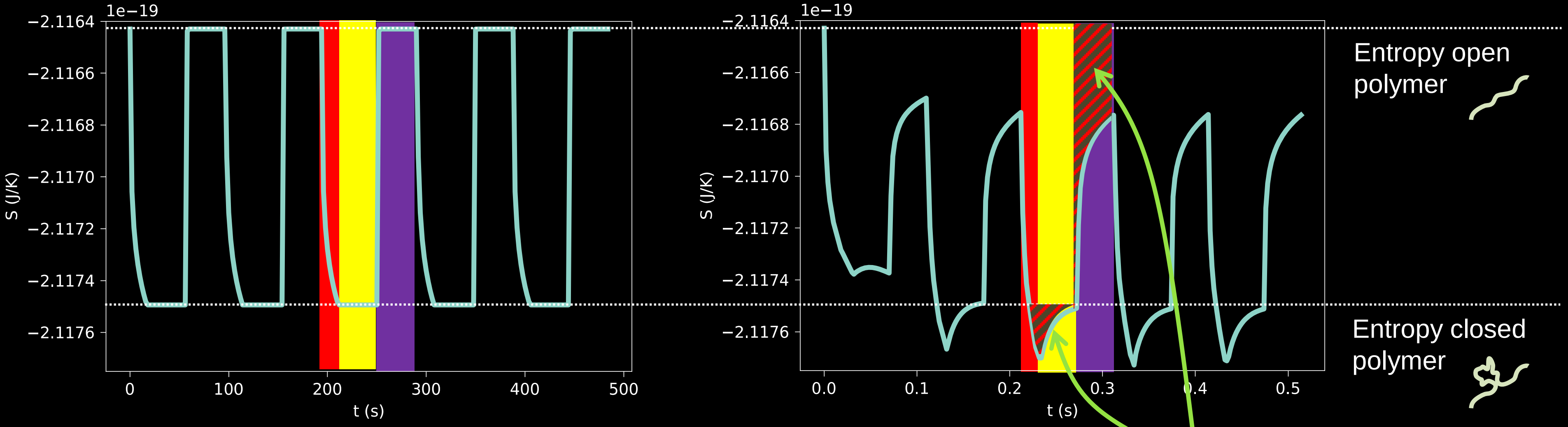


What are the quantitative effects of interplay between activity and relaxation with broken detailed balance?

Dissecting the entropy dynamics

Slow extrusion

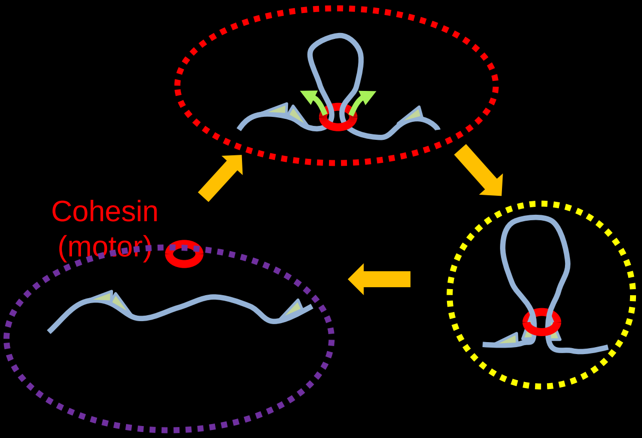
~ 100x Faster extrusion



Entropy open polymer

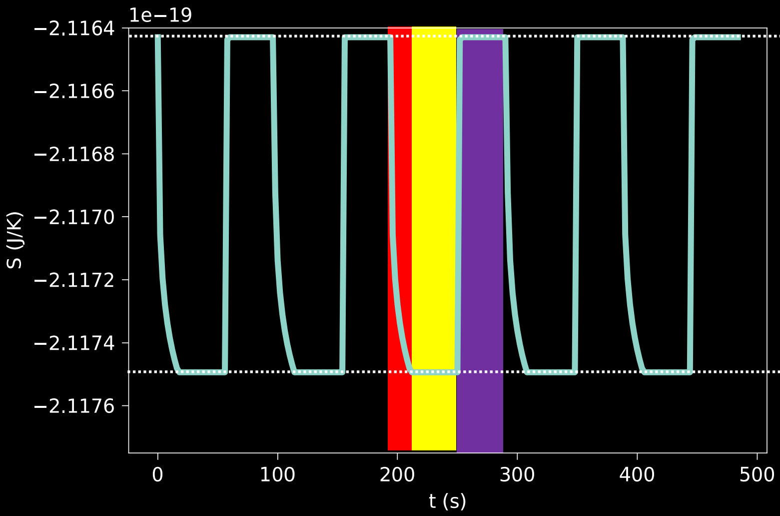
Entropy closed polymer

Out of equilibrium effects related to relaxation

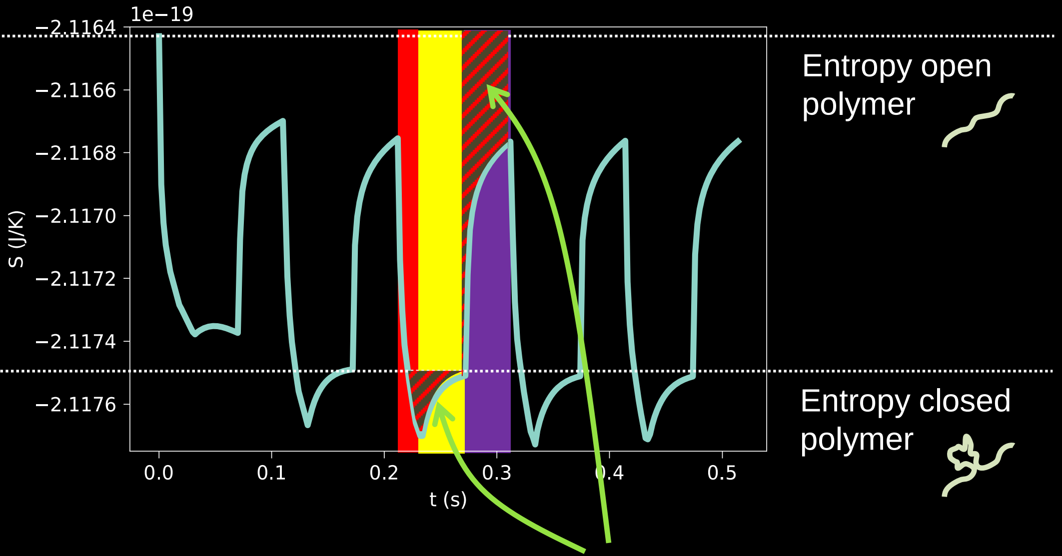


Dissecting the entropy dynamics

Slow extrusion



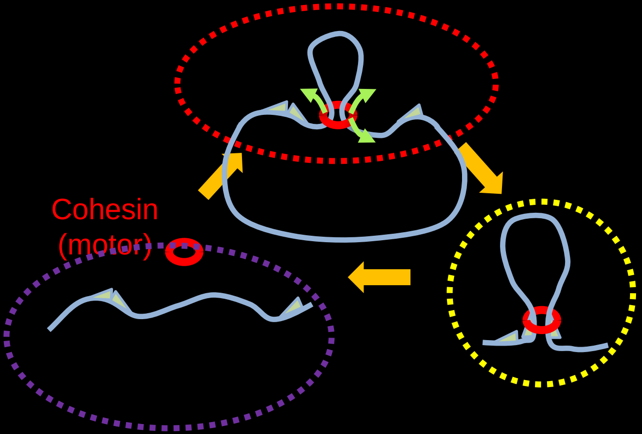
~ 100x Faster extrusion



Entropy open polymer

Entropy closed polymer

Out of equilibrium effects related to relaxation

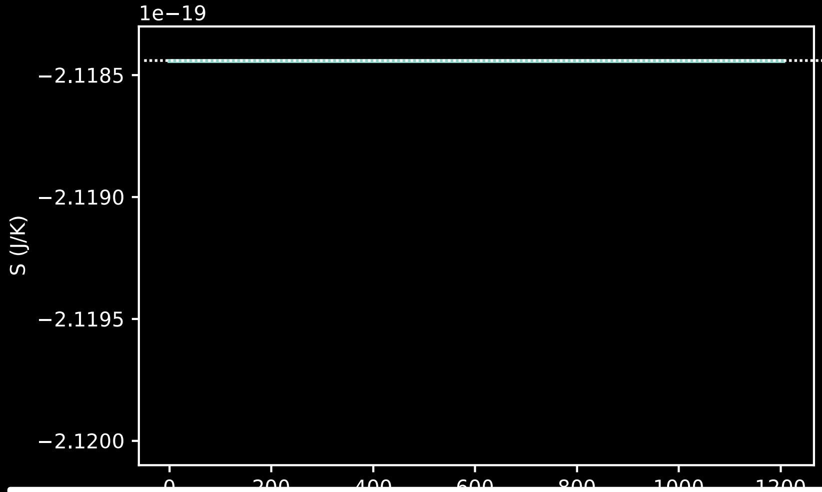


An even simpler toy-version of loop extrusion?

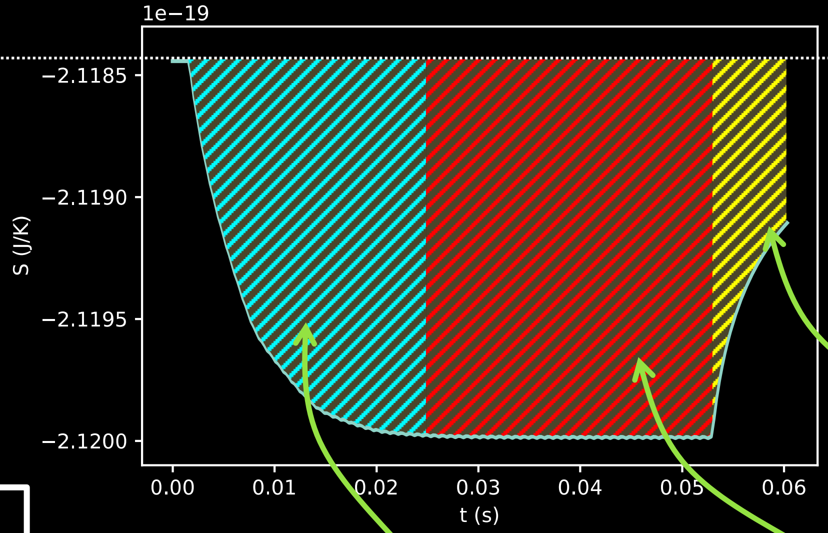
The model of loop-transition

An even simpler toy-version of loop extrusion, to theoretically visualize the breaking of detailed balance

Slow extrusion



~ 1000x Faster extrusion

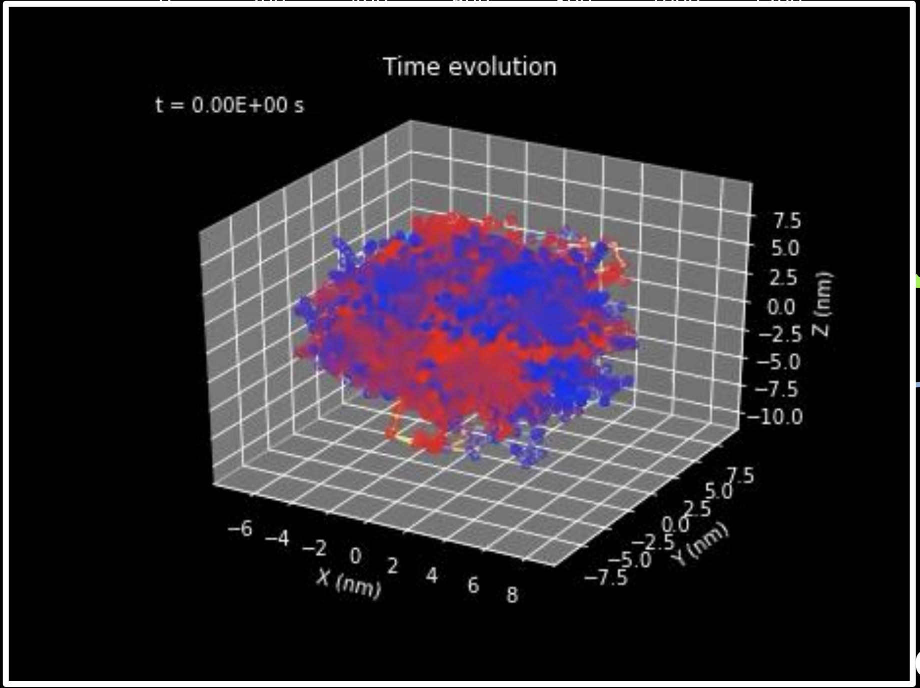


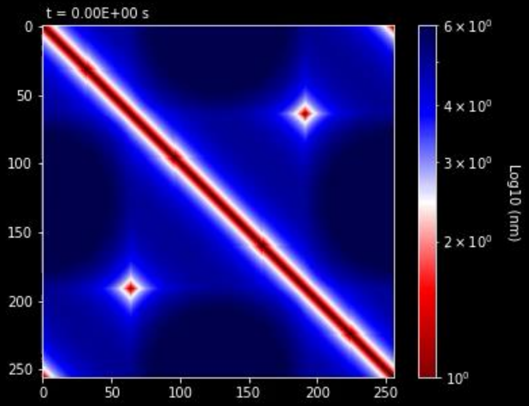
Entropy with loops

relaxation

transition

Out of equilibrium steady state



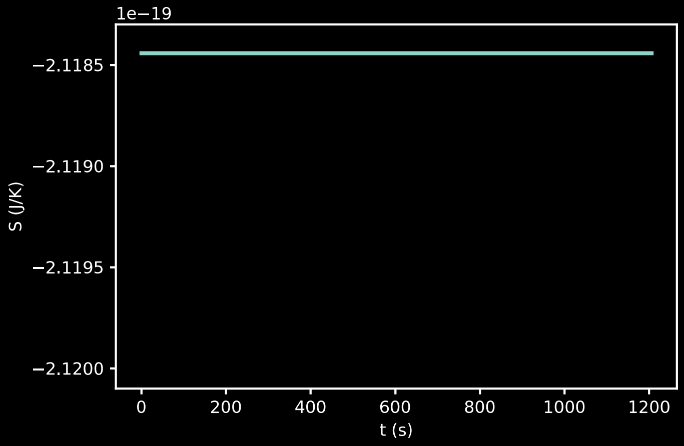


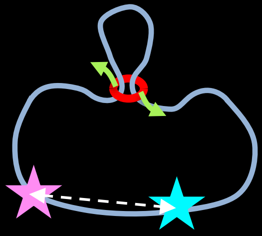
4x cycles

op-transition

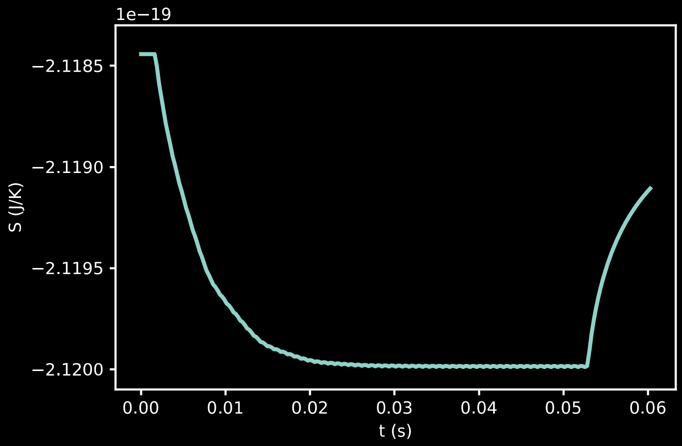
Theoretically visualize the breaking of detailed balance

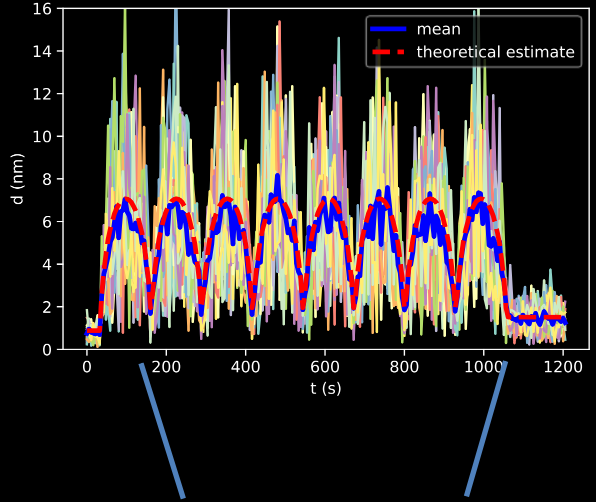
Slow extrusion



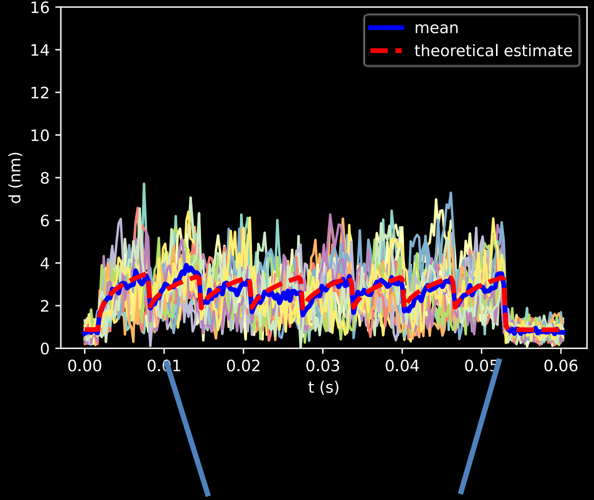


~ 1000x Faster extrusion





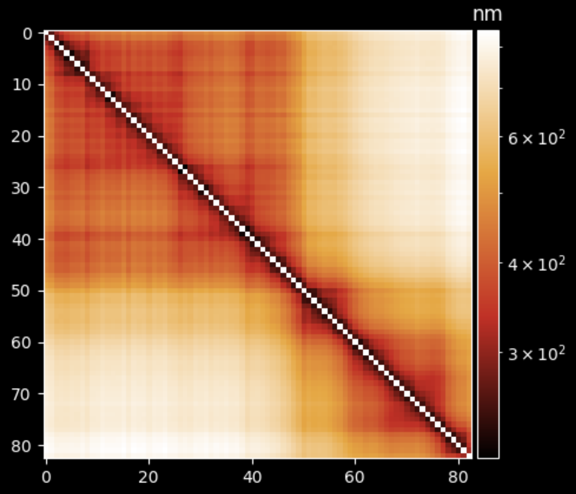
Time reversal symmetry



Breaking of time reversal symmetry

What does it takes to reproduce the following kind of data?

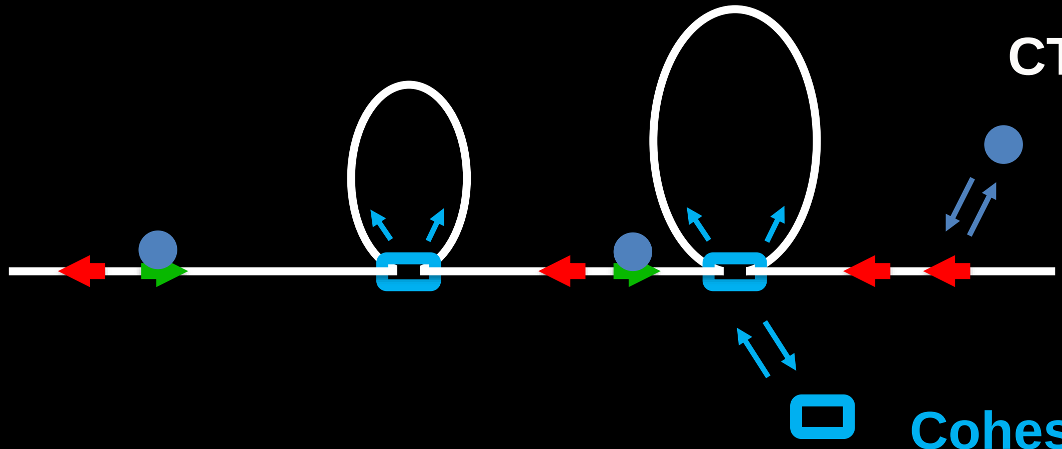
Mean distance matrix



Bintu et al. (2018)

simLoop: Minimal model of loop dynamics (secondary structure)

5 parameters



CTCF

- **Position, orientation** → ChIP-seq + motifs
- **kon** → same for all sites
- **koff** → occupancy from ChIP-seq data

Cohesin

- **kon**
- **koff**
- **speed**

Hypotheses:

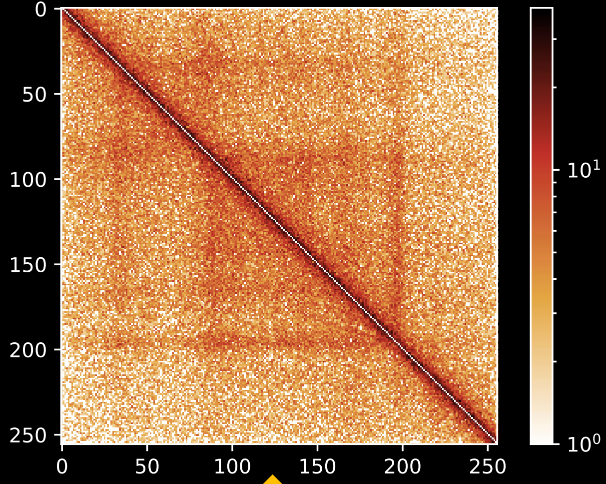
- Cohesin loading → uniform … biased?
- Extrusion → two-sided … one-sided?
  → constant speed … biased random walk?
- Upon collision
  cohesin/cohesin → stalling dissociation? stabilization?
  cohesin/CTCF → stalling stabilization?

Going 1D to 3D: Combining simLoop with polyFlux

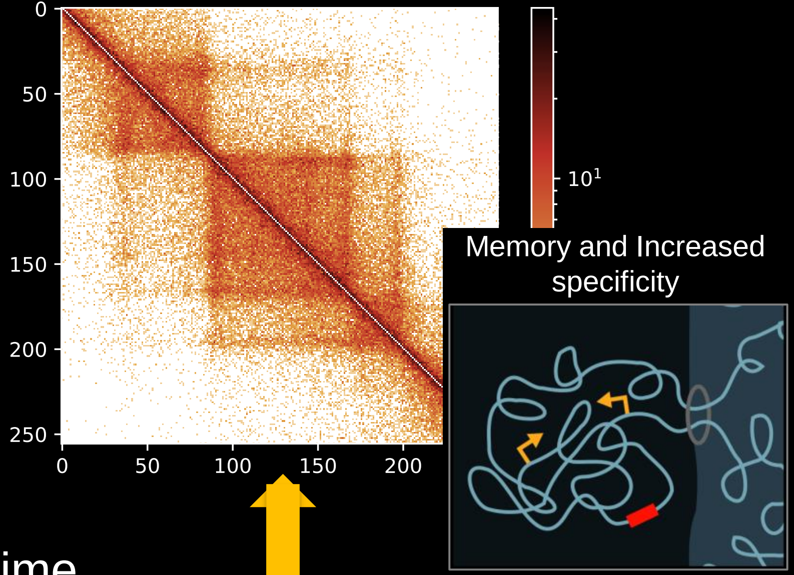
simLoop input

3D contact maps

Fast relaxation



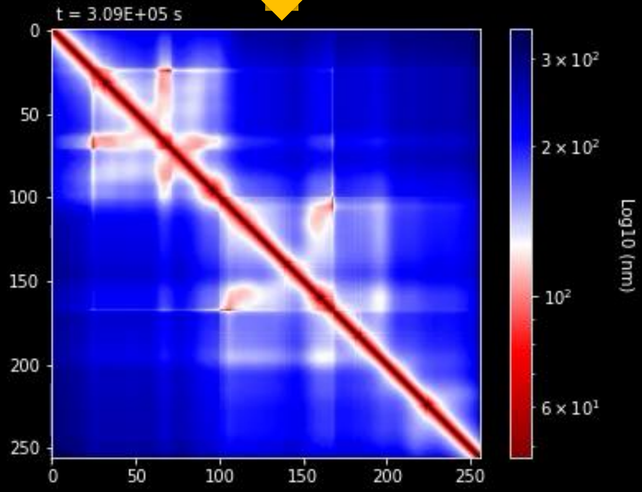
Slow relaxation



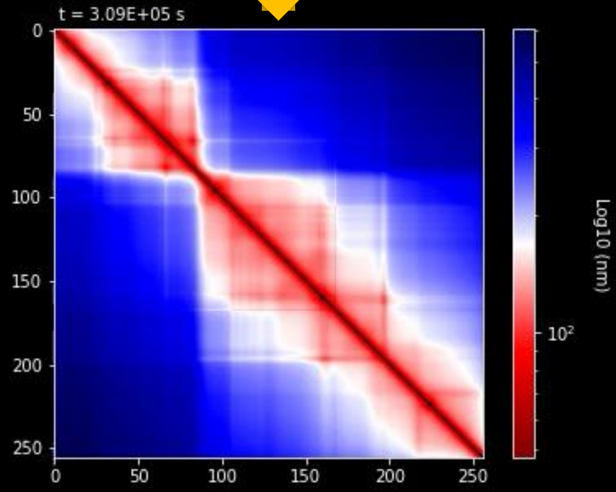
Memory and Increased specificity

polyFlux distributions in time

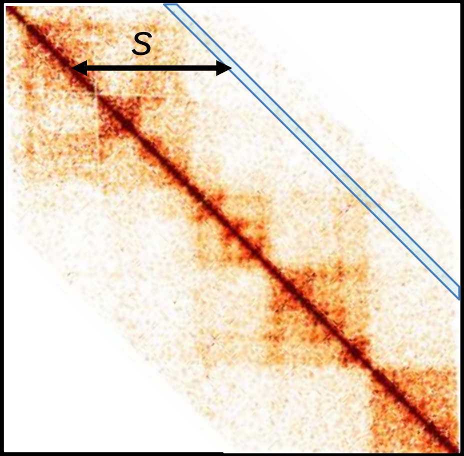
$\int dt$

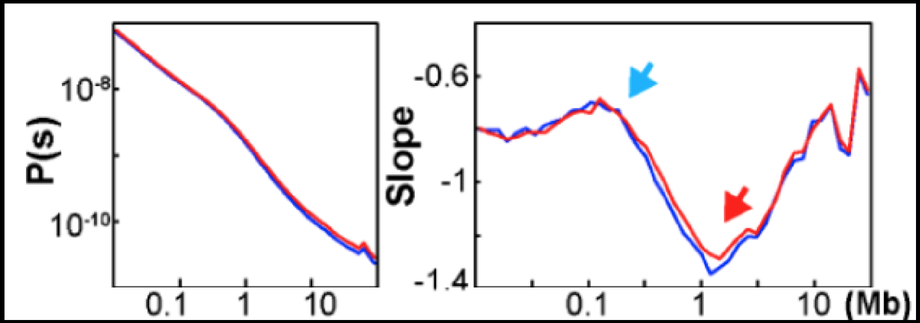


$\int dt$



Out-of-equilibrium effects reproduce the main landmarks of contact maps at the TADs scale

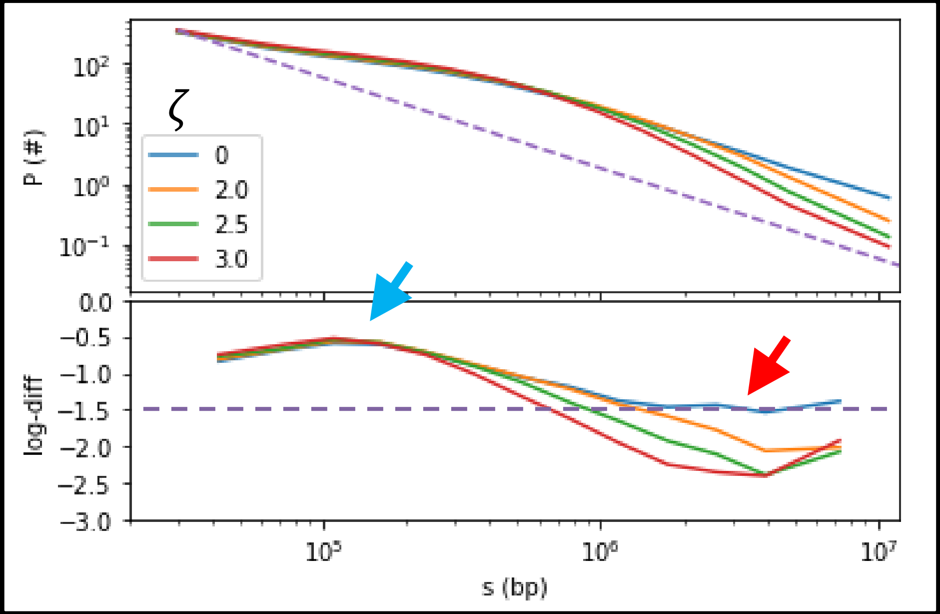




From Liu and Dekker (Nat Cell Biol) 2022

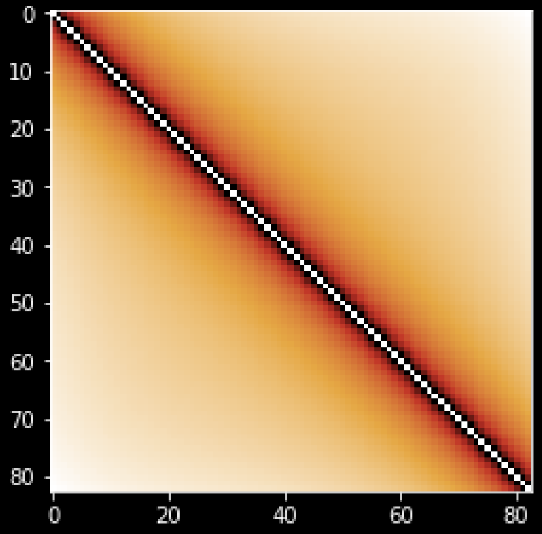
See also Gassler et al. (EMBO) 2017 and Scolari et al. (PRL) 2018

Simulations, P(s) in function of viscosity scale (log10)

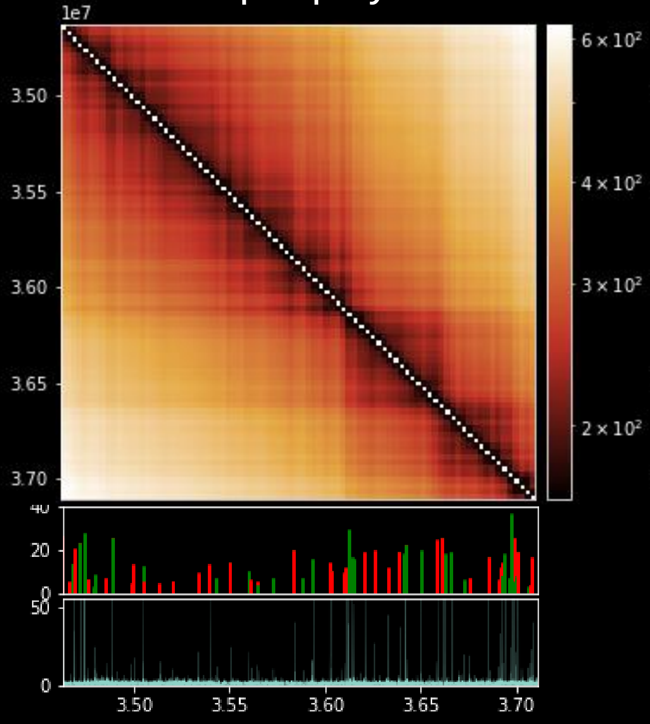


The allows to simulate contact maps within minutes

No cohesin model

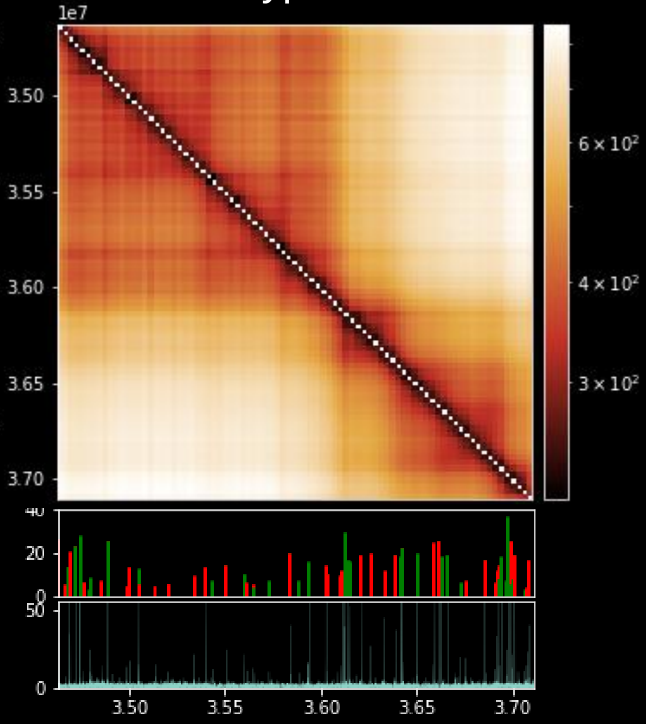


simLoop + polyFlux



HCT116, chr21, 1e7·bp

Wild type data



Oligopaint [Bintu et al. (2018) *Science*]

CTCF ChIP [Rao (2017) Cells]

Future: free and open source, community driven, development

Work in progress!!

Thanks!

Institut Curie

Genome functions in Space and Time Group

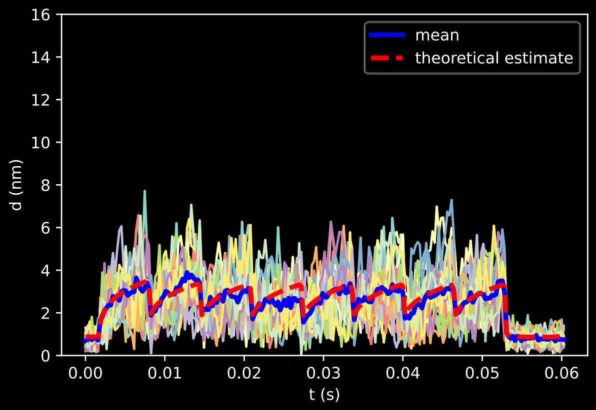
The Physics unit (PCC) – theory and experiments

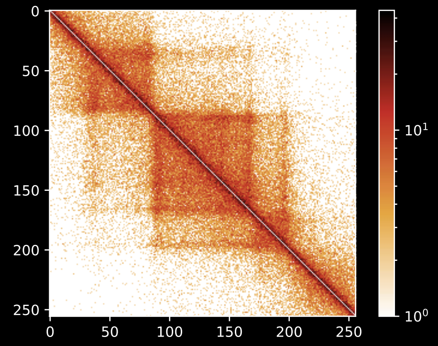
Ecosystems -> Full-organism -> Tissues -> Cells -> Molecules

The Nuclear dynamics unit:

We do the anything about the (cell) nucleus

Forces shaping chromatin in the nucleus





Credits: CoulonLab, funding bodies

Antoine Coulon, Kyra Borgman, Julia Ronsch, Ilham Ladid
Ex members that I thank: Lorena Kolar, Veer Keizer

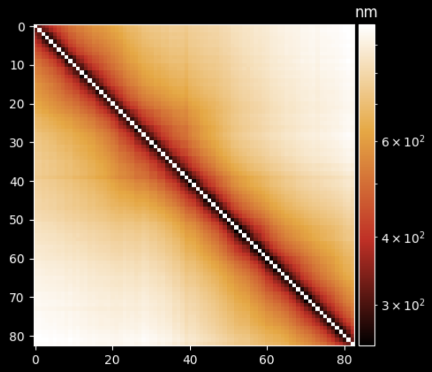
Contacts:
vittore.scolari@curie.fr

Is the model missing anything?

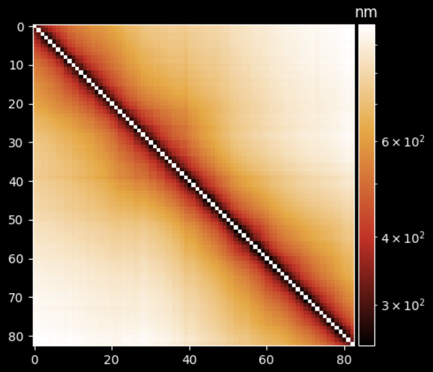
Original data

$$P(R) = \frac{1}{Z}\exp\left(-\frac{1}{2}\boldsymbol{R}^T\,\Xi^{-1}\boldsymbol{R}\right)$$

Distance matrix

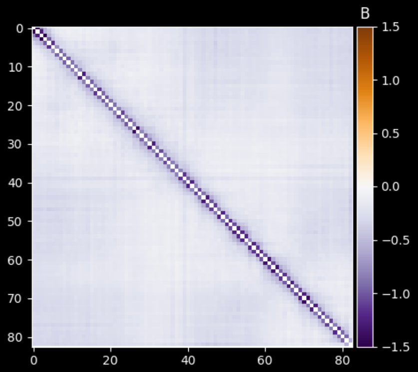


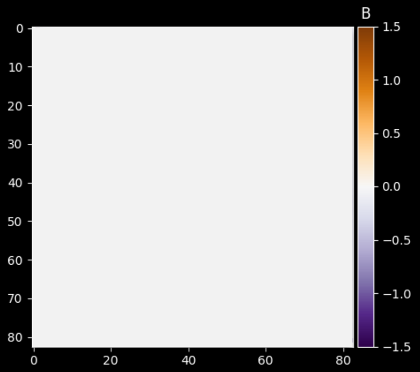
Bintu et al. (2018)



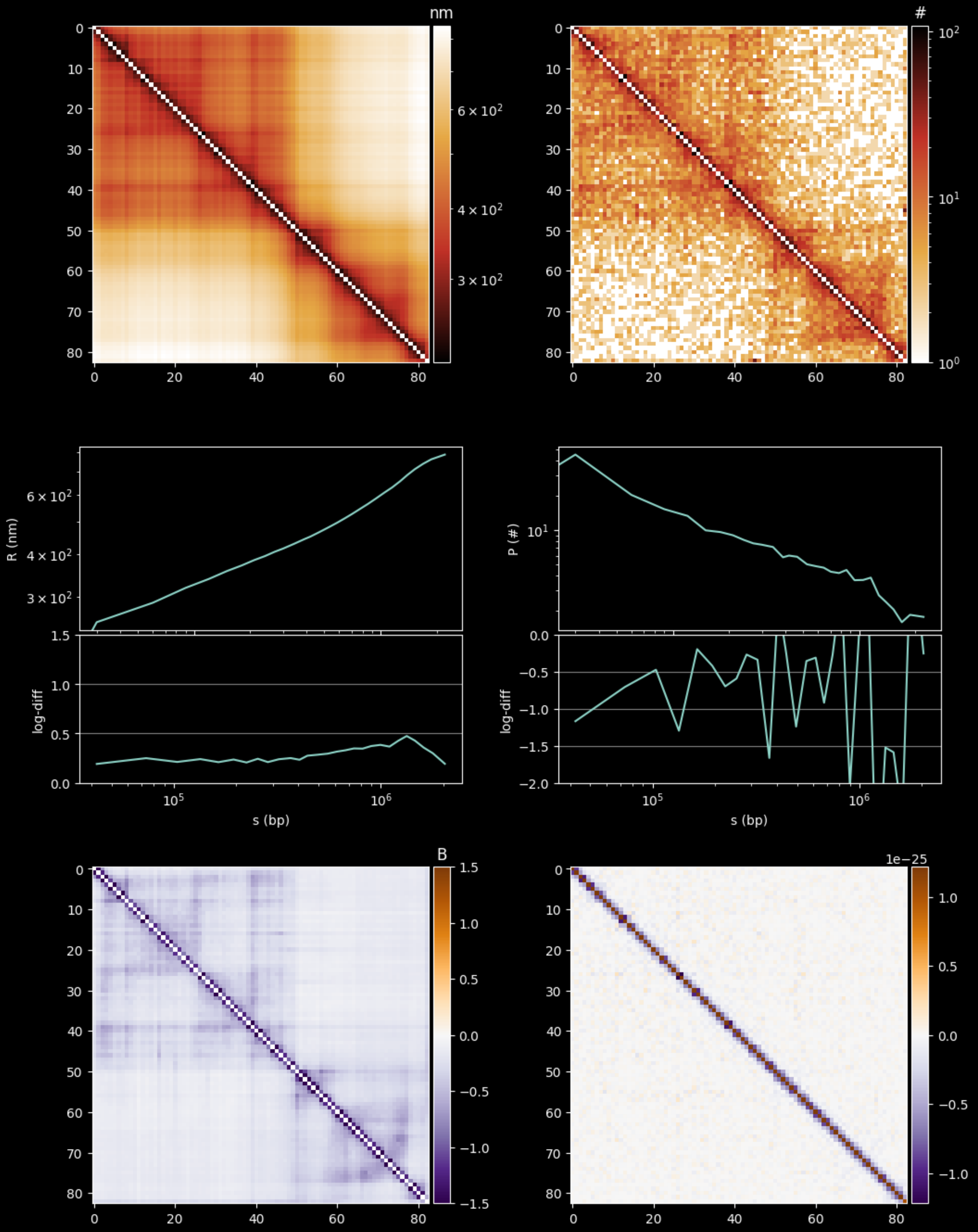
6889 parameters to “fit” 6889 mean observables

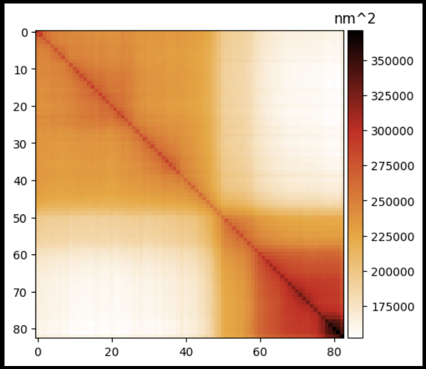
Binder cumulant





the Gaussian is a good model


nm
6 × 10^2
4 × 10^2
3 × 10^2
#
10^2
10^1
10^0
R (nm)
P (#)
log-diff
s (bp)
10^5
10^6
B
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
1e−25
0
20
40
60
80


nm^2
350000
325000
300000
275000
250000
225000
200000
175000
0
20
40
60
80