INSTITUTE FOR MATHEMATICS

STRUCTURES CLUSTER OF EXCELLENCE

UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386

Michael Bleher

Institute for Mathematics, Heidelberg University

– DAGSTUHL 24122 – MARCH 20, 2024 –

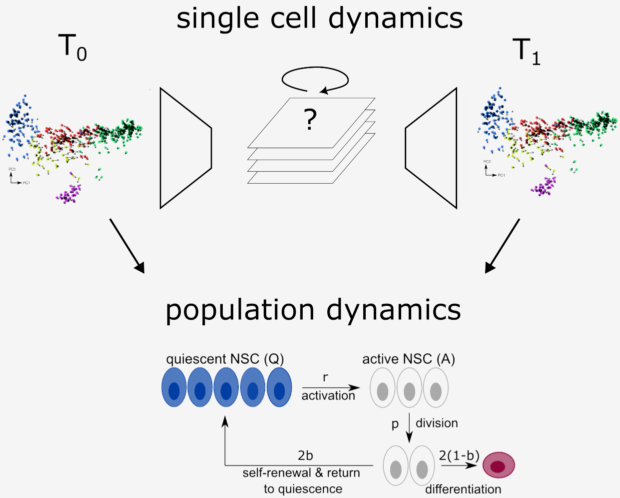
RNA VELOCITY EMBEDDINGS IN CURVED SPACES

EXPLORING CELLULAR DYNAMICS

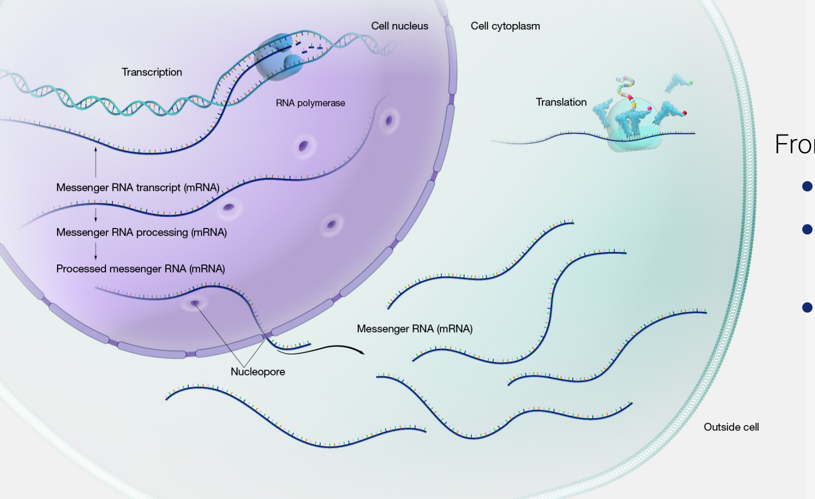
From single cells to populations

joint w/

Anna Marciniak-Czochra (Heidelberg)
Anna Wienhard (MPI Leipzig)
Fred Hamprecht (Heidelberg)



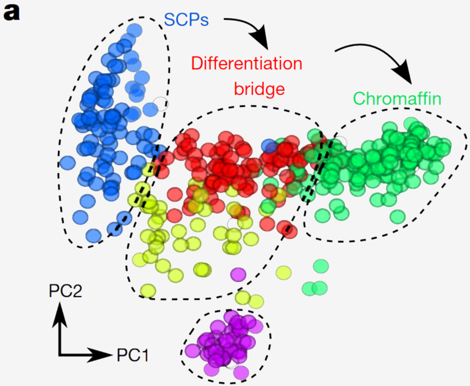
Gene Expression



From code to function

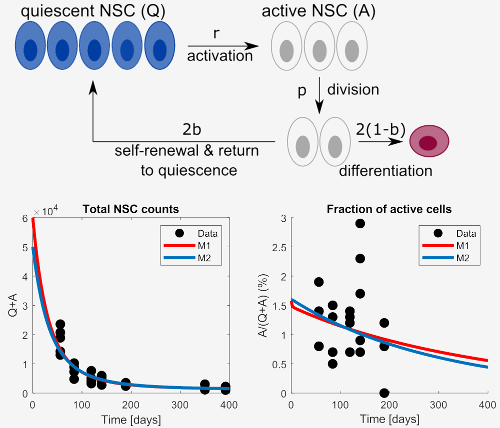
- DNA $\rightarrow$ mRNA $\rightarrow$ proteins
- gene expression
  $\simeq$ # mRNA snippets
- proxy for cell's current biological state $x_i \in \mathbb{R}^N$

Population Dynamics



La Manno, G. et al. (2018) 'RNA velocity of single cells', Nature, 560(7719), pp. 494–498. Fig 2.

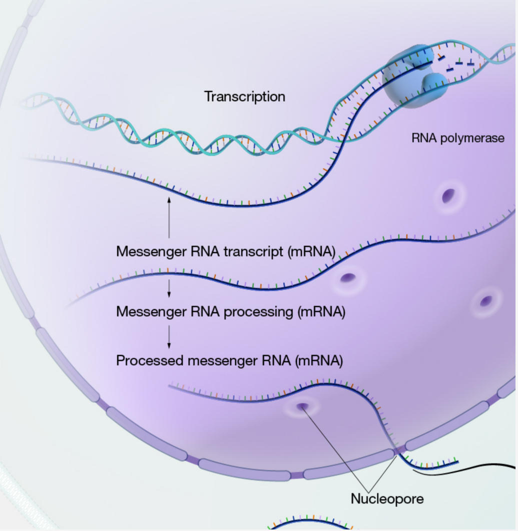
Model schematic

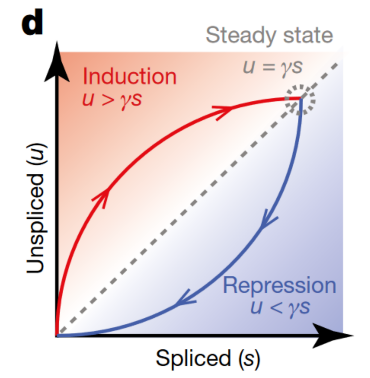


Danciu, D.-P. et al. (2023) 'Mathematics of neural stem cells: Linking data and processes', Cells & Development, 174, p. 203849. Fig 1 & Fig 5A.

"Standard" RNA velocity

La Manno, G. et al. (2018) 'RNA velocity of single cells', Nature, 560(7719), pp. 494–498. Fig 1.





$\rightsquigarrow v_i \in \mathbb{R}^N$ **RNA velocity**

"Standard" Visualizations

Given RNA-velocity data

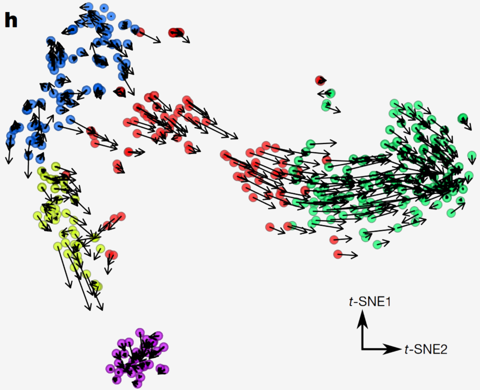
$$(x_i, v_i) \in \mathbb{R}^N \times \mathbb{R}^N$$

1. get low dimensional representation

$$x_i \mapsto y_i \in \mathbb{R}^2$$

2. "pushforward" of velocities $v_i \mapsto w_i \in \mathbb{R}^2$ st.

$$\mathrm{Similarity}(v_i, x_j - x_i) \simeq \mathrm{Similarity}(w_i, y_j - y_i)$$



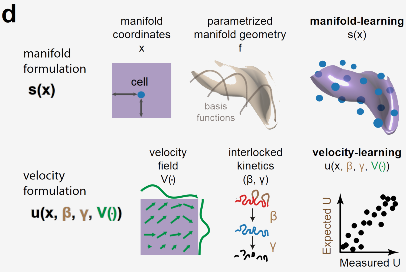
La Manno, G. et al. (2018) 'RNA velocity of single cells', Nature, 560(7719), pp. 494–498. Fig 2.

Common Criticism of "standard" approach

RNA velocities v_i are not tangent to data manifold.

Manifold Constrained RNA Velocity

1. Choose cell-state manifold M (e.g. $M = S^1$)
2. Assume gene expressions depend only on M
3. compute representation of scRNA-seq data in M (think: coordinization $x(y)$, $s(y)$, $u(y)$)
4. infer RNA velocity in low-dimensional representation
5. (optional) pull velocities back to $\mathbb{R}^N$ for downstream analysis



$$(\nabla_x s_g) \cdot V(x(t)) = \beta_g u_g(x(t)) - \gamma_g s_g(x(t))$$

Lederer, A.R. et al. (2024) 'Statistical inference with a manifold-constrained RNA velocity model uncovers cell cycle speed modulations'.

Successfully tested for cell cycle. RNA velocities generally point in the expected direction (and show interesting speed modulations).

Taking a step back

"Standard" position and velocity pairs are points in tangent bundle (aka phase space)

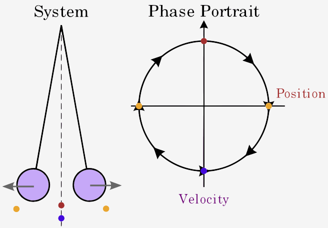
$$(x_i, v_i) \in \mathbb{R}^N \times \mathbb{R}^N = T\mathbb{R}^N$$

1. Manifold assumption

Data is noisy sample from $TM \hookrightarrow T\mathbb{R}^N$
$\rightsquigarrow$ low-dim representations in tangent bundles

2. "Geometric Dynamics" assumption

Time evolution is determined by flow on TM
$\rightsquigarrow$ can flows in dimensionally reduced representation capture principal dynamical components?



Note: This kind of data is ubiquitous

- Meteorology (pressure and wind velocity)
- Astronomy (space velocity of stars)
- Velocity Obstacle Problem in robotic motion planning
- Traffic Flow Dynamics
  (often includes acceleration
  $\rightsquigarrow (x_i, v_i, a_i) \in J^2 M$ second Jet bundle)

The Sasaki Metric

(M, g) – Riemannian manifold.

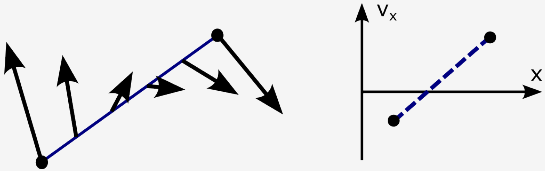
Naturally induced **Sasaki Metric** on TM

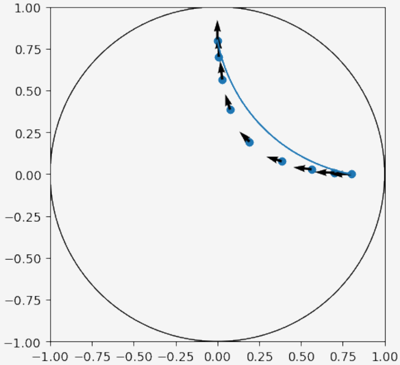
$\exists$ 2-parameter family of natural metrics

$$g_{\text{Sasaki}} = g \oplus g : T(TM) \simeq \underbrace{TM}_{\text{horizontal}} \oplus \underbrace{TM}_{\text{vertical}} \to \mathbb{R}$$

Examples

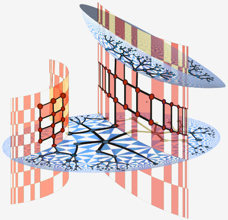
- Euclidean space $M = \mathbb{R}^N, TM = \mathbb{R}^{2N}$, $g_{\text{Sasaki}} = g_{\mathbb{R}^{2N}}$





- Poincaré disk

Choice of Latent Space




- trees fit into Hyperbolic spaces ($\kappa < 0$)
- grids fit into Euclidean spaces ($\kappa = 0$)
- cycles fit into Spheres ($\kappa > 0$)

Idea

Provide several components $M = M_0 \times \ldots \times M_k$ with different curvature to make data feel at home. E.g. Symmetric Spaces $\rightsquigarrow$ decompose into $M = M_{\kappa \leq 0} \times \mathbb{R}^n \times M_{\kappa \geq 0}$.

(also simplifies a bunch of other things)

Intuition

Good approximation of main geometric content of the data; denoising of irrelevant curvature.

E.g. with metric MDS $\simeq$ **principal curved coordinate analysis**

Low-dimensional representation with interpretable geometry, can use geometric tools.

Hamiltonian Single Cell Dynamics ...

Time evolution of a point $p = (x, v) \in TM$ is determined by a Hamiltonian flow

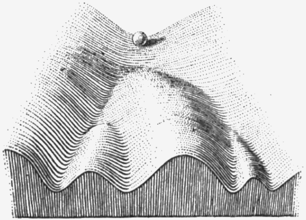
$$\dot{\gamma} = X_H \, , \; \gamma(0) = p$$

Here H is some Hamiltonian ($\simeq$ energy) and X_H is the associated Hamiltonian vector field defined by

$$\omega(X_H, \cdot) = dH$$

Example

- $H(x, v) = \frac{1}{2} g(v, v) \rightsquigarrow$ Geodesic Flow
- $H(x, v) = V(x) \rightsquigarrow$ Waddington's landscape



... and Magnetic Systems

Hamiltonian dynamics depends on choice of symplectic structure

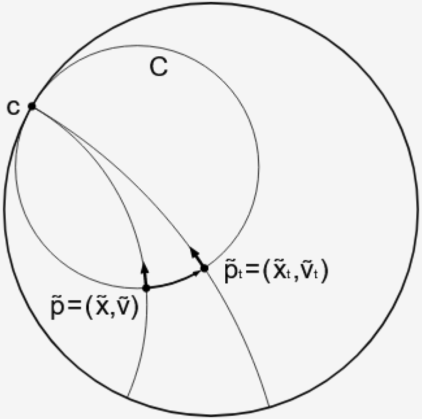
- standard symplectic structure on T^*M

$$\omega_0 = dp \wedge dq$$

- deformation by magnetic field $\mu \in \Omega^2(M)$

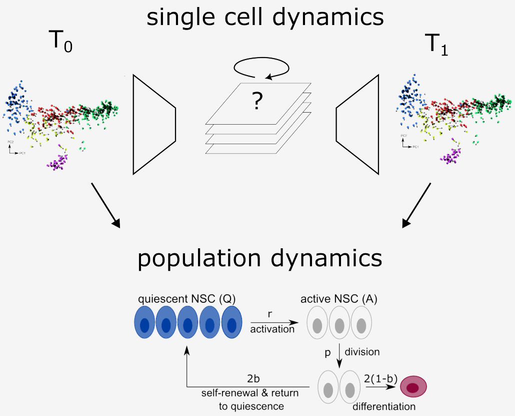
$$\omega = \omega_0 + \pi^* \mu$$

Magnetic fields can lead to **motion on cycles**
$\rightsquigarrow$ model for cell cycle?



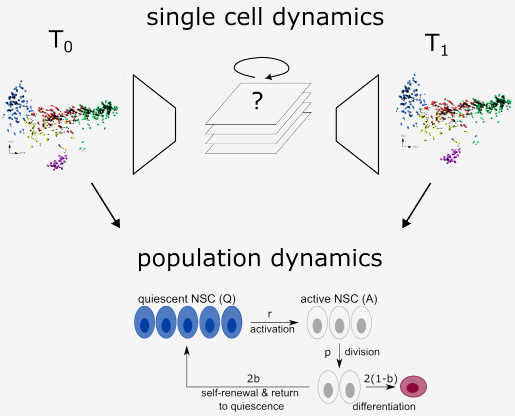
Summary

- RNA velocity embeddings into Sasakian Geometry
- Principal curved coordinate analysis
- Hamiltonian and Magnetic Systems on principal curved coordinates
- many other concepts and ideas from differential geometry can be explored



Summary

- RNA velocity embeddings into Sasakian Geometry
- Principal curved coordinate analysis
- Hamiltonian and Magnetic Systems on principal curved coordinates
- many other concepts and ideas from differential geometry can be explored



Thank you!