SHAPES AND DYNAMICS OF BIOLOGICAL SYSTEMS

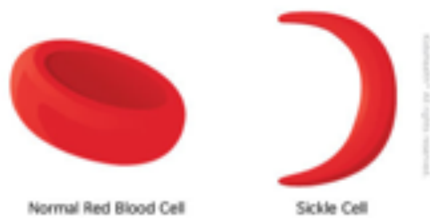
Patrice KOEHL

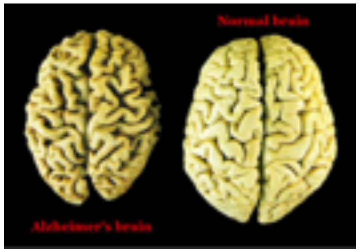
Department of Computer Science
Genome Center

UC Davis

Biology = Quantitative Science ….

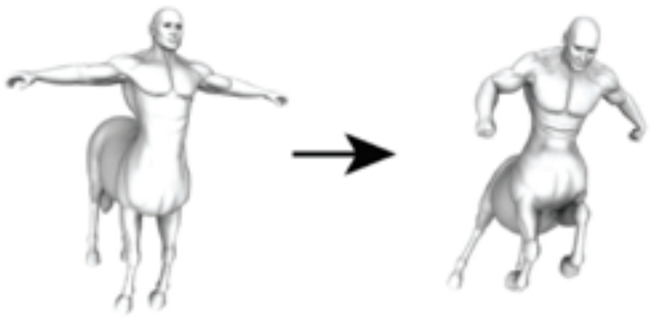
Acetylaminofluorene



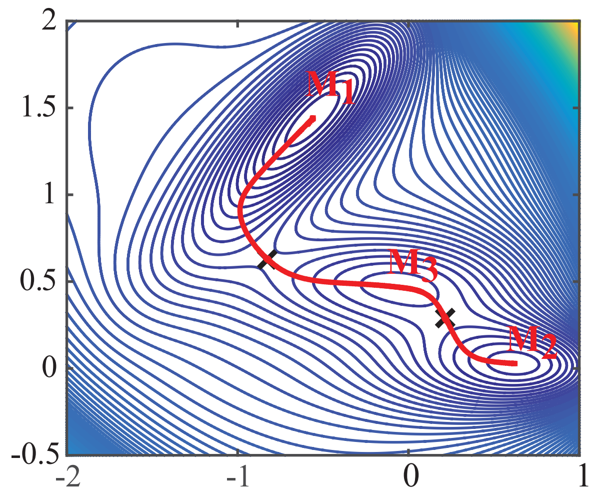


Comparing (biological) shapes

Part I: Optimal diffeomorphism

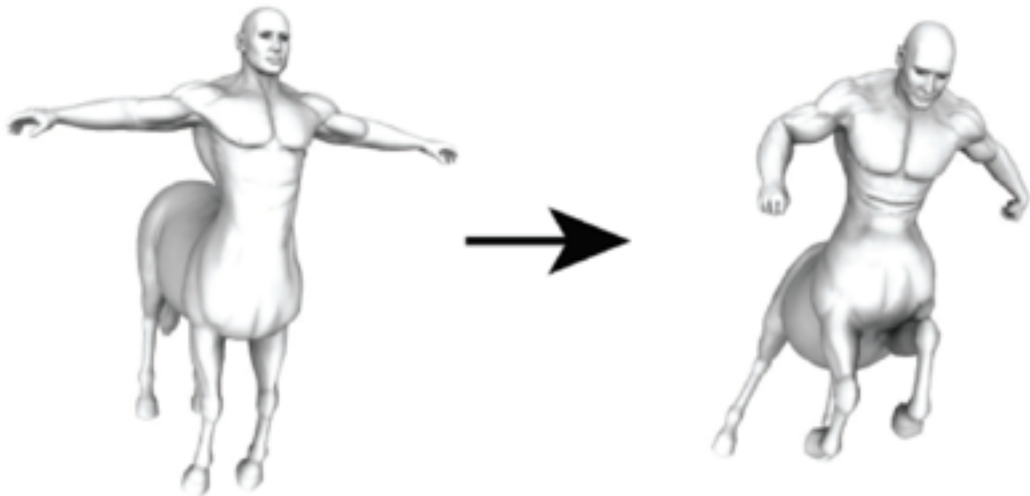


Part II: Geodesics in shape space



Optimal diffeomorphims

We want to compare two surfaces by finding an *optimal diffeomorphism* between them.



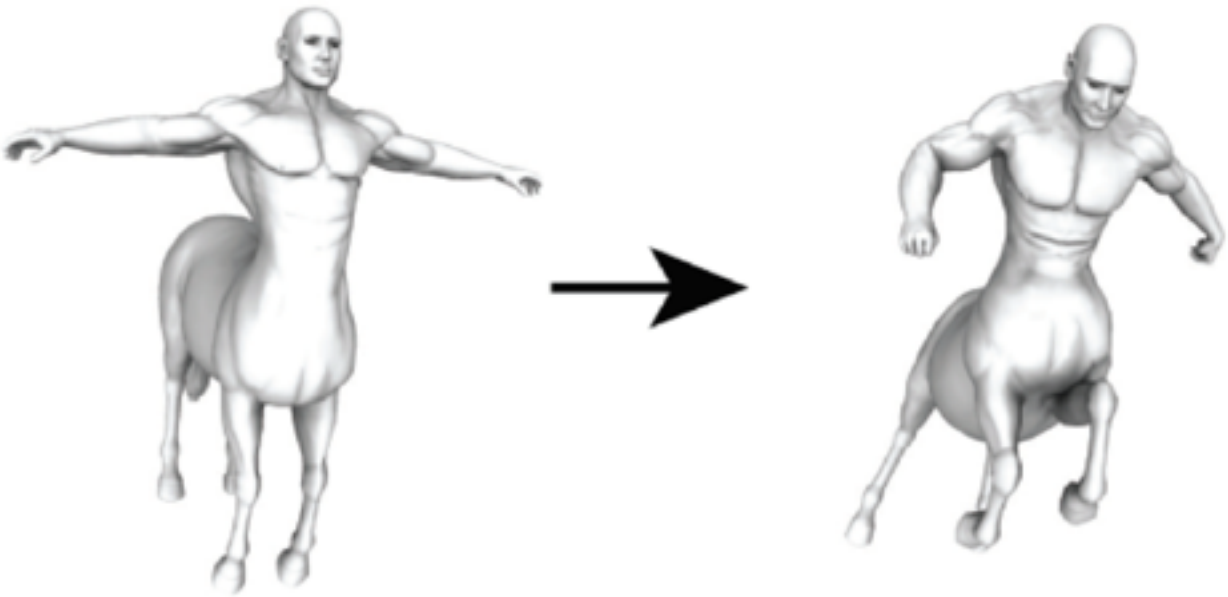
If the surfaces have identical geometry then the optimal diffeomorphism is given by an isometry.

But what if they have different geometries?

What map is *closest* to being an isometry?

Optimal diffeomorphims

Optimal diffeomorphisms do more than give a distance. They also give a *correspondence*.



Diffeomorphic mapping

- General maps between two surfaces deform lengths and angles

- Isometries conserve lengths and angles…. but they are rarely appropriate

- Conformal maps are the next best options, as they distort lengths but preserve angle

The Uniformization Theorem

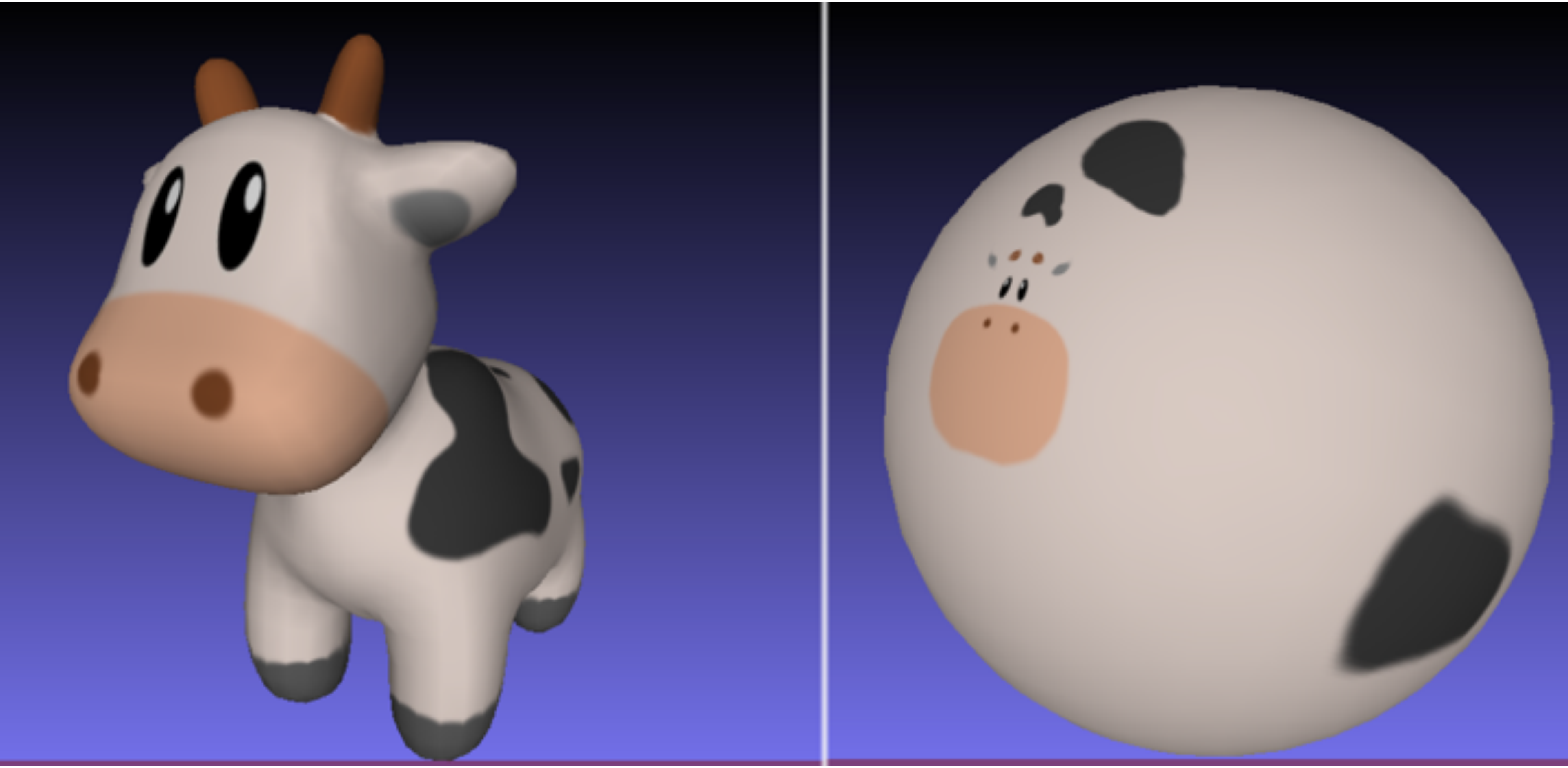
Theorem [Poincaré, Koebe]

Any two genus-zero surfaces are conformally equivalent

Given any two shapes (with no holes), there is a map from one to the other that preserves angles.

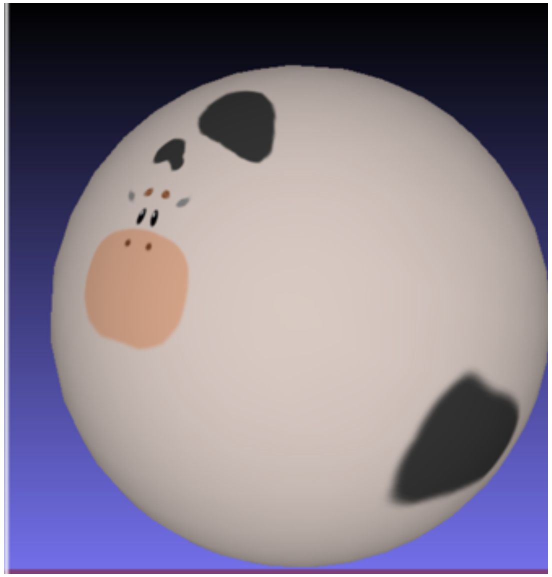
The UC Davis Version…

The UC Davis Version…



Mapping Genus 0 Surfaces to the Sphere

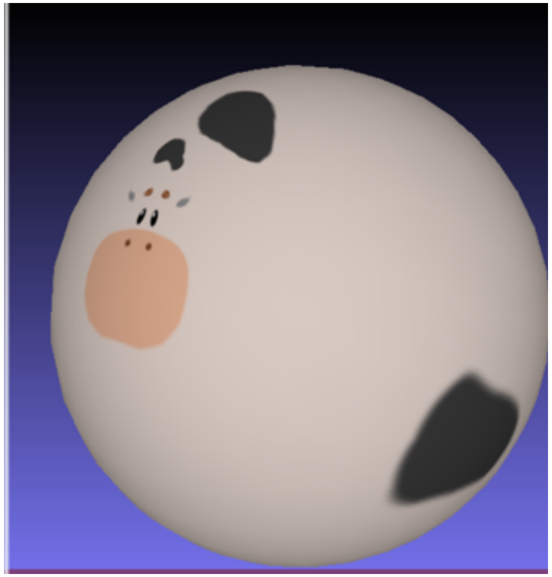
g



$\tilde{g}$

$$\tilde{g} = e^{2u} g$$

Mapping Genus 0 Surfaces to the Sphere



g / g: E -> R+
(i,j) -> l_{ij}

$\tilde{g}$ / g': E -> R+
(i,j) -> l'_{ij}

Continuous:

$$\tilde{g} = e^{2u} g$$

Discrete:

$$l'_{ij} = e^{u(i)+u(j)} l_{ij}$$

Mapping Genus 0 Surfaces to the Sphere

Many algorithms exist:

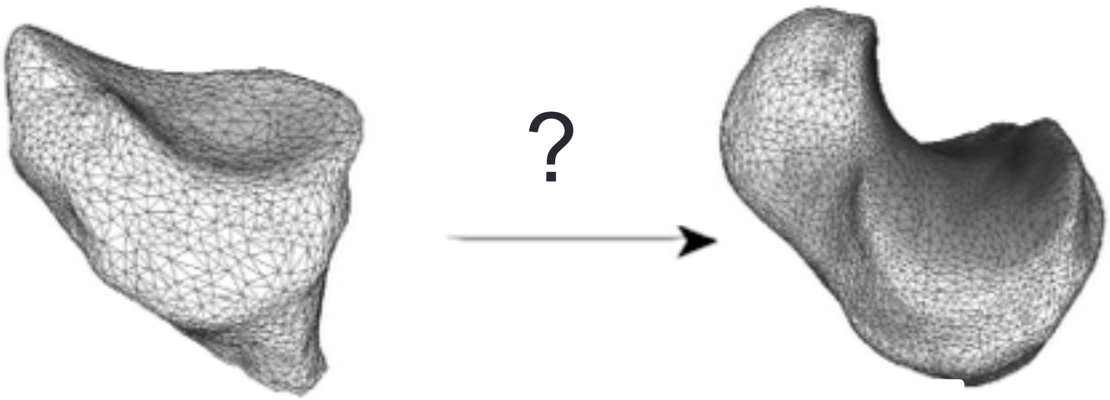
1. Discrete Ricci Flow
2. Discrete Yamabe Flow
3. Conformal Mean Curvature Flow
4. Harmonic Maps
5. Finite Elements
6. Optimize a cost function
7. Discrete Differential Equation
8. Wilmore Flow
9. Circle Packings

Mapping Genus 0 Surfaces to the Sphere

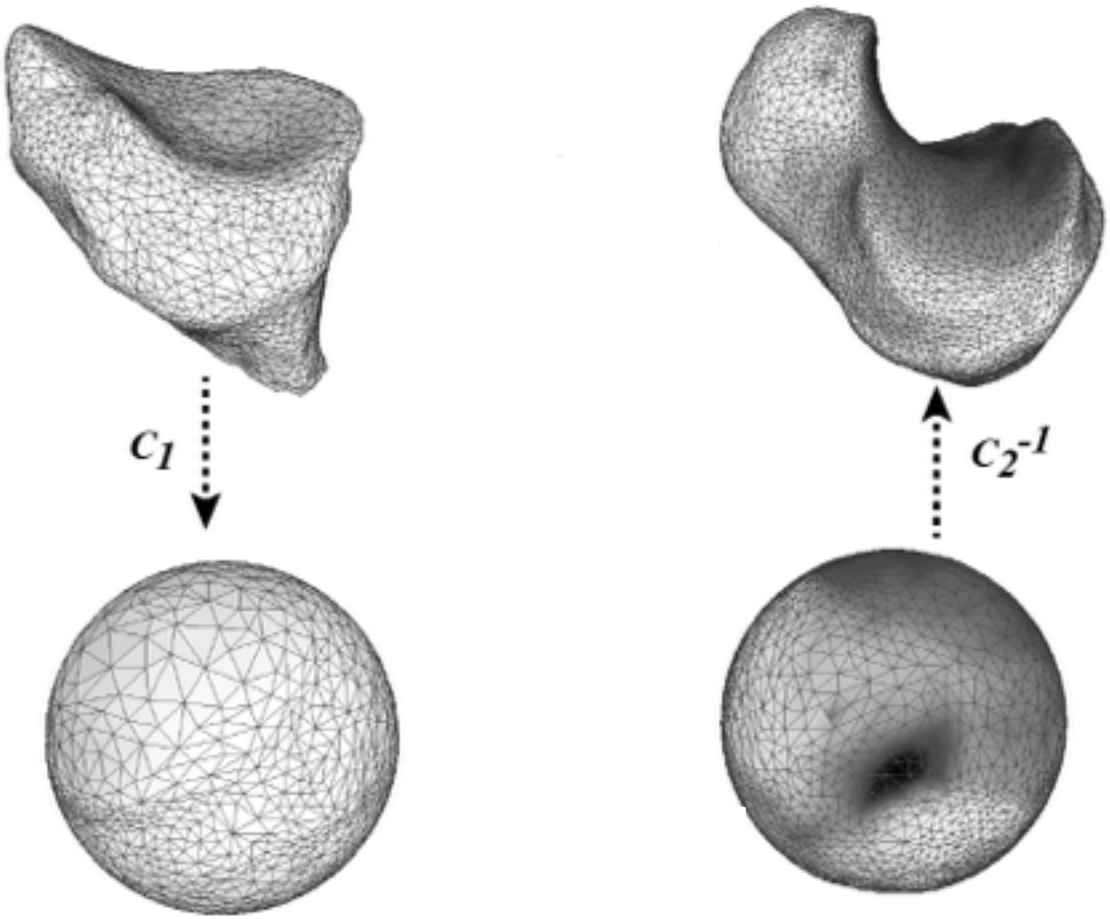
Many algorithms exist:

1. Discrete Ricci Flow
2. Discrete Yamabe Flow
3. Conformal Mean Curvature Flow
4. Harmonic Maps
5. Finite Elements
6. ***Optimize a cost function*** ← ***(Springborn et al, 2008)***
7. Discrete Differential Equation
8. Wilmore Flow
9. Circle Packings

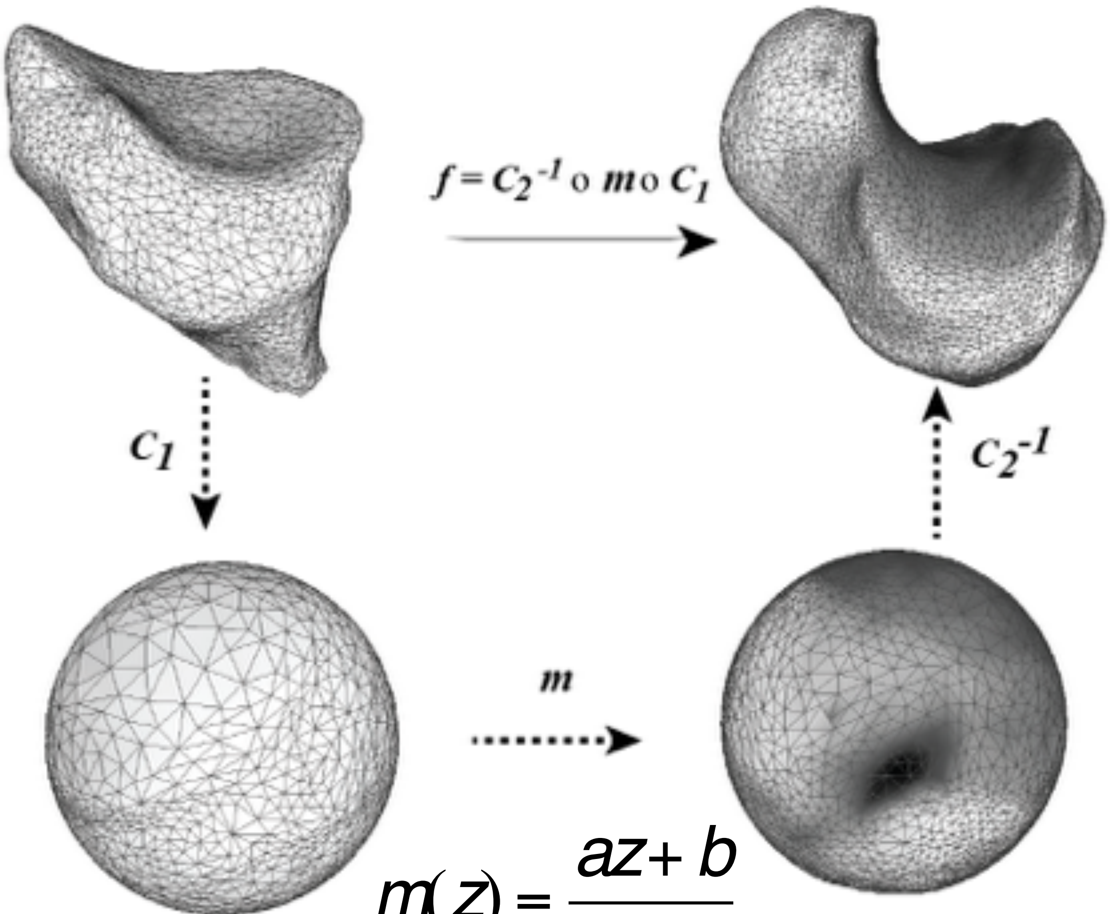
Parametrizing a conformal map between two surfaces



Parametrizing a conformal map between two surfaces

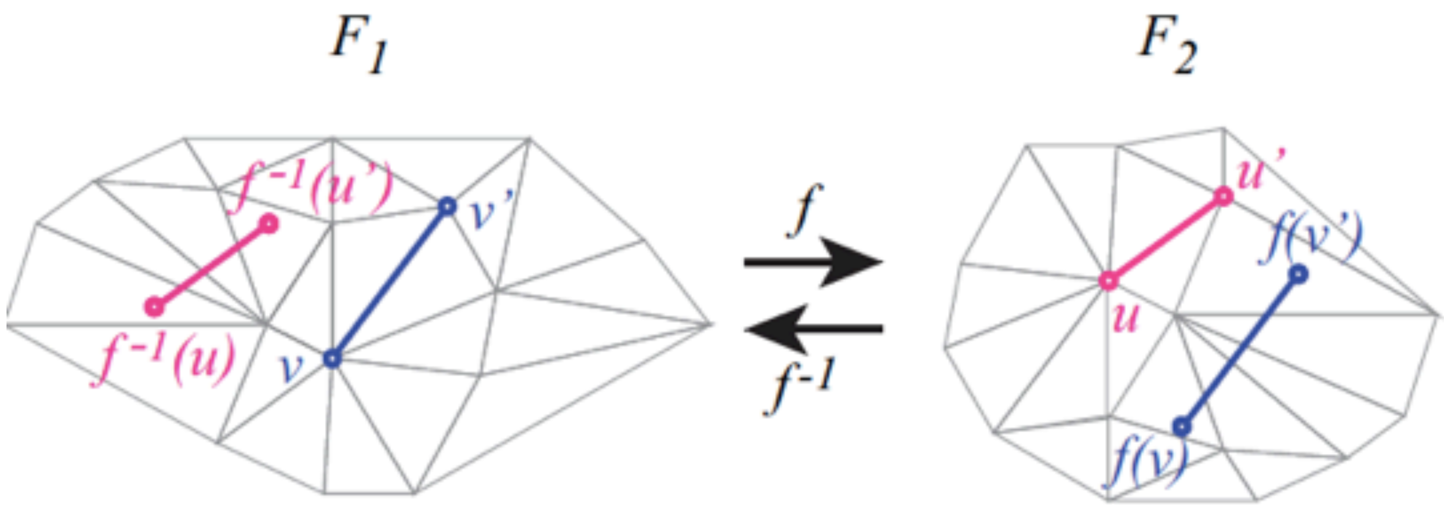


Parametrizing a conformal map between two surfaces



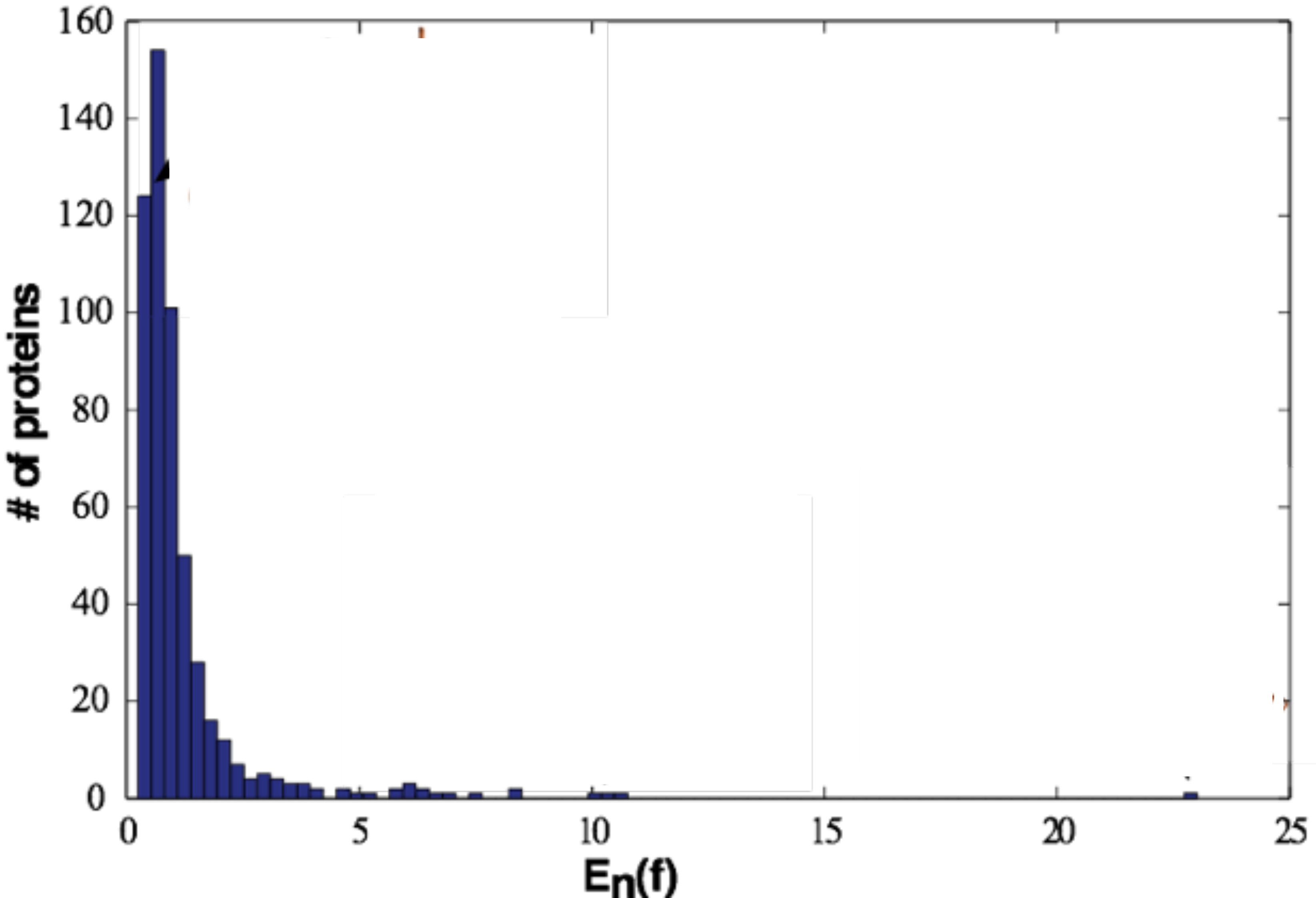
$$m(z) = \frac{az + b}{cz + d}$$

Discrete stretching energy

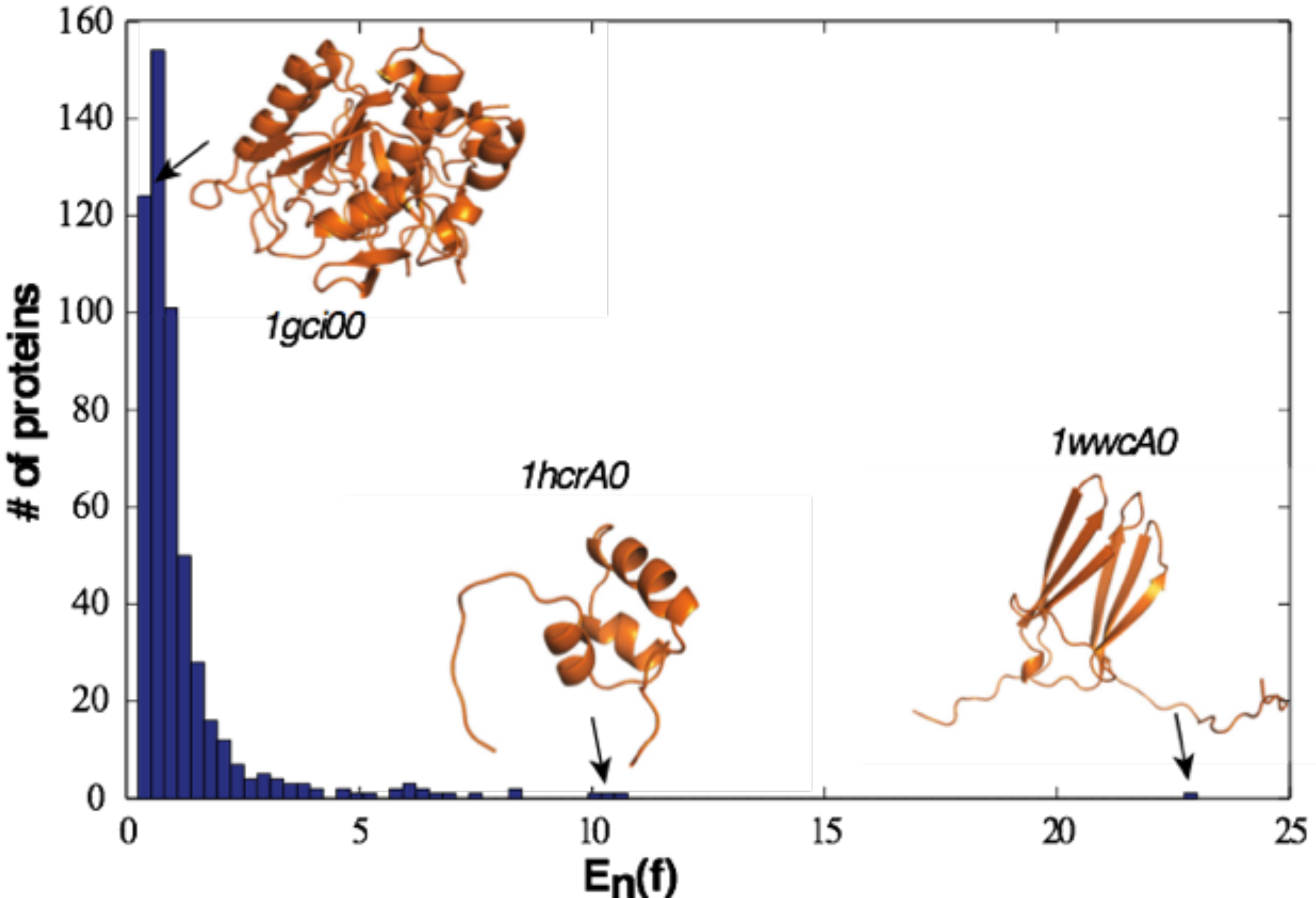


$$E_{sd}(f) = \sqrt{\sum_{(v,v')\in F_1} \frac{A_{vv'}}{3}\left(\frac{l\left(f(v),f(v')\right)}{l(v,v')}-1\right)^2} + \sqrt{\sum_{(u,u')\in F_1} \frac{A_{uu'}}{3}\left(\frac{l\left(f^{-1}(u),f^{-1}(u')\right)}{l(u,u')}-1\right)^2}$$

How round are proteins?



How round are proteins?

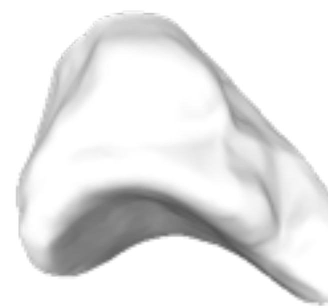


Analysis of anatomical data

Dataset of proximal first metatarsals from 38 prosimian primates, and 23 New and Old World monkeys

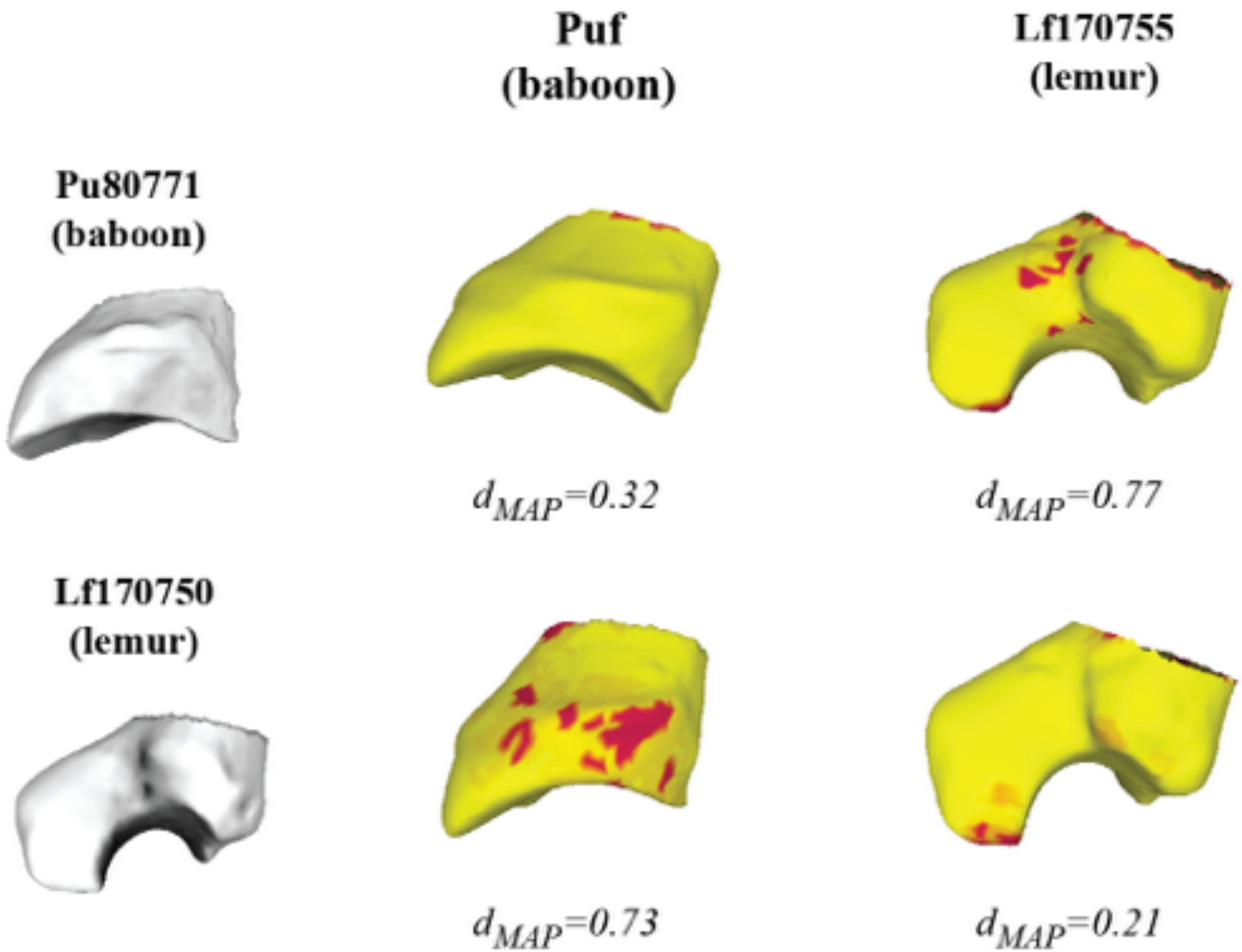
Prosimian:
lemur

Simian:
Cape baboon
(old world)

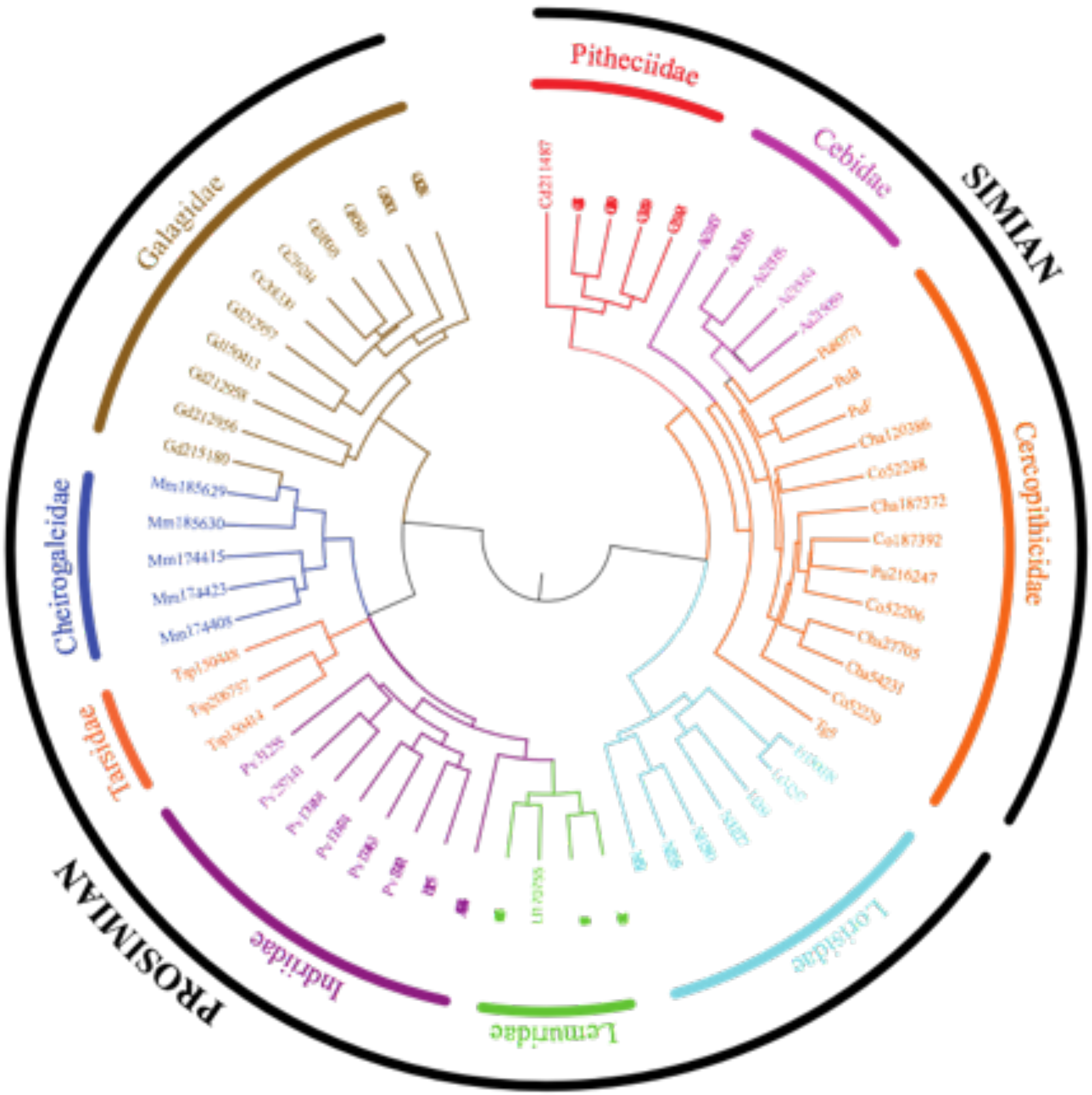


Simian:
White eared titi
(new world)

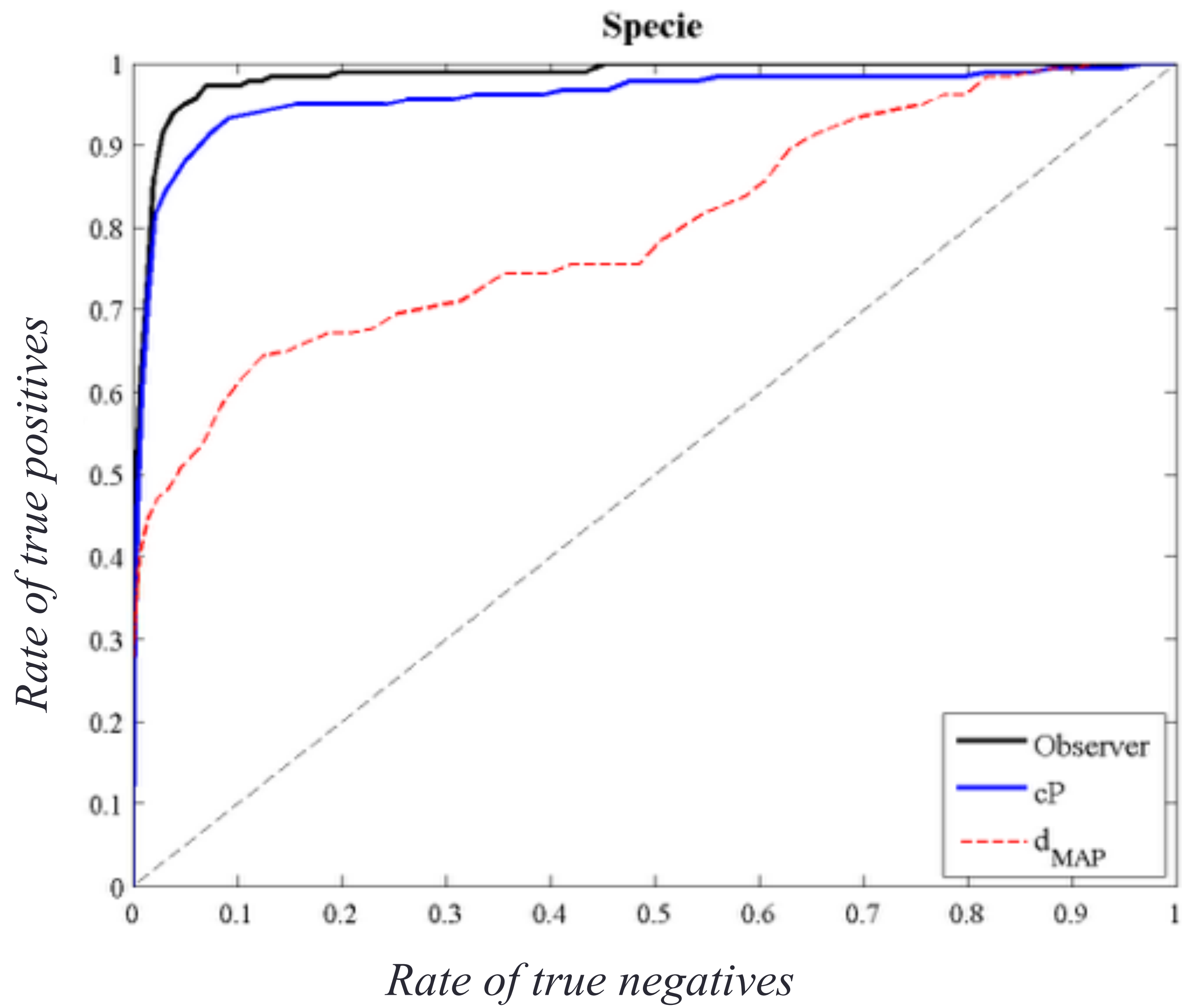
Analysis of anatomical data



Analysis of anatomical data

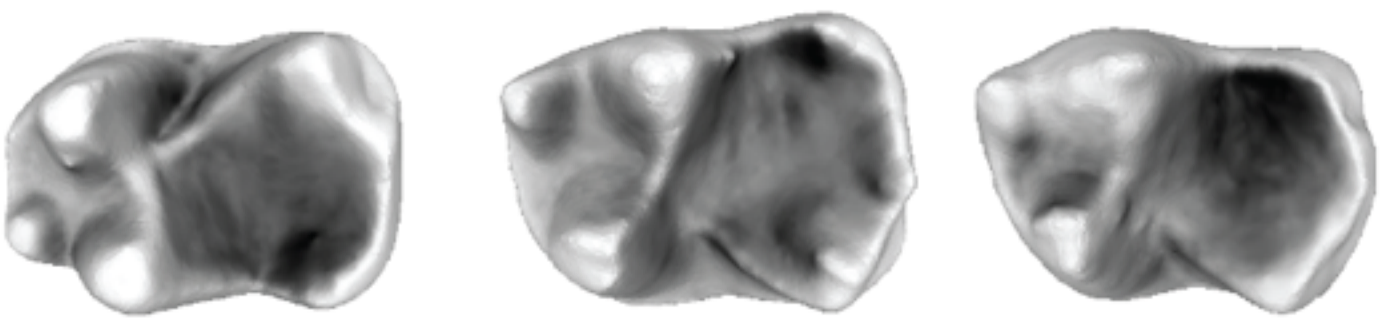


Comparing lower molars from primates



Comparing lower molars from primates:

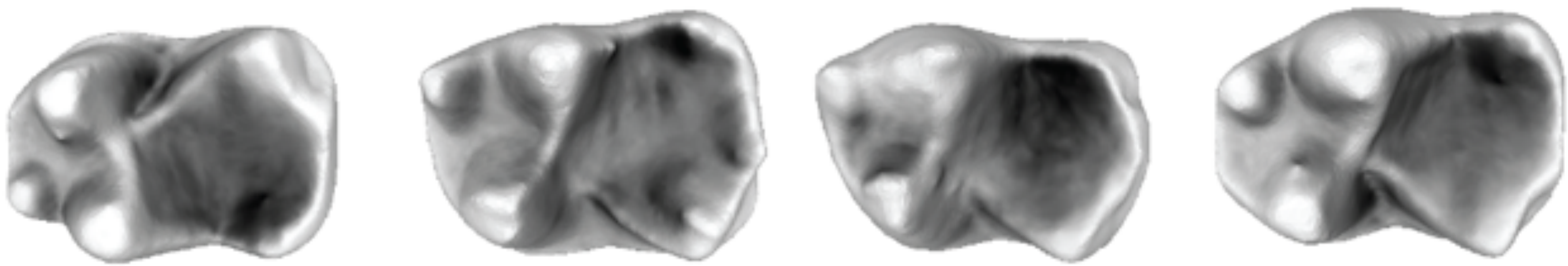
A10-A13: same order, same family
Q06 : same order, different family



	A10	A13	Q06
A10	0.0	**0.38**	**0.30**

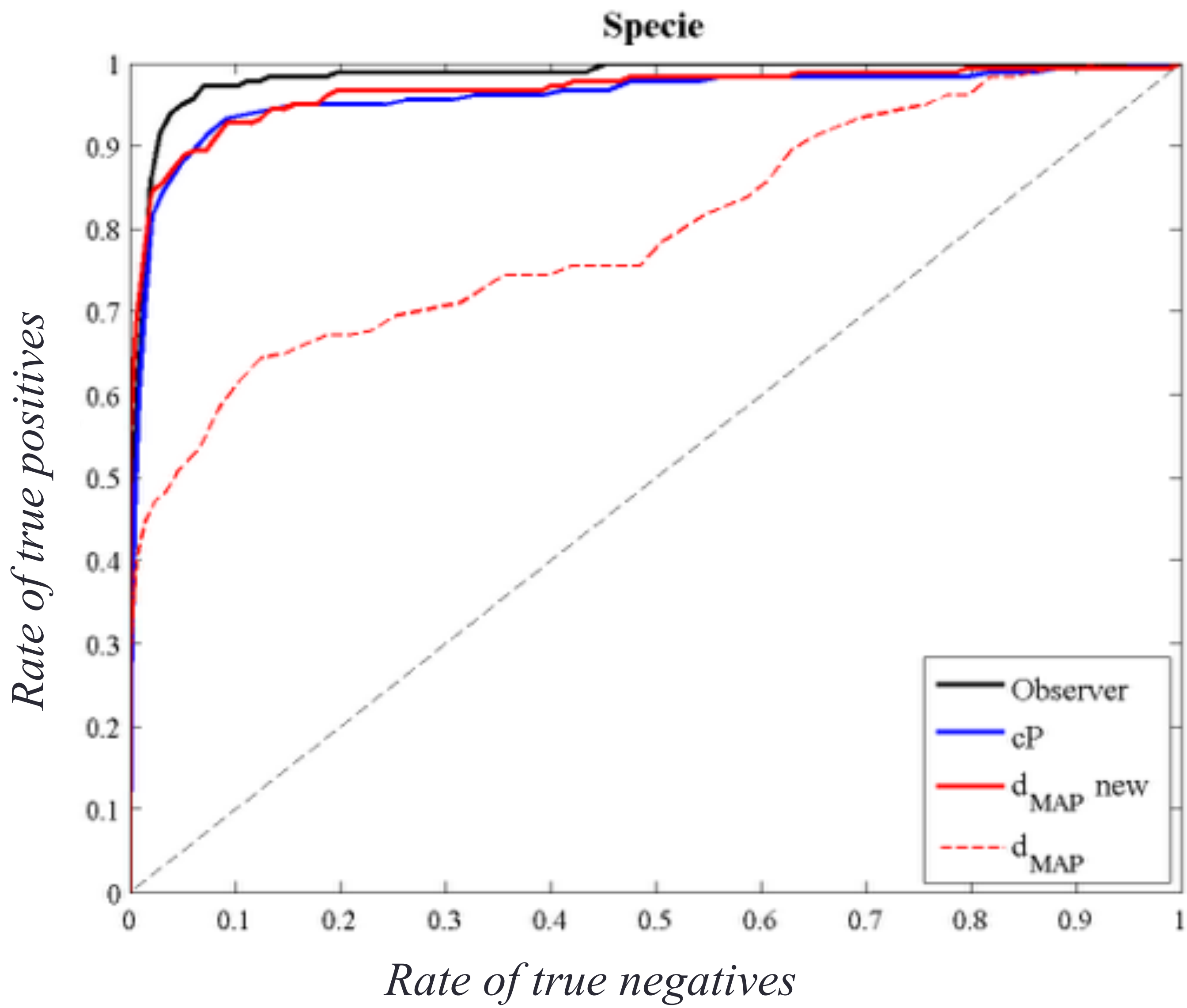
Comparing lower molars from primates:

A10-A13: same order, same family
Q06   : same order, different family



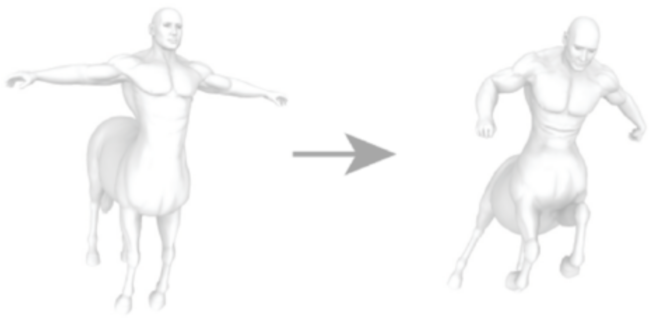
	A10	A13	Q06	A10R
A10	0.0	**0.38**	**0.30**	**0.36**
A10R	**0.36**	**0.18**	**0.22**	**0.0**

Comparing lower molars from primates

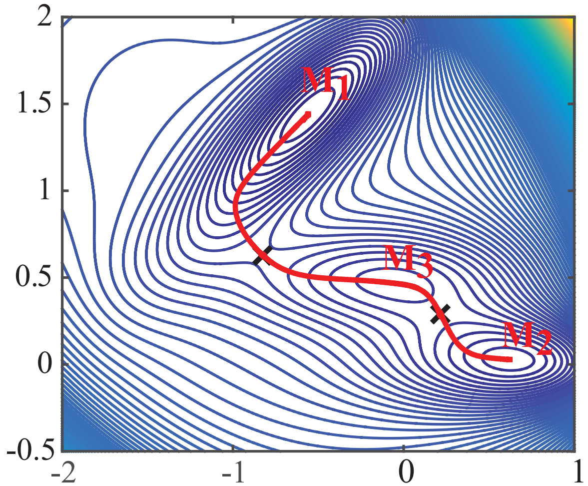


Comparing (biological) shapes

Part I: Optimal diffeomorphism



Part II: Geodesics in shape space



Minimum Action Paths and Shape Similarity

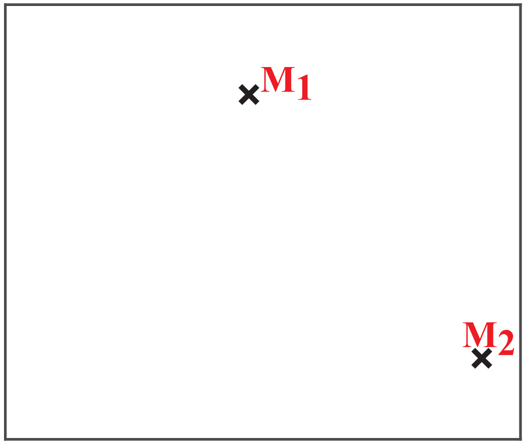
1. Defining a (geodesic) distance between shapes

2. Applications to simple 2D potentials

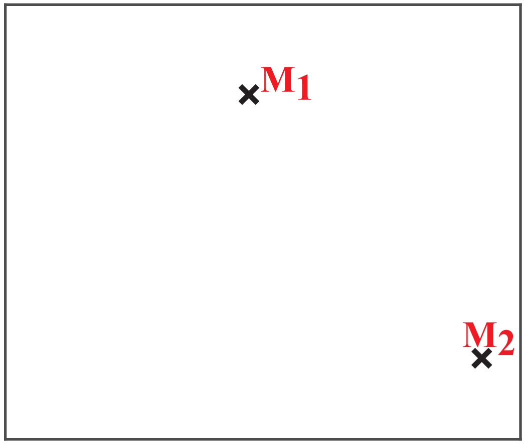
3. Applications to proteins: a simplified potential

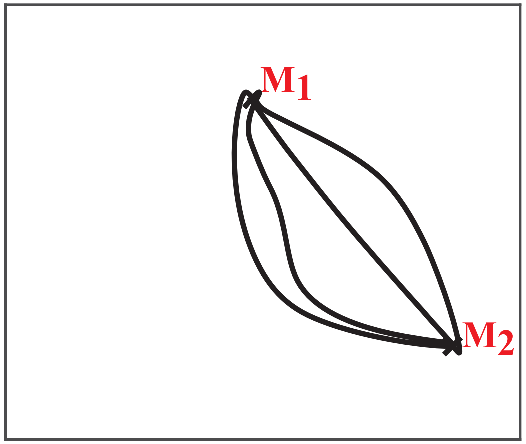
4. Applications to large shapes: more simplifications

Distance between Shapes…

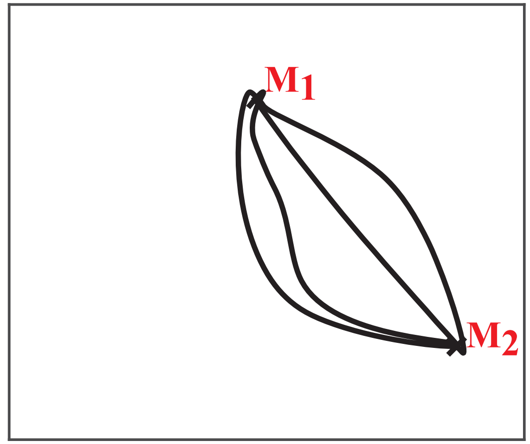


$$D = \sqrt{\left(x_1 - x_2\right)^2 + \left(y_1 - y_2\right)^2}$$





$$D = \sqrt{\left(x_1 - x_2\right)^2 + \left(y_1 - y_2\right)^2}$$

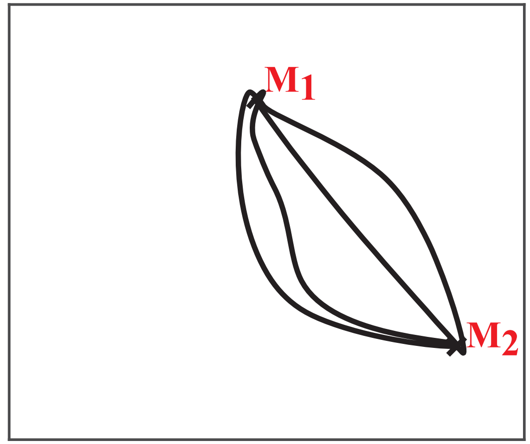


$$L = \sqrt{\left(\frac{dX}{dt}\right)^2}$$

$$\frac{\partial L}{\partial X} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{X}}\right)$$

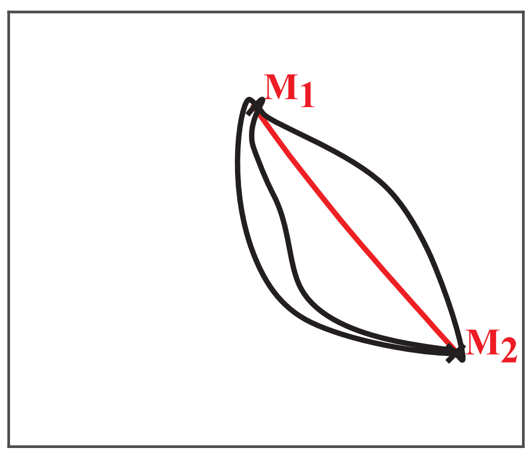
$$D = \sqrt{\left(x_1 - x_2\right)^2 + \left(y_1 - y_2\right)^2}$$

$$D = \sqrt{\left(x_1 - x_2\right)^2 + \left(y_1 - y_2\right)^2}$$



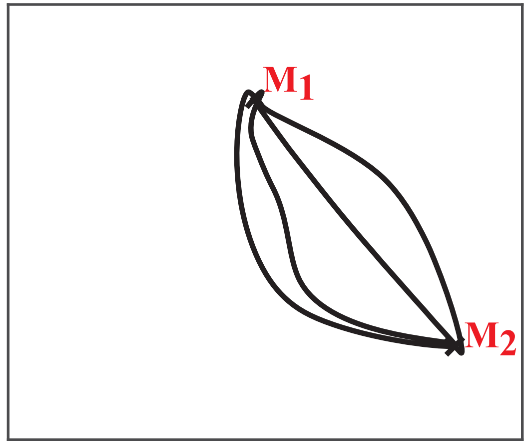
$$L = \sqrt{\left(\frac{dX}{dt}\right)^2}$$

$$\frac{\partial L}{\partial X} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{X}}\right)$$



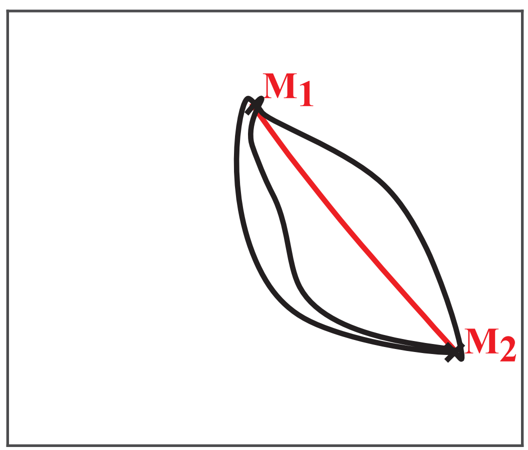
$$S = \int_0^T L\,dt = \sqrt{\left(x_1 - x_2\right)^2 + \left(y_1 - y_2\right)^2}$$

$$D = \sqrt{\left(x_1 - x_2\right)^2 + \left(y_1 - y_2\right)^2}$$

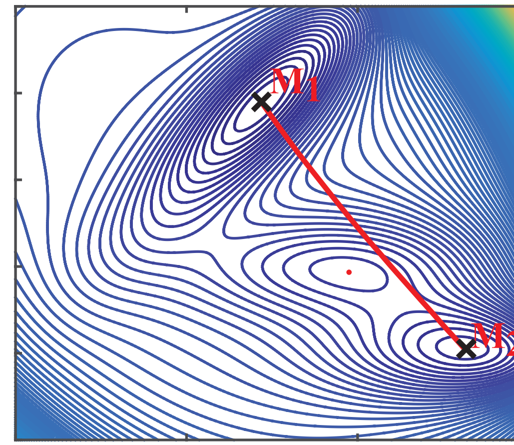


$$L = \sqrt{\left(\frac{dX}{dt}\right)^2}$$

$$\frac{\partial L}{\partial X} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{X}}\right)$$



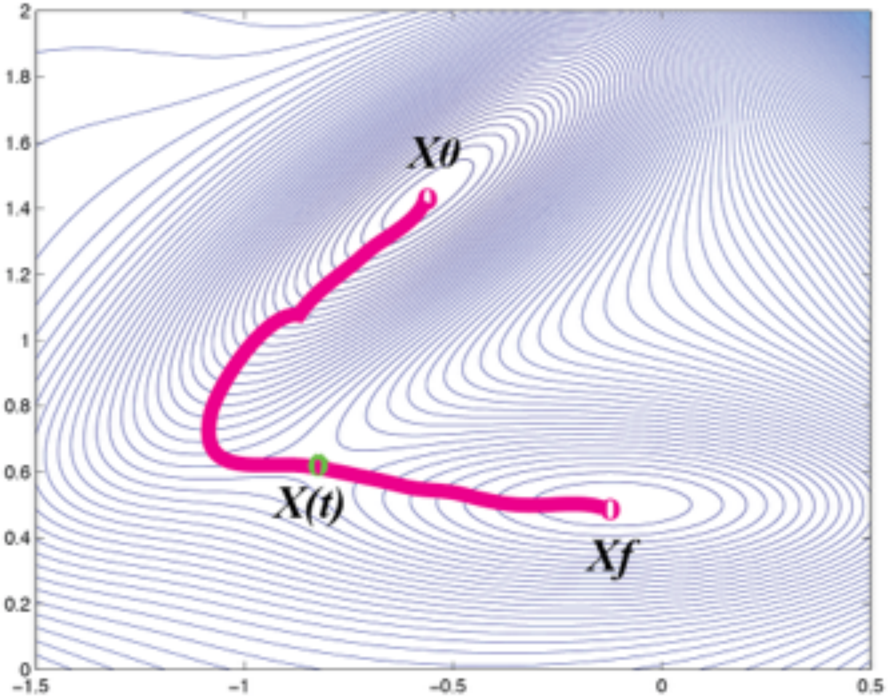
$$S = \int_0^T L\,dt = \sqrt{\left(x_1 - x_2\right)^2 + \left(y_1 - y_2\right)^2}$$



Diffusive form of Langevin Equation:

$$\eta M \frac{dX}{dt} = -\nabla U(X) + B$$

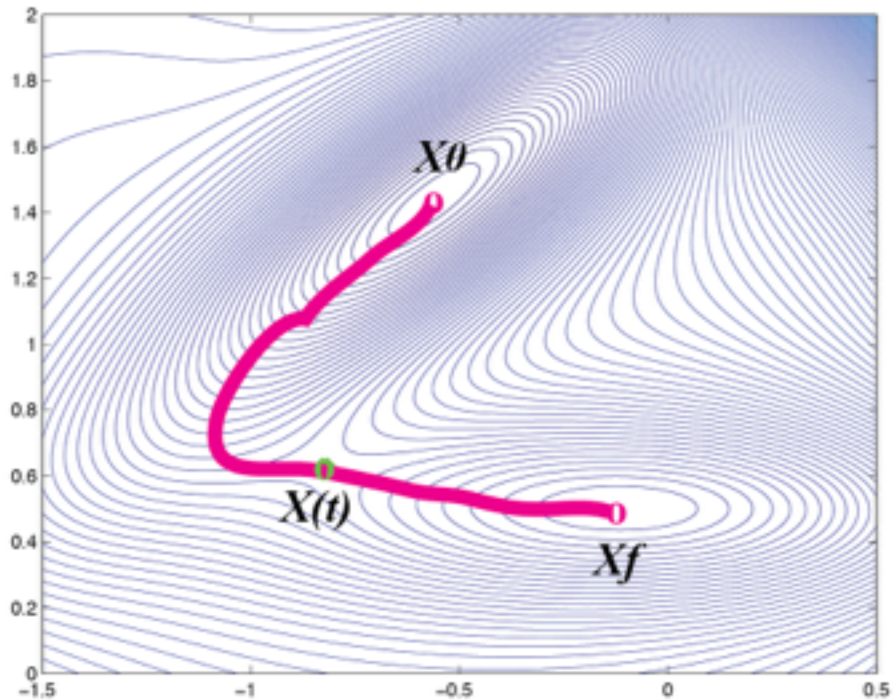
η: *friction*
M: diagonal mass matrix
U: potential energy
B: random force



For a trajectory

$$X(0) \rightarrow X(t) \rightarrow X_f(T_f)$$

$$P\left(X(0) \rightarrow X_f(T_f)\right) \propto \exp\left(-\frac{S}{k_B T}\right)$$



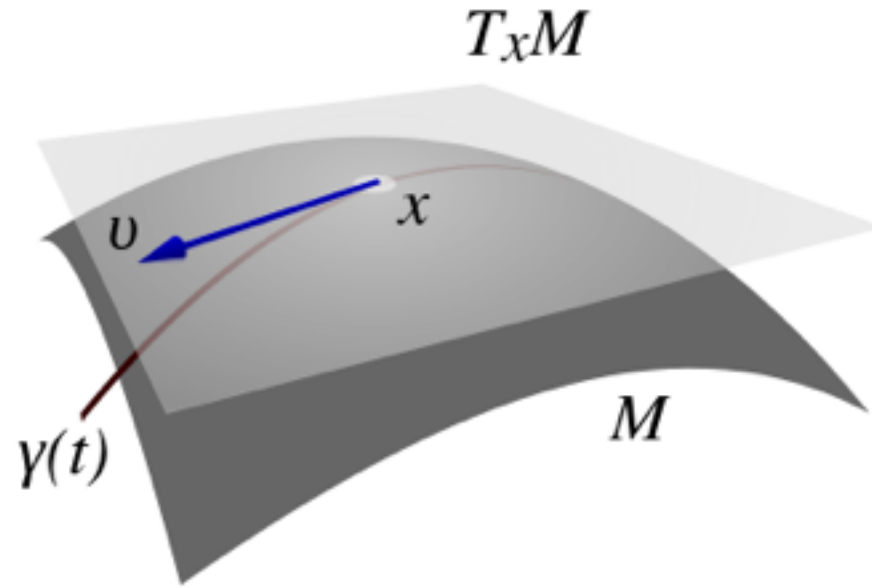
where the action, S, is given by (Onsager and Machlup, 1953):

$$S = \frac{1}{2\eta}\int_0^{Tf}\left(\eta M \frac{dX}{dt} + \nabla U(X)\right)^2 dt$$

Corresponding Lagrangian: $$L = \left(\eta M \frac{dX}{dt} + \nabla U(X)\right)^2$$

Let M be a (smooth) manifold and

E a function from (M,TM) то [0,∞);



Let X_0 and X_F be two points on M. Then

$$d(X,Y) = \inf\left\{ \int_0^T E(\gamma(t),\dot{\gamma}(t))\,dt \;\middle|\; \gamma \in C^1([0,T],M), \gamma(0) = X_0, \gamma(T) = X_F \right\}$$

defines an intrinsic quasi-metric on M.

There always exist length minimizing curves on (M,E). Such curves can always be reparametrized to be geodesics, and any geodesic must satisfy the Euler-Lagrange equation for F(γ):

$$F(\gamma)=\int_{X_0}^{X_F} L(\gamma(t),\dot{\gamma}(t))dt=\int_{X_0}^{X_F}\left[E(\gamma(t),\dot{\gamma}(t))\right]^2 dt$$

There always exist length minimizing curves on (M,E). Such curves can always be reparametrized to be geodesics, and any geodesic must satisfy the Euler-Lagrange equation for F(γ):

$$F(\gamma) = \int_{X_0}^{X_F} L(\gamma(t), \dot{\gamma}(t))dt = \int_{X_0}^{X_F} \left[E(\gamma(t), \dot{\gamma}(t)) \right]^2 dt$$

Onsager and Machlup (1953) action:

$$S = \frac{1}{2\eta} \int_0^{Tf} \left(\eta M \frac{dX}{dt} + \nabla U(X) \right)^2 dt$$

Corresponding Lagrangian:

$$L = \left(\eta M \frac{dX}{dt} + \nabla U(X) \right)^2$$

Lagrangian:

$$L = \left(\eta M \frac{dX}{dt} + \nabla U(X) \right)^2$$

Euler Lagrange equations:

$$\frac{\partial L}{\partial X} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{X}} \right)$$

$$\frac{d^2 X}{dt^2} = \nabla\nabla U(X) \nabla U(X)$$

Boundary conditions:

$$X(0) = X_0 \qquad X(T_f) = X_f$$

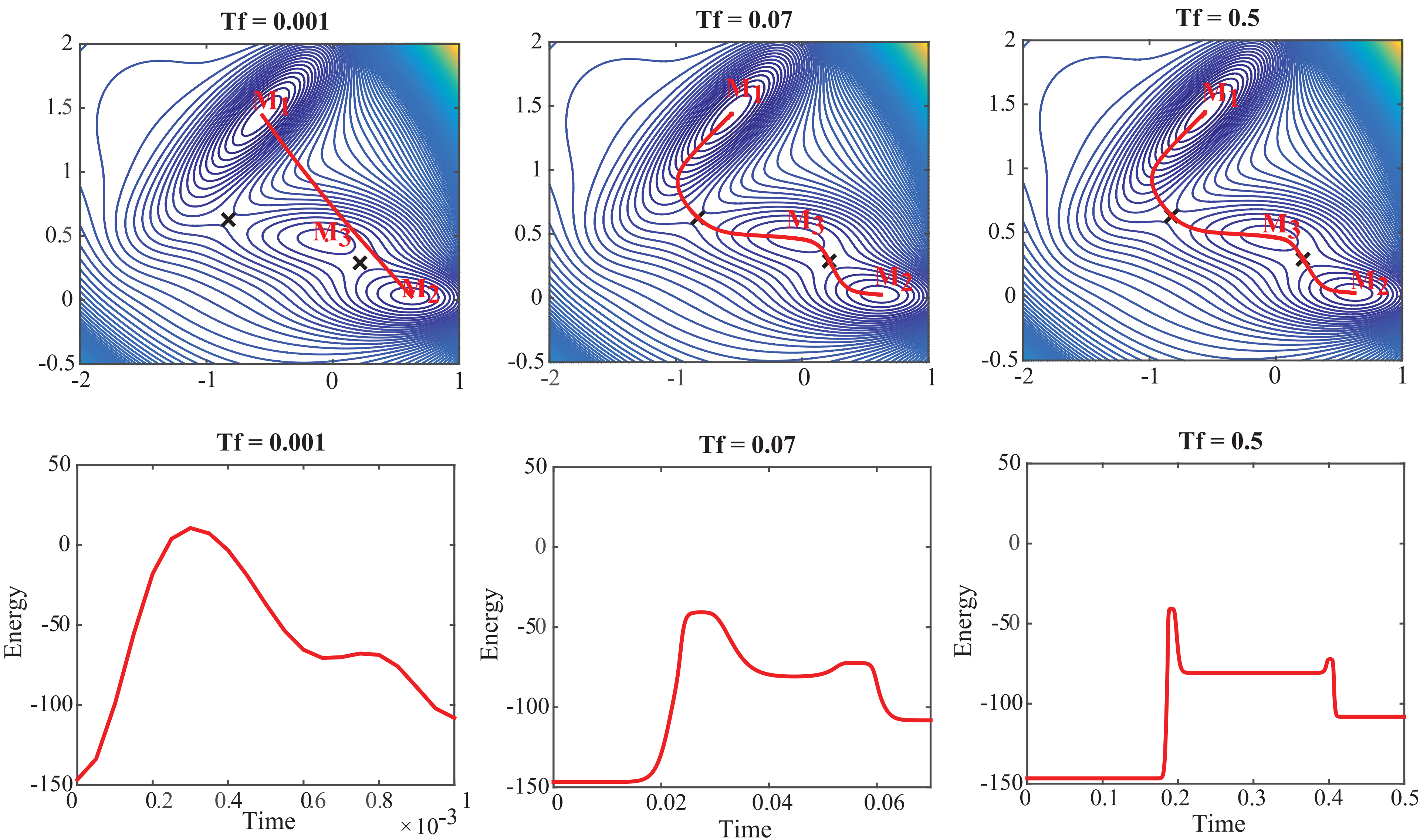
Minimum Action Paths and Shape Similarity

1. Defining a (geodesic) distance between shapes

2. Applications to simple 2D potentials

3. Applications to proteins: a simplified potential

4. Applications to large shapes: more simplifications



Tf = 0.001
Tf = 0.07
Tf = 0.5
M1
M2
M3
Energy
Time
×10-3

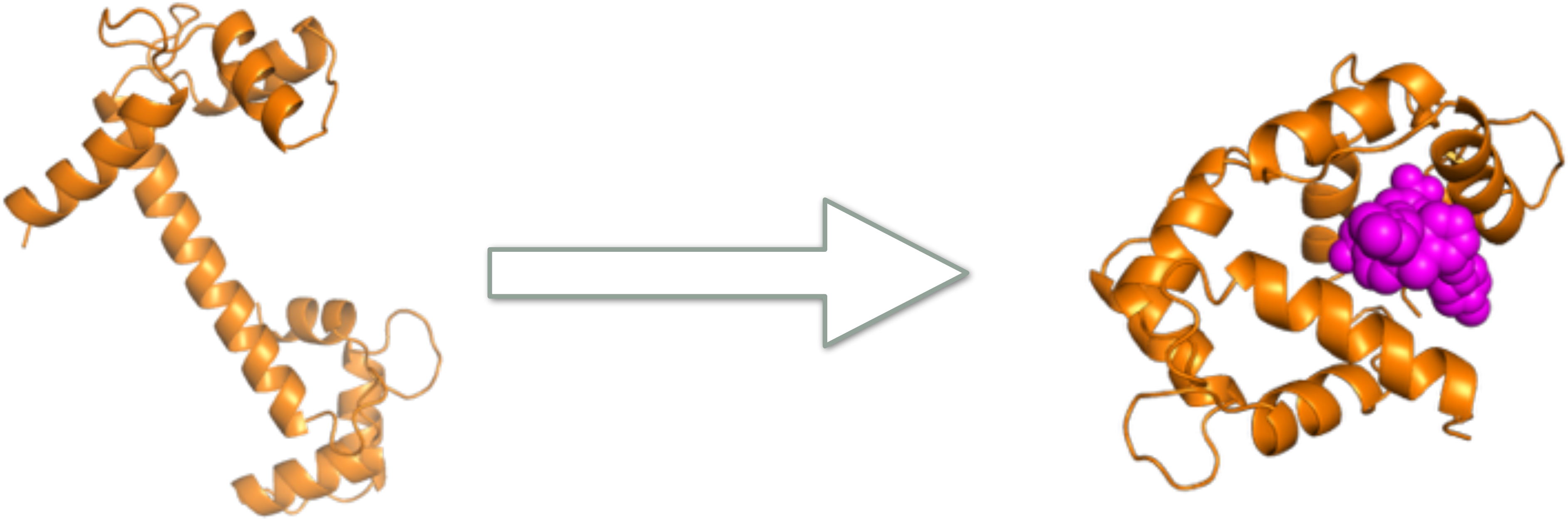
Minimum Action Paths and Shape Similarity

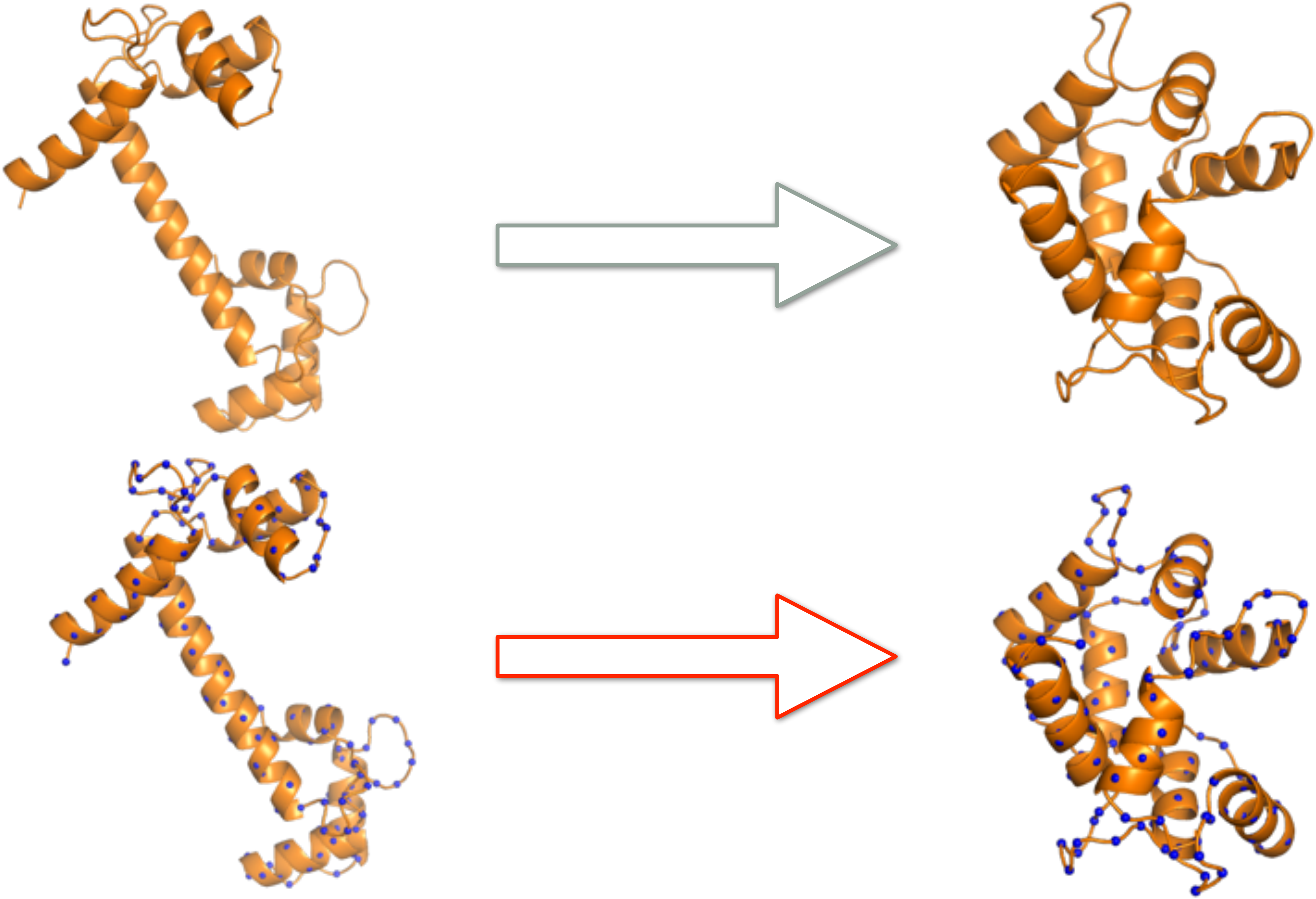
1. Defining a (geodesic) distance between shapes

2. Applications to simple 2D potentials

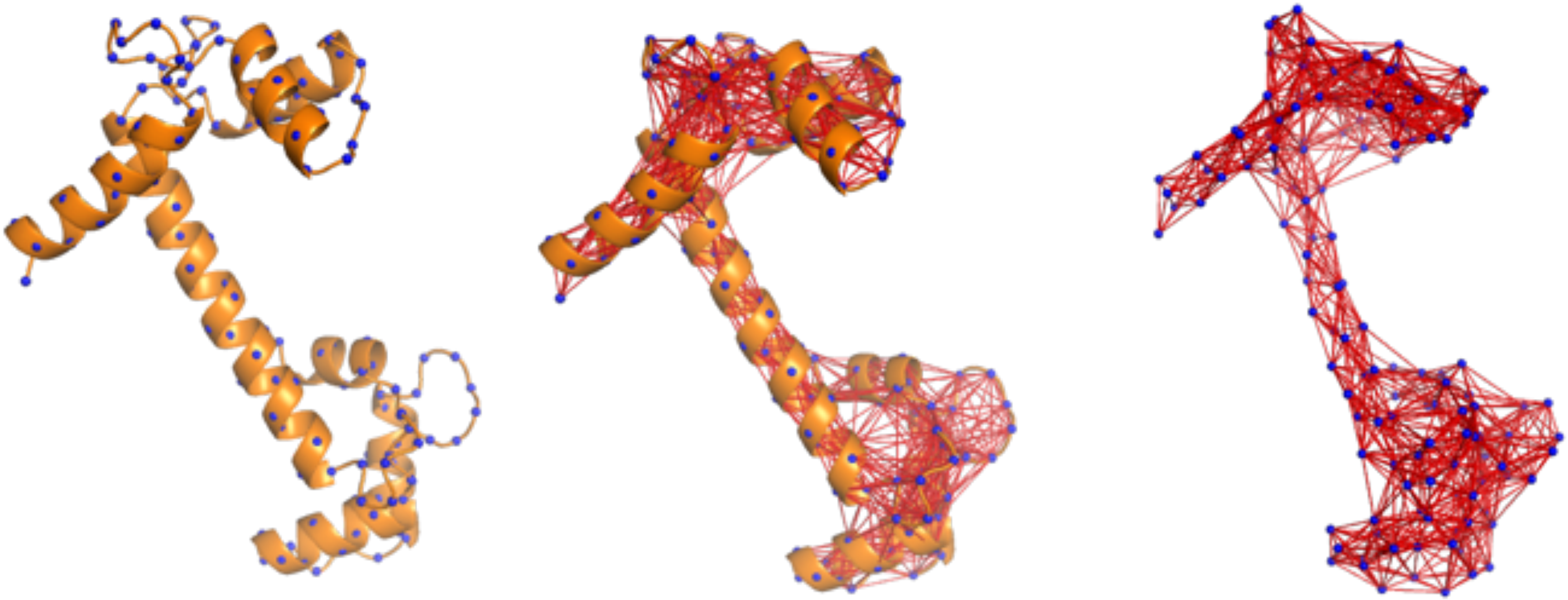
3. Applications to proteins: a simplified potential

4. Applications to large shapes: more simplifications





Elastic network for biomolecules:



Elastic network for biomolecules:

Elastic potential:

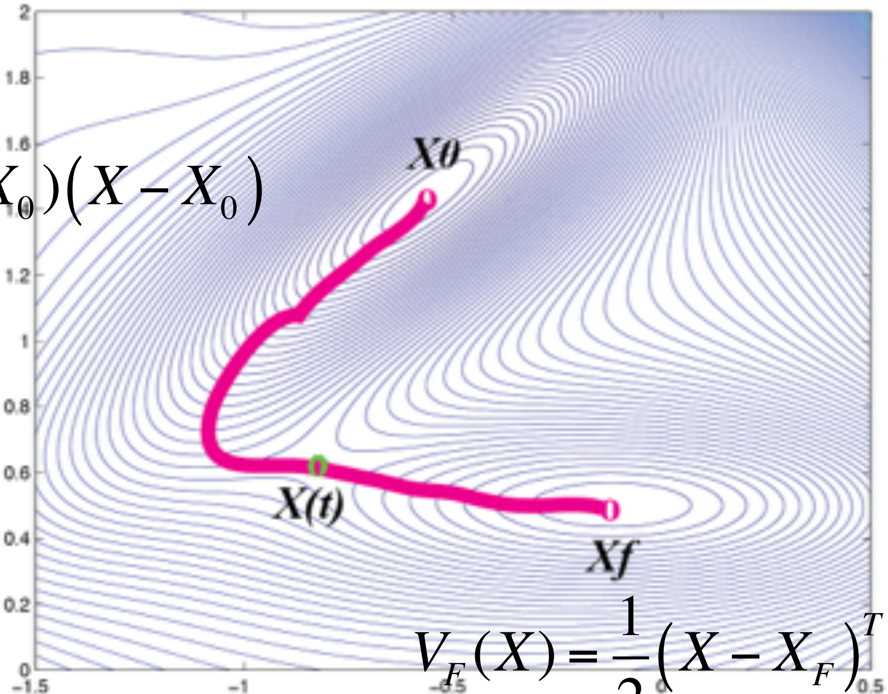
$$V_{ENM}(X) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j>i} k_{ij}(r_{ij} - r_{ij}^{0})^2$$

2nd order Taylor expansion:

$$V_{ENM}(X) \approx V_{ENM}(X^0) + \nabla V_{ENM}(X^0)^T(X - X^0) + \frac{1}{2}(X - X^0)^T H(X - X^0)$$

$$V_{ENM}(X) \approx \frac{1}{2}(X - X^0)^T H(X - X^0)$$

Mixing potential for transition path:



$$V_0(X) = \frac{1}{2}\left(X - X_0\right)^T H(X_0)\left(X - X_0\right)$$

$$V_F(X) = \frac{1}{2}\left(X - X_F\right)^T H(X_F)\left(X - X_F\right)$$

Mixing potential:

$$U(X) = -\log\left(e^{-V_0(X)} + e^{-V_F(X)}\right)$$

Solve:

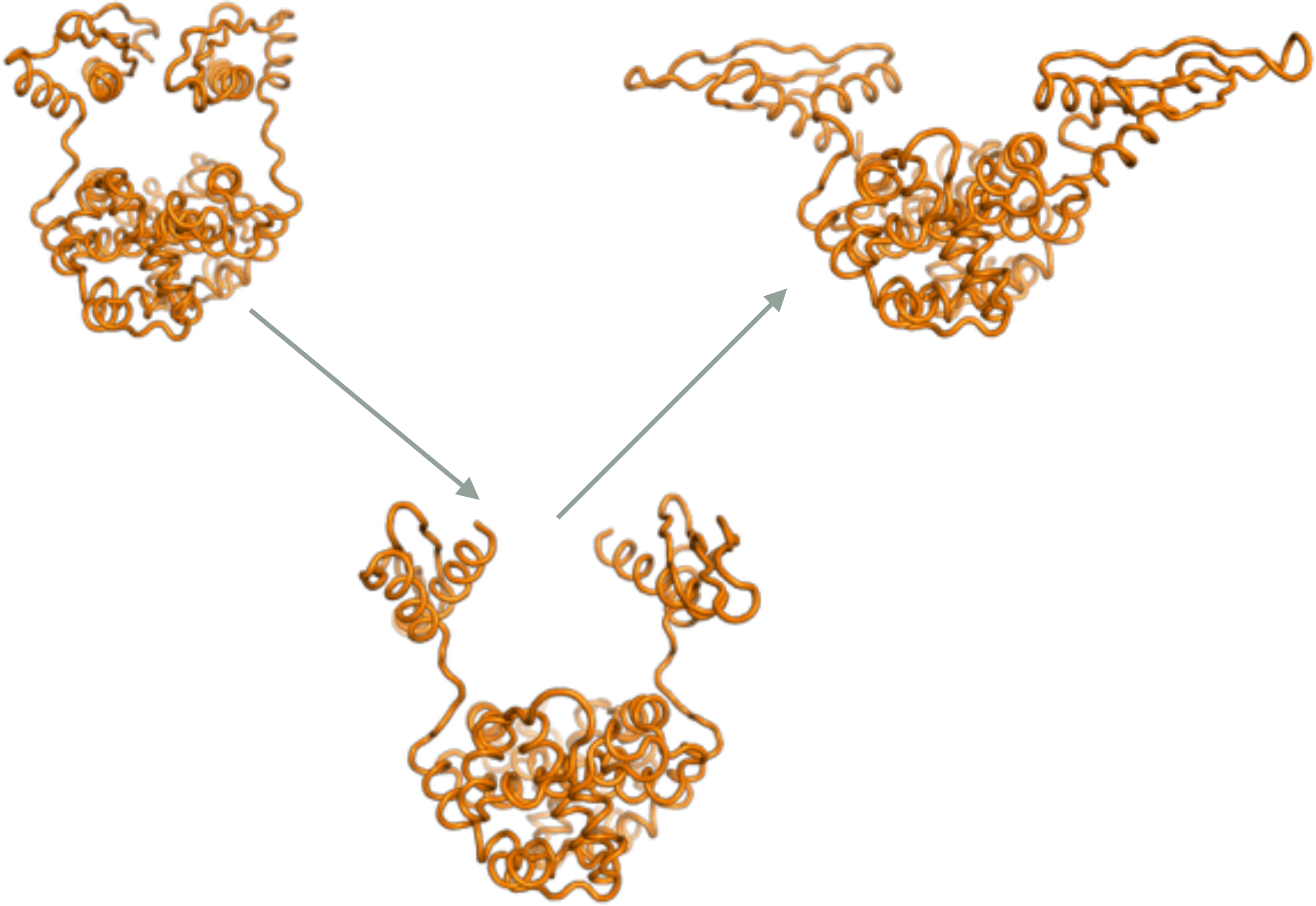
$$\frac{d^2X}{dt^2} = \nabla\nabla U(X)\nabla U(X)$$

with boundary conditions:

$$X(0) = X_0 \qquad X(T_f) = X_f$$

using a relaxation method.

Transition Path for a Ribonuclease III



Minimum Action Paths and Shape Similarity

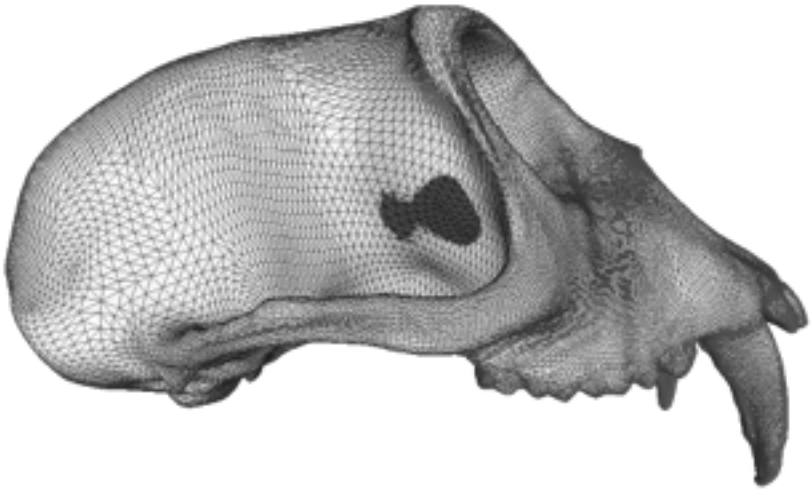
1. Defining a (geodesic) distance between shapes

2. Applications to simple 2D potentials

3. Applications to proteins: a simplified potential

4. Applications to large shapes: more simplifications

An elastic model for shapes

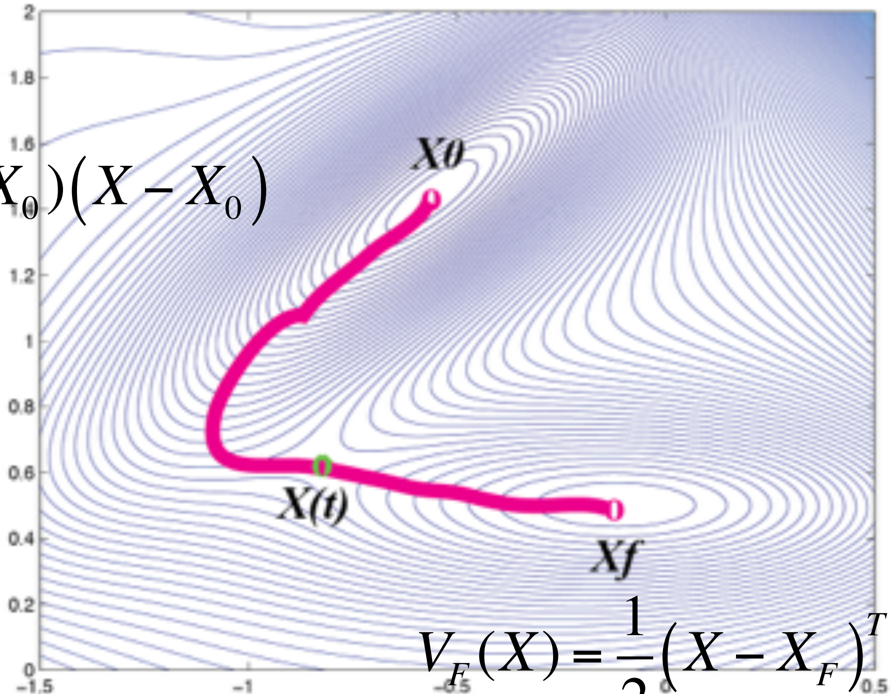


$$U_1(X) = \sum_{edges(i,j)} \left(d_{ij}(X) - d_{ij}(X_0) \right)^2$$

$$\begin{aligned} U(X) &\approx E(X_0) + \nabla E(X_0)^T + \frac{1}{2}\left(X(t) - X_0\right)^T H(X_0)\left(X(t) - X_0\right) \\ &= \frac{1}{2}\left(X(t) - X_0\right)^T H(X_0)\left(X(t) - X_0\right) \end{aligned}$$

Mixing potential for transition path:

$$V_0(X) = \frac{1}{2}\left(X - X_0\right)^T H(X_0)\left(X - X_0\right)$$



$$V_F(X) = \frac{1}{2}\left(X - X_F\right)^T H(X_F)\left(X - X_F\right)$$

Mixing potential:

$$U(X) = \min\left(V_0(X), V_1(X)\right)$$

Euler-Lagrange equations for stationary action:

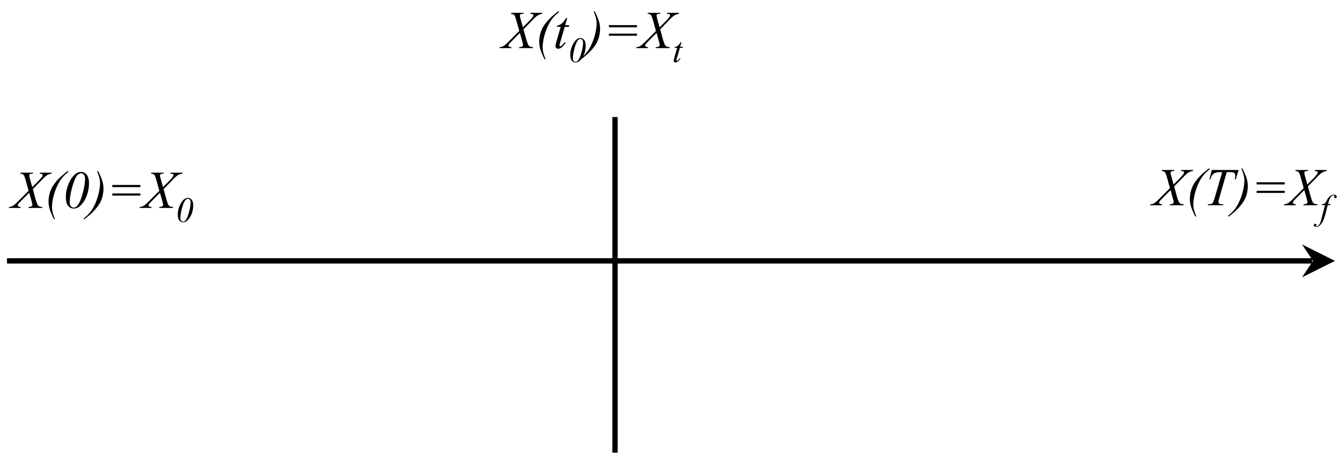
$$\frac{\partial L}{\partial X} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{X}}\right) \Rightarrow \begin{cases} \ddot{X} = H(X_0)^T H(X_0)(X - X_0) & t \leq t_0 \\ \ddot{X} = P(X_f)^T P(X_f)(X - X_f) & t \geq t_0 \end{cases}$$

$$X(t \rightarrow t_0) = X_t = X(t_0 \leftarrow t)$$

$$\dot{X}(t \rightarrow t_0) = \dot{X}(t_0 \leftarrow t)$$

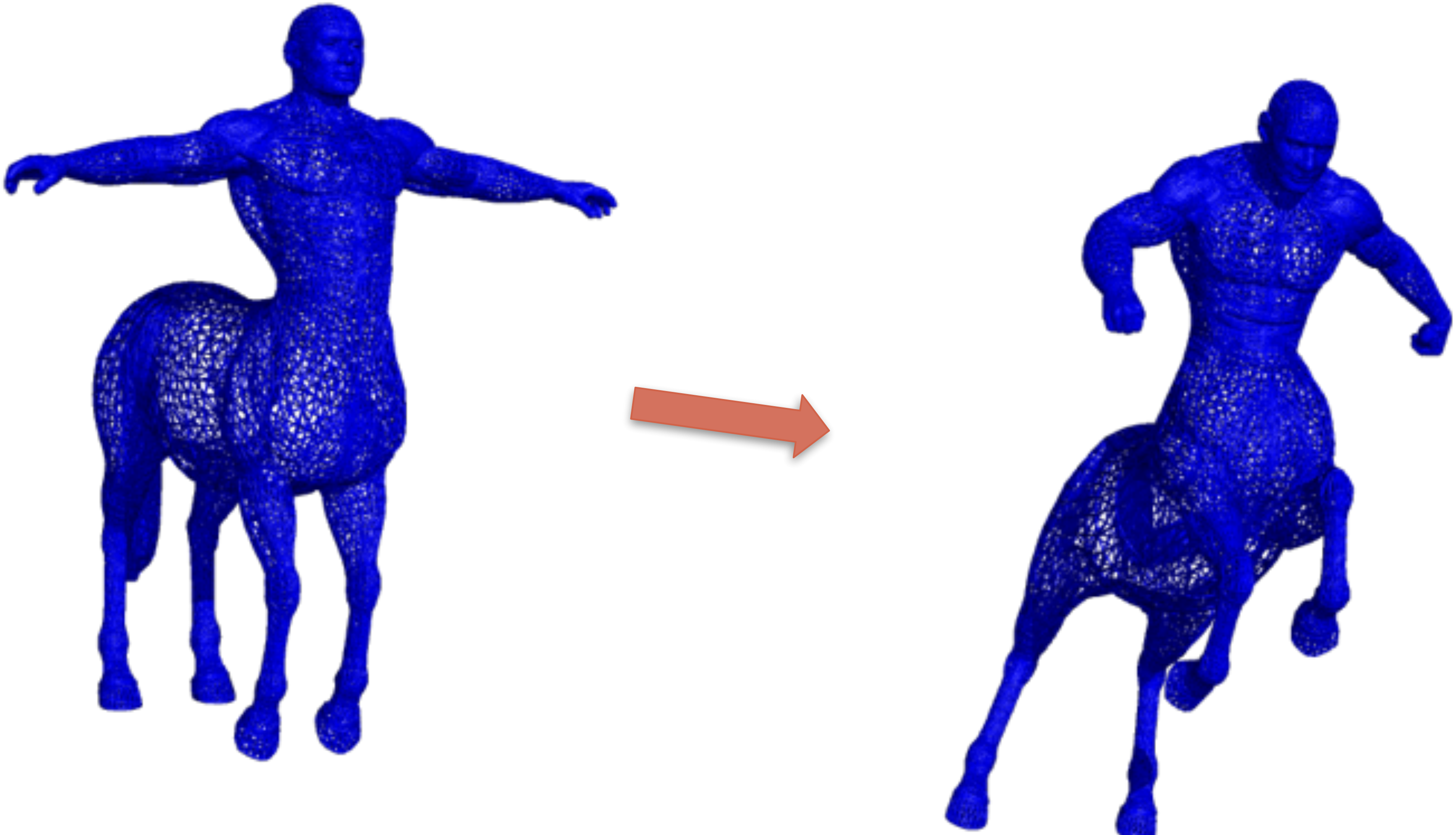
$$U(X(t \rightarrow t_0)) = U(X(t_0 \leftarrow t))$$

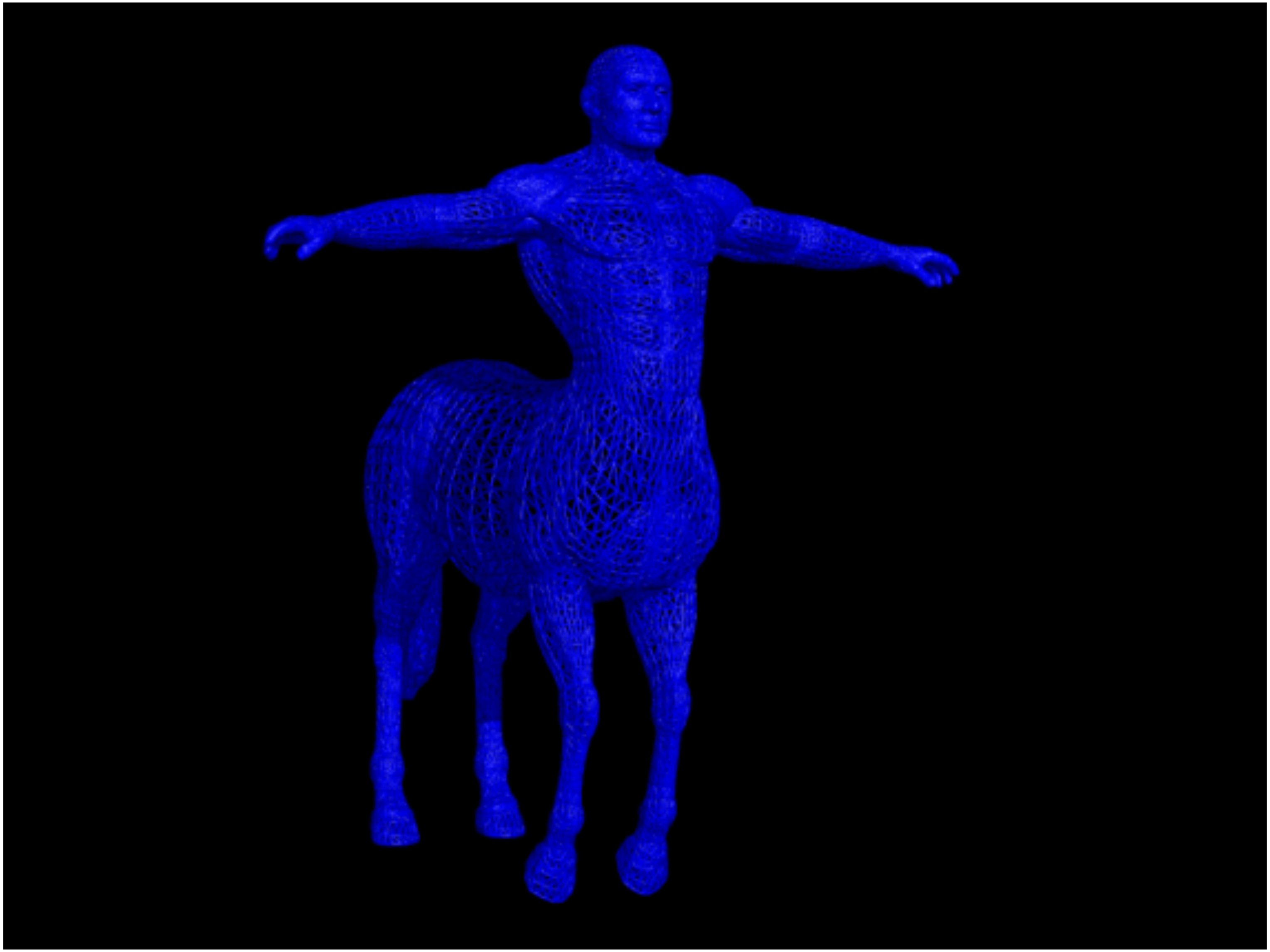
Transitions between two states



$$X(t) = \sinh(Ht)\operatorname{csch}(Ht_0)(X_t - X_0) + X_0$$

$$X(t) = \sinh(P(t-T))\operatorname{csch}(P(T-t_0))(X_f - X_t) + X_f$$



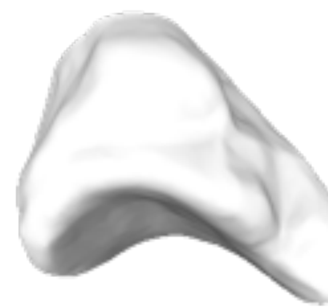


Analysis of anatomical data

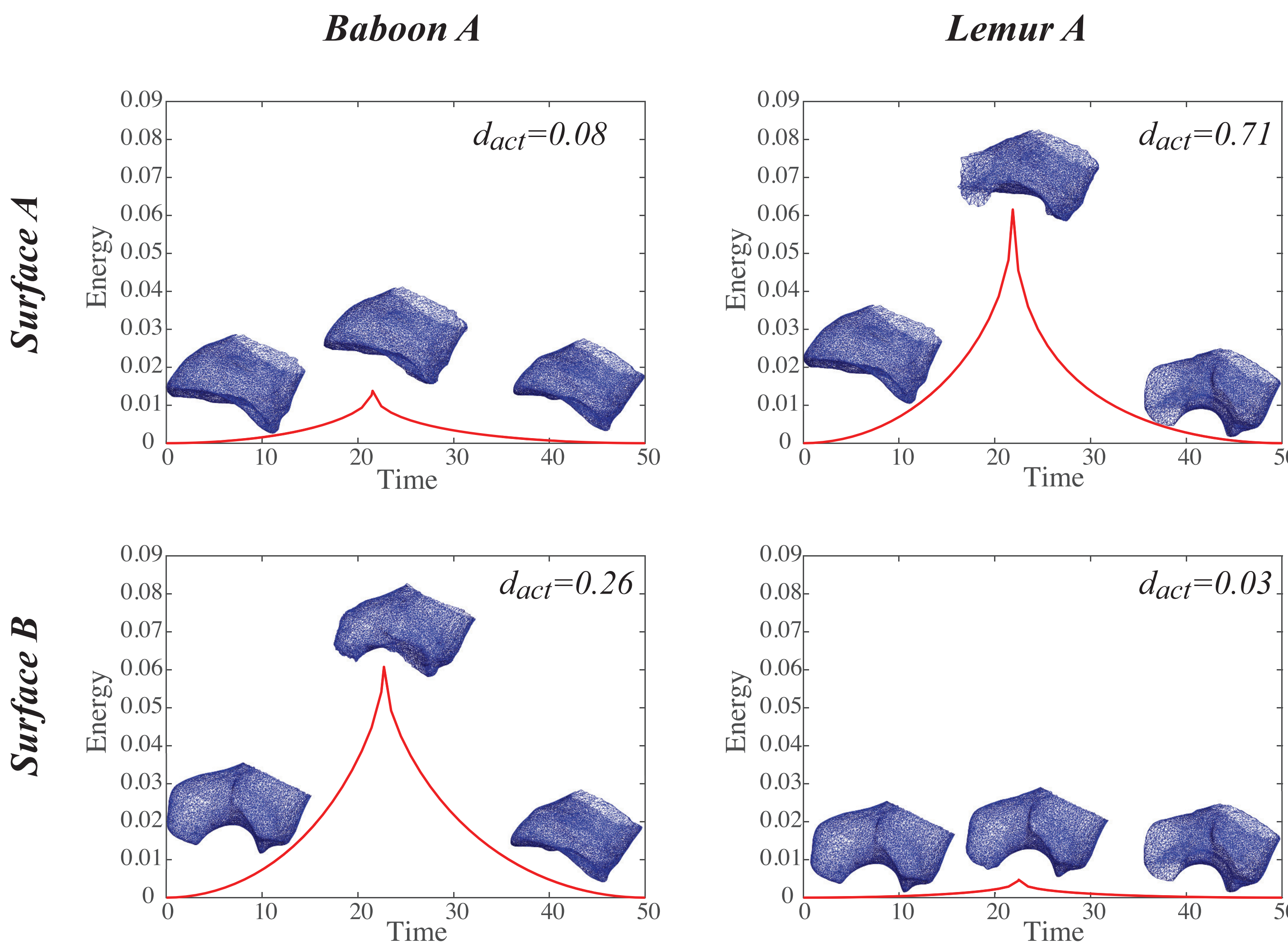
Dataset of proximal first metatarsals from 38 prosimian primates, and 23 New and Old World monkeys

Prosimian:
lemur

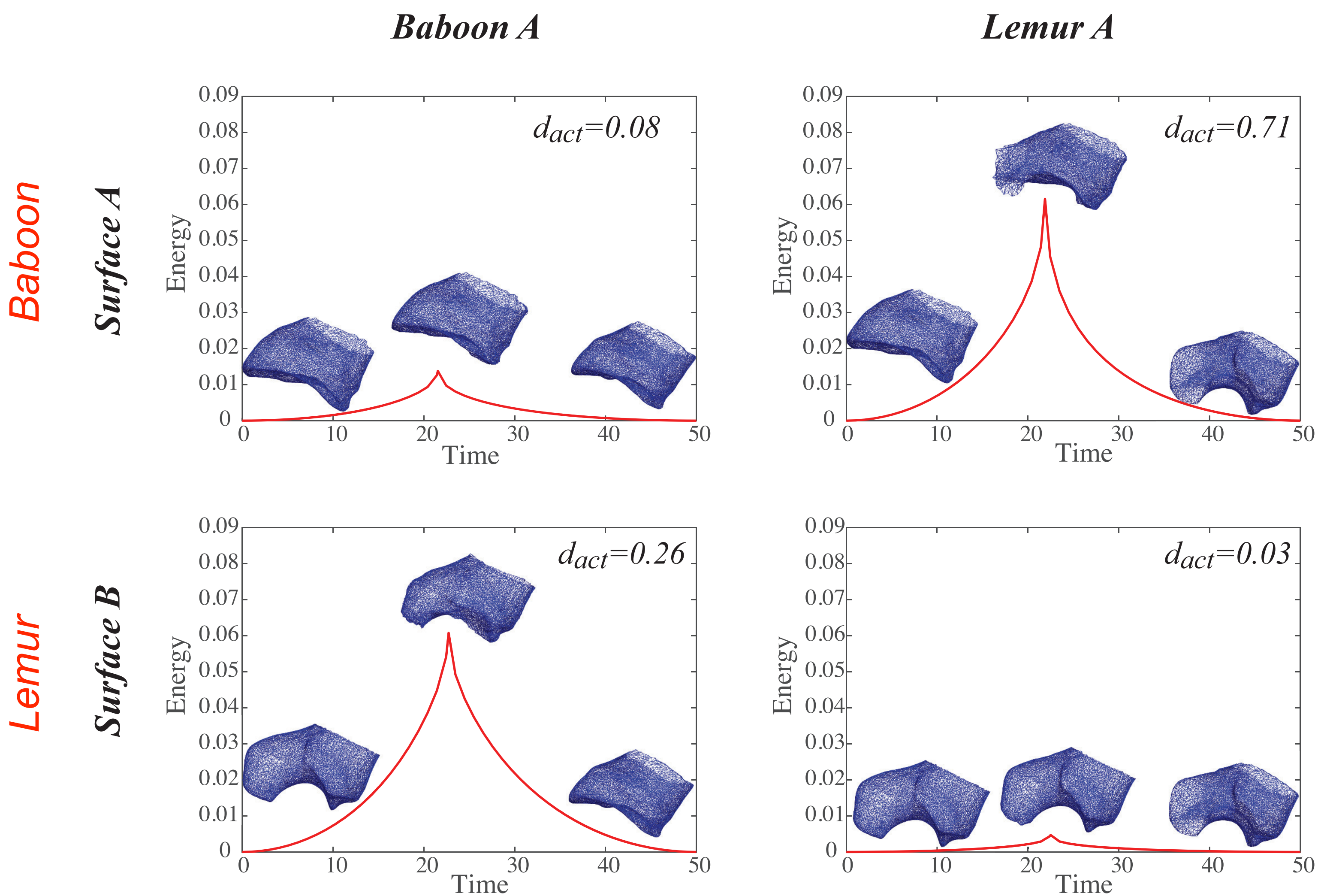
Simian:
Cape baboon
(old world)



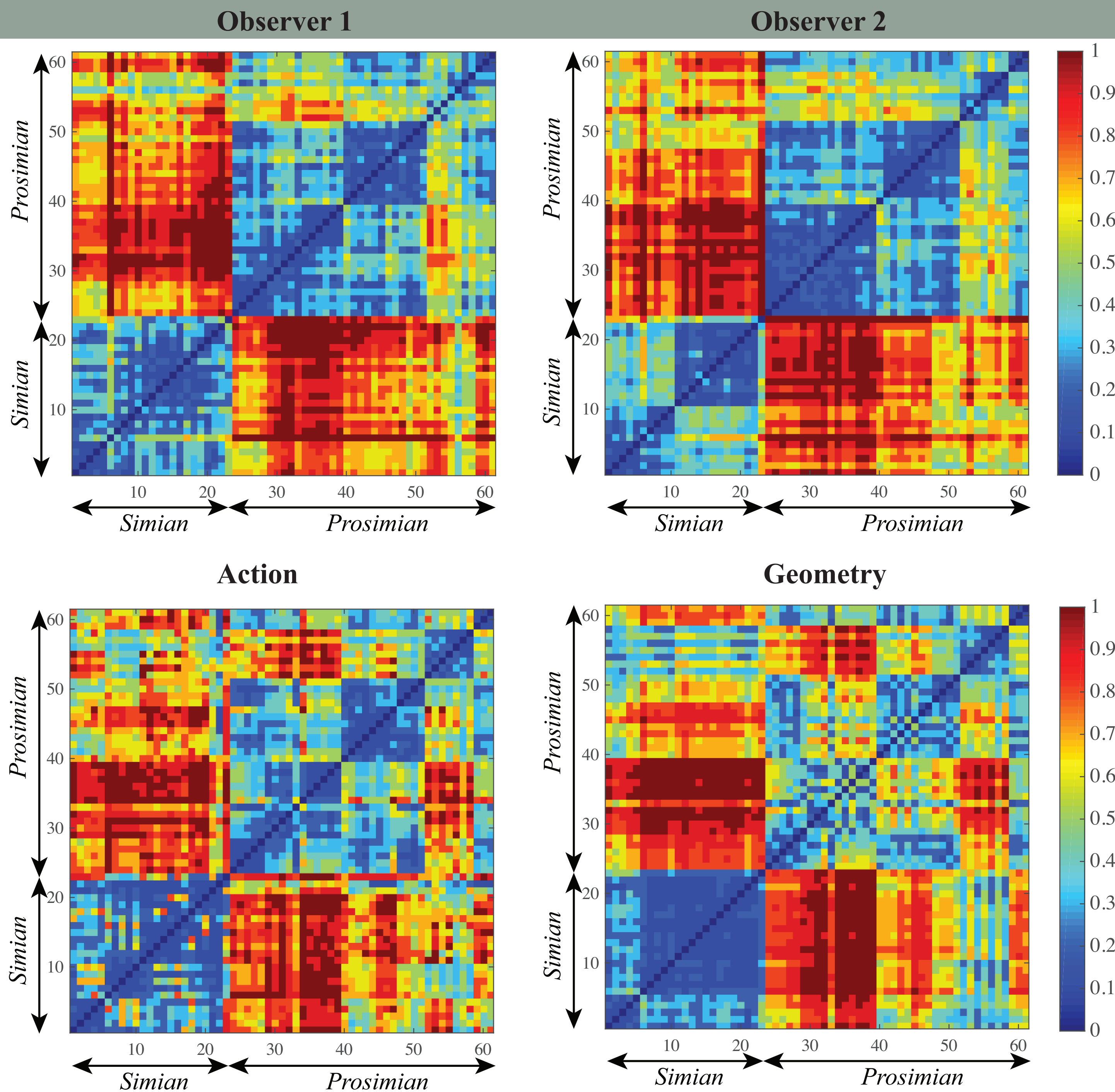
Simian:
White eared titi
(new world)



Baboon A
Lemur A
Surface A
Surface B
Energy
Time
d_act=0.08
d_act=0.71
d_act=0.26
d_act=0.03

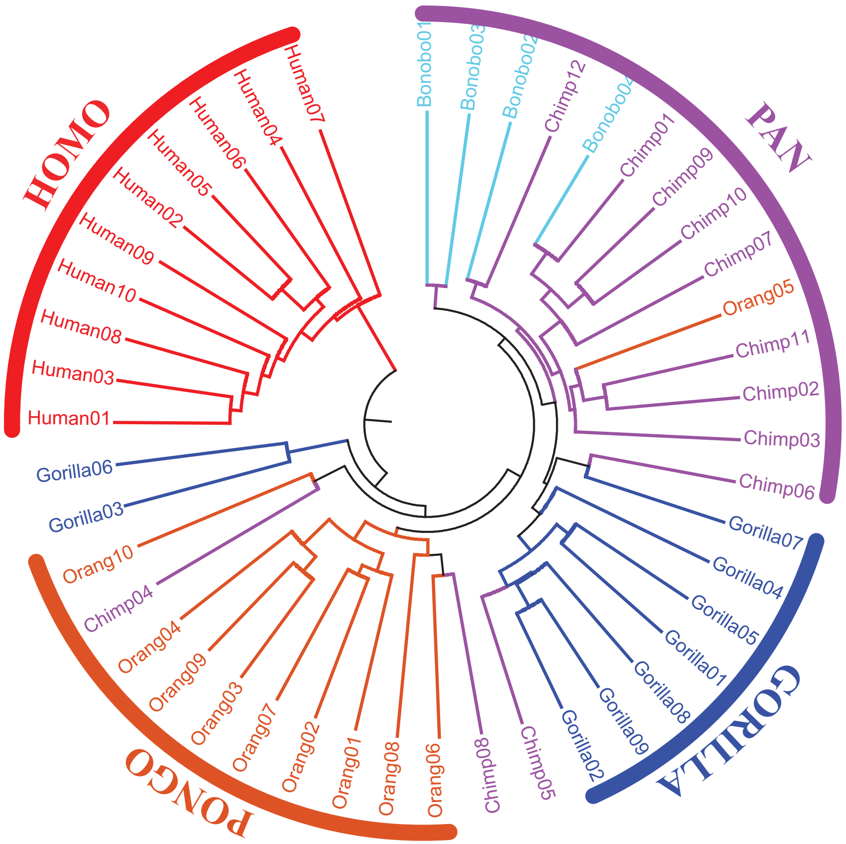


Baboon A
Lemur A
Baboon
Surface A
Lemur
Surface B
d_{act}=0.08
d_{act}=0.71
d_{act}=0.26
d_{act}=0.03
Energy
Time

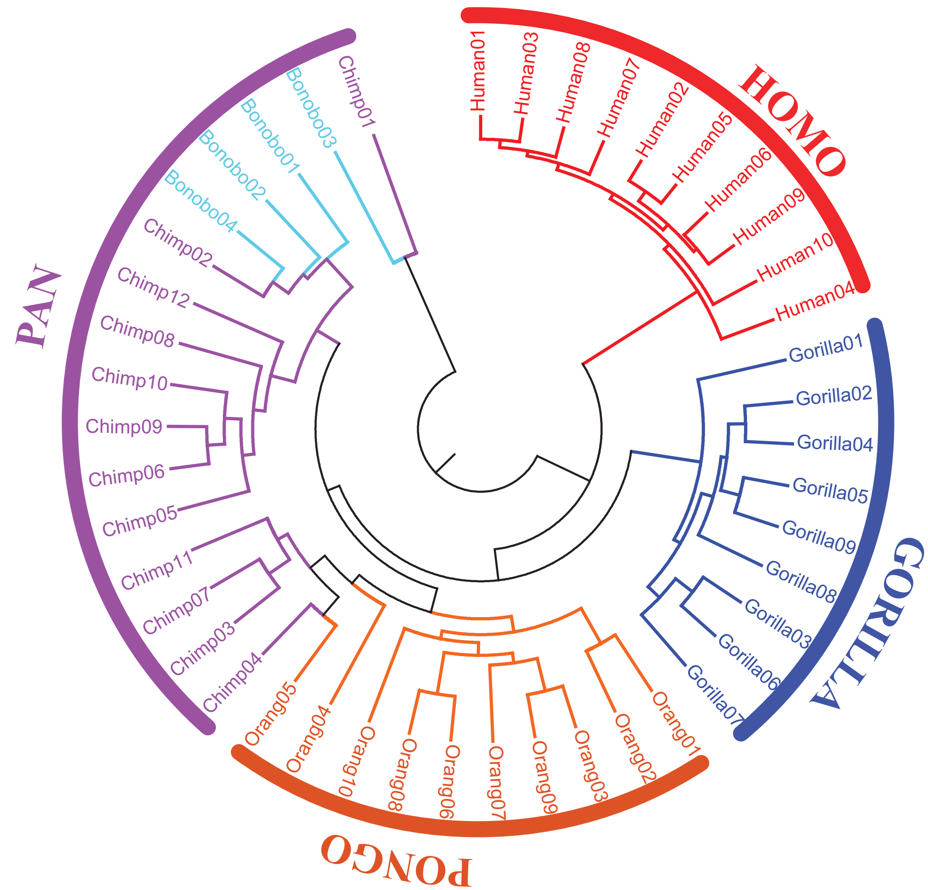


Observer 1
Observer 2
Action
Geometry
Prosimian
Simian
Prosimian
Simian
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
10
20
30
40
50
60

Trained morphometrist

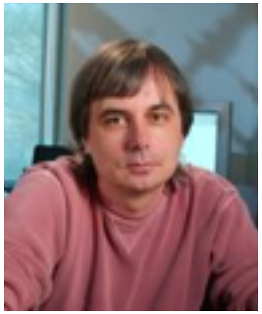


Variational distance



Thank You

Marc Delarue, Institut Pasteur, Paris
Henri Orland, CEA, Saclay



Herbert Edelsbrunner (IST Austria),
Seb Doniach (Stanford University)

Michael Levitt, Stanford U.
Joel Hass, UC Davis

Funding: NIH, NSF, Sloan Foundation, NU Singapore