The circular SiZer, inferred persistence of shape parameters and application to stem cell stress fibre structures

Stephan Huckemann[1], Kwang-Rae Kim, Axel Munk[1], Florian Rehfeldt[2], Max Sommerfeld[1], Joachim Weickert[4], Carina Wollnik[2]

[1]University of Göttingen - Institute for Mathematical Stochastics

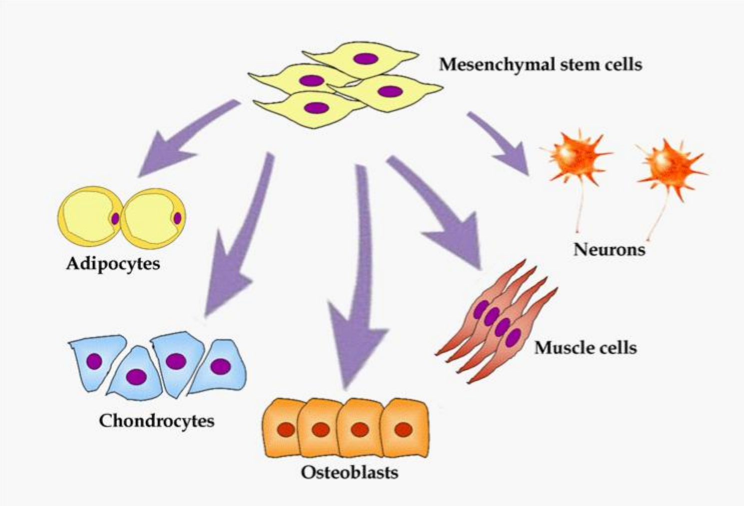
[2]University of Göttingen - Third Institute of Physics - Biophysics

[3]University of Nottingham

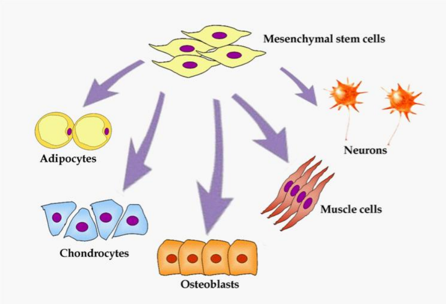
[4]Saarland University

October 20, 2015

Human Mesenchymal Stem Cells (hMSC)

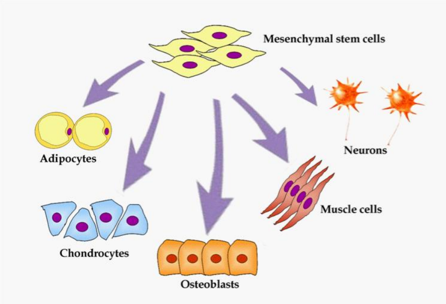


How does Differentiation work?



Long known answer: Chemistry!

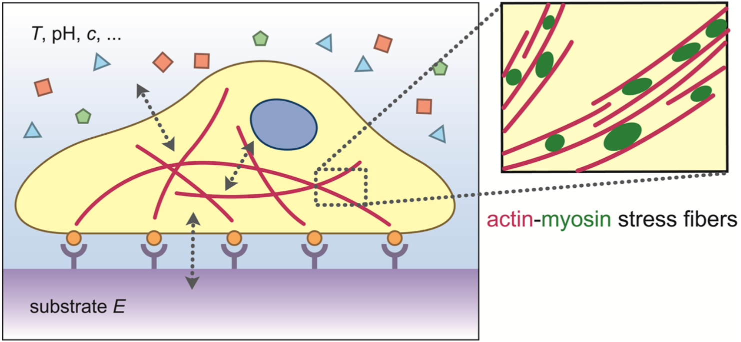
How does Differentiation work?

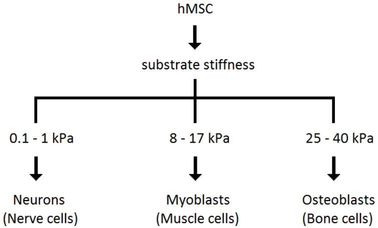


Long known answer: Chemistry!

... only chemistry?

Mechano-Chemical Environment

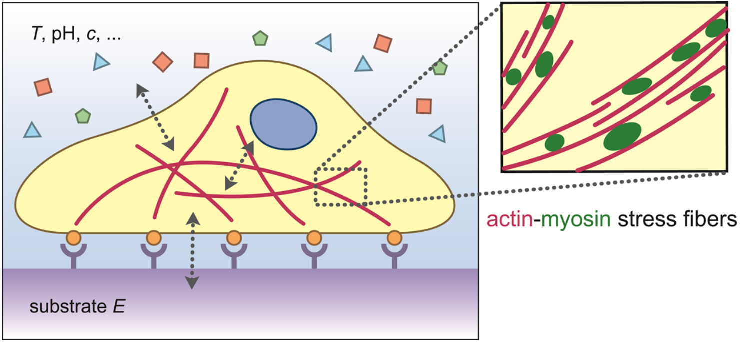


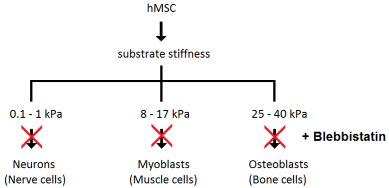


(A.J. Engler et al., *Cell* (2006))

- substrate stiffness influences differentiation
- actin-myosin stress fibers are the key players

Mechano-Chemical Environment

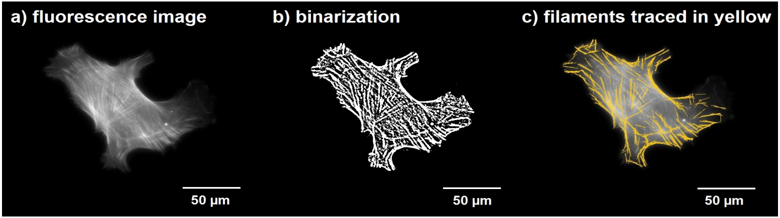




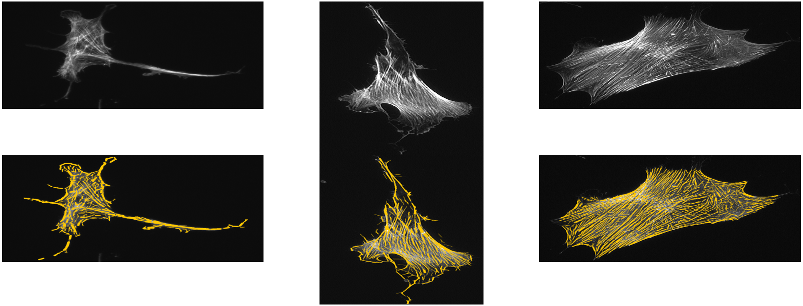
(A.J. Engler et al., *Cell* (2006))

- substrate stiffness influences differentiation
- actin-myosin stress fibers are the key players

Filament tracing



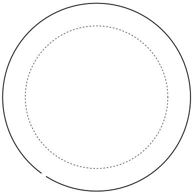
Software *FilPicker*: Gottschlich et al., *IEEE TIFS* (2009), Eltzner et al., *Plos One* (2015)



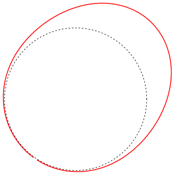
Soft - Fat or Neuron Medium - Muscle Hard - Bone

Orientations of Stress Fibers

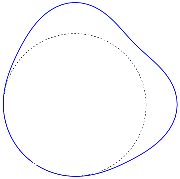
Hypothesis (Zemel, Rehfeldt et al., Nat. Phys. (2010)): The *distribution of filament orientations* distinguishes tissue types (neuron, muscle, bone)



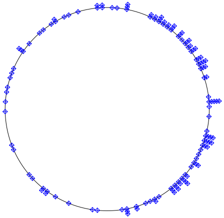
Soft 1 kPa - Fat, Neuron
No mode



Medium 11 kPa - Muscle
One mode



Hard 34 kPa - Bone
Two (or more) modes



Filament orientations

$Z_1, Z_2, \ldots, Z_n \overset{i.i.d.}{\sim} Z$ with values in

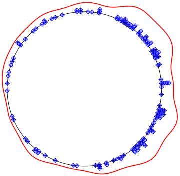
$$S^1 = \{z \in \mathbb{C} : |z| = 1\}$$

Kernel Density Estimator

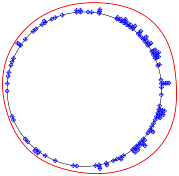
$\widehat{f}_h(z) = \frac{1}{n} \sum_{j=1}^{n} K_h(z \cdot Z_j^{-1})$

With a <u>circular</u> kernel K_h.

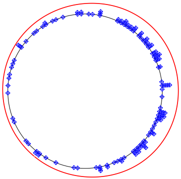
Fat, bone or muscle?



Bandwidth $h = 0.02$



Bandwidth $h = 0.1$



Bandwidth $h = 0.33$

Circular SiZer - Inference on the smoothed density $K_h * f$

Linear SiZer (Significant Zero Crossings): Chaudhuri and Marron, *AoS* (2000)

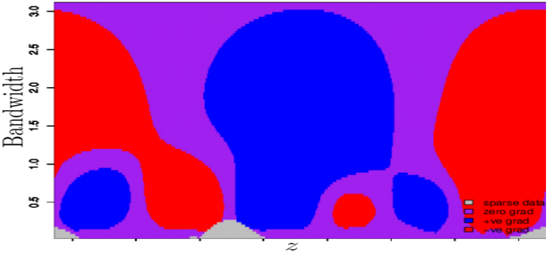
- **Distributional Limit:**

$$\sup_{z, h \geq h_0} \sqrt{n} \left| \partial_z \widehat{f}_h(z) - \partial_z (f * K_h)(z) \right| \to Y$$

  Y = supremum of Gaussian process on S^1.

- Test <u>simultaneously</u> for all z and $h \geq h_0$ the hypotheses

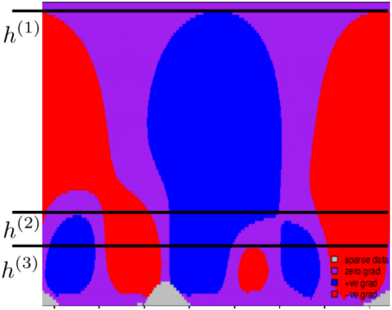
$$H_0^{(z,h)} : \partial_z (f * K_h)(z) = 0 \quad vs. \quad H_1^{(z,h)} : \partial_z (f * K_h)(z) \neq 0$$



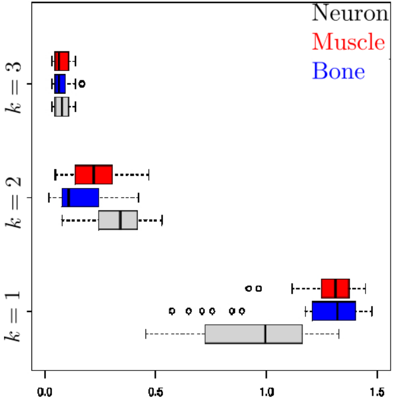
significant decrease / increase, no significance

Scale space persistence

Critcal bandwidth $h^{(k)}$:= largest bandwidth for which $\geq k$ modes are significant



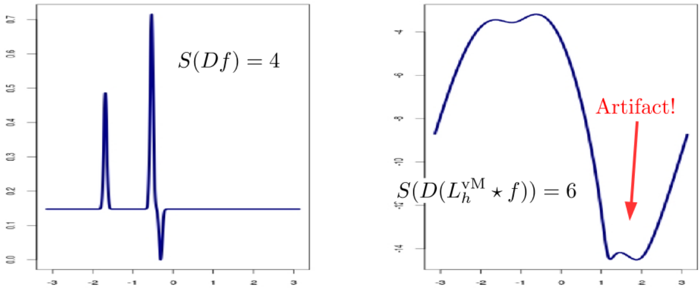
Data analysis



Causality

Definition of critical bandwidths $h^{(k)}$ only makes sense if we have **Causality:** No new modes appear as h increases.

Von-Mises $L_h^{vM} \propto I_o(h)^{-1} \exp(h \cos(x))$ does *not* have this property.



Causality

Definition of critical bandwidths $h^{(k)}$ only makes sense if we have **Causality:** No new modes appear as h increases.

- for linear data ($\in \mathbb{R}$): scale space theory - causality + semi-group property (essentially) uniquely characterizes the Gaussian kernel
- in higher dimensions (≥ 2) the Gaussian kernel does not satisfy causality
- if causality is replaced by non-enhancement of local extrema it is again unique

Uniqueness of the wrapped Gaussian

We say that a family of kernels $\{K_h : h > 0\}$

- has the causality property if

$$h \mapsto \#\text{Sign Changes}(K_h * f)$$

  is decreasing for all smooth f
- semi-group if $K_h * K_{h'} = K_{h+h'}$
- symmetric if $K_h(z) = K_h(-z)$
- regular if for some $r > 0$: $\lim_{h \downarrow 0} \left[\sup_{k \in \mathbb{Z} \setminus 0} \frac{|\hat{K}_{h,k} - 1|}{h|k|^r} \right] = 0$

Theorem

The only symmetric and regular semi-group on the circle that has the causality property is (up to rescaling of h) the family of wrapped Gaussians

$$K_h(e^{it}) = \frac{1}{\sqrt{2\pi h}} \sum_{k \in \mathbb{Z}} \exp\left(-\frac{(t + 2\pi k)^2}{2h} \right)$$

Another approach to circular data - work in progress

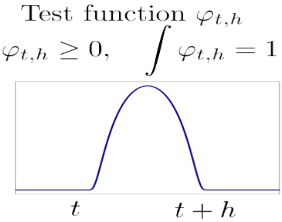
- kernel methods (such as WiZer) do not satisfy multisclae optimality bounds (see Axel Munk's talk)
- we can not handle contamination by random error (deconvolution)

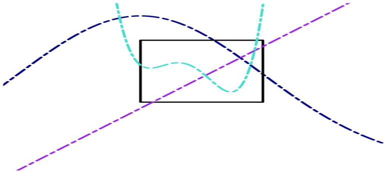
Circular deconvolution problem:

$$Y_i = Z_i \cdot \varepsilon_i, \quad i = 1, \ldots, n$$

If $Z_i \sim f$ and $\varepsilon_i \sim f_\varepsilon$ then $Y \sim g = f * f_\varepsilon$
Multiscale inference for qualitative features of f?

Idea: Do *not* reconstruct f but instead consider $\int(\mathcal{A}f)\varphi_{t,h}$ for a family of test functions $\varphi_{t,h}$





$$a \leq \int(\mathcal{A}f)\varphi_{t,h} \leq b \quad \Rightarrow \quad \mathrm{Graph}(\mathcal{A}f) \cap \{[t, t+h] \times [a, b]\} \neq \emptyset$$

Question: How can we access $\int (\mathcal{A}f)\varphi$?

$$\int (\mathcal{A}f)\varphi = \int f(\mathcal{A}^*\varphi) = \sum_{\xi\in\mathbb{Z}} \widehat{\mathcal{A}^*\varphi}(\xi)\widehat{f}(\xi)$$

$$= \sum_{\xi\in\mathbb{Z}} f_\varepsilon^{-1}(\xi)\,\widehat{\mathcal{A}^*\varphi}(\xi)\,\widehat{g}(\xi) = \int (\mathrm{op}(f_\varepsilon^{-1})\mathcal{A}^*\varphi)g$$

$$= E[(\mathrm{op}(f_\varepsilon^{-1})\mathcal{A}^*\varphi)(Y)]$$

This is accessible through samples of Y!

Roadmap:

- use local-global equivalence of the *periodic pseudo-differential* operator $\mathrm{op}(f_\varepsilon^{-1})\mathcal{A}^*$ to link to classical pseudo-differential operator
- use linear theory to obtain simultaneous multiscale confidence sets for $\int (\mathcal{A}f)\varphi_{t,h}$ (Schmidt-Hieber et al., *AoS* (2013), Dümbgen and Walther, *AoS* (2008))
- obtain confidence rectangles

Modehunting with confidence rectangles

$$T_{t,h} = n^{-1/2} \sum_{j=1}^{n} \mathrm{op}(f_\varepsilon^{-1}) \mathcal{A}^* \varphi_{t,h}(Y_j)$$

$$T_n = \sup_{t, l_n \le h \le u_n} w_{t,h} |T_{t,h} - ET_{t,h}| - c_h$$

Theorem (Schmidt-Hieber et al., AoS (2013) $T_n \to T^\infty$ in distribution; T^∞ distribution free