Memory and state dependent switching in biological diffusion

Scott A. McKinley, University of Florida, Mathematics.

Frontier Probability Days, May 19, 2014. Tucson, Arizona.

- Diffusion in Biological Fluids: Modeling, Simulation and Analysis

-- Diffusion in human mucus

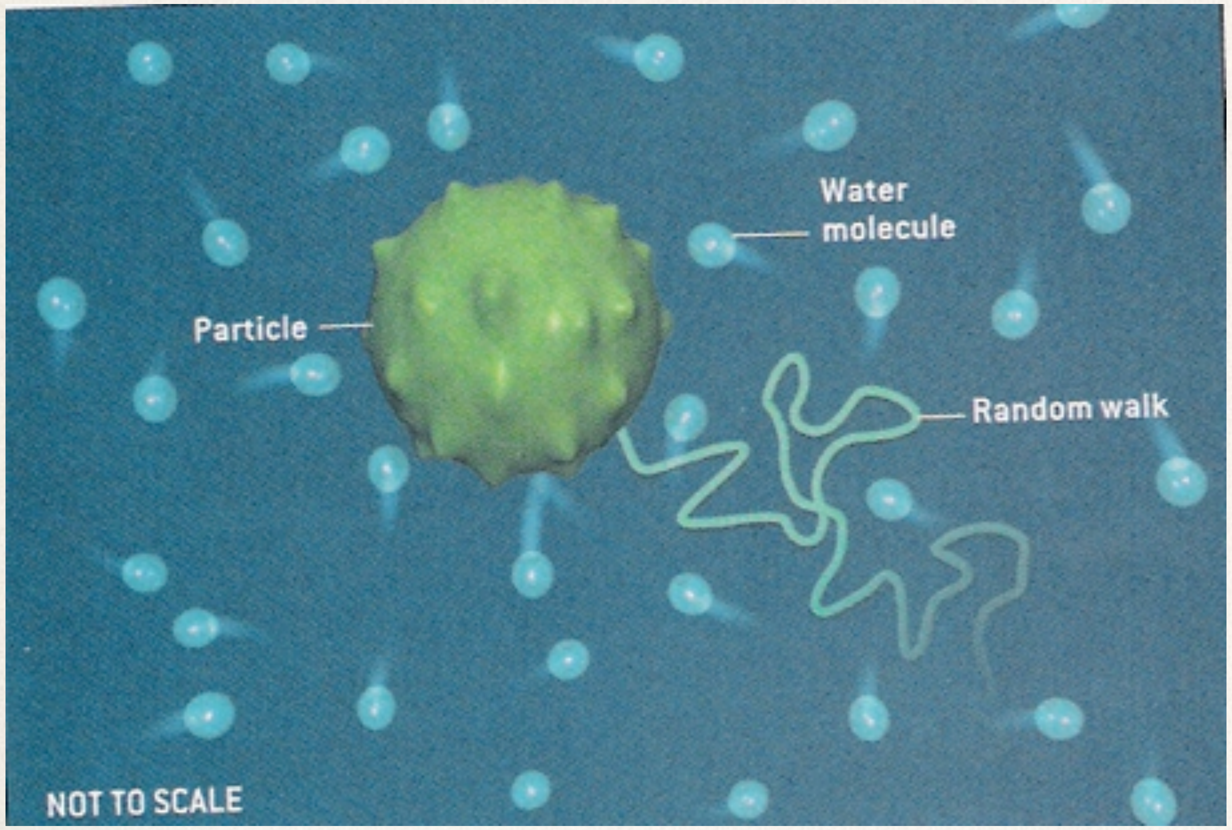
GLE: M.-, Yao & Forest. J Rheology (2009)
Whittle Estimation: Didier, M.-, Hill & Fricks. J Time Ser Anal (2012)
Data Analysis: D. Hill (CF Center), G. Forest (UNC), M.- and others PLOS ONE (2014)
Model comparison: M. Lysy (Waterloo), N. Pillai (Harvard), M.- in preparation.

-- Multi-particle diffusion in viscoelastic fluids

Work in progress with C. Hohenegger (Utah).

-- Geometric ergodicity of a bead-spring system

Mattingly, M.- & Pillai. Stoch Proc. & Appl. 122 (2012) 3953-3979.

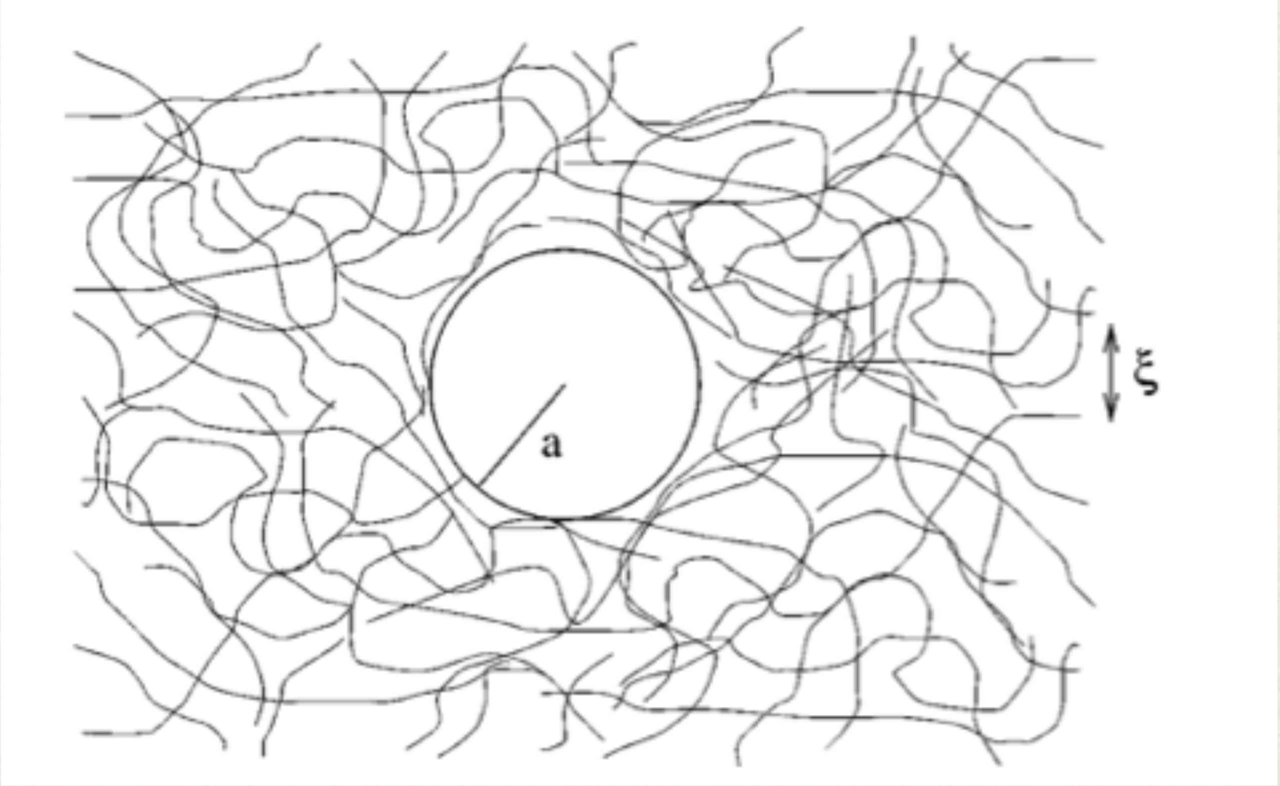


Viscous Diffusion

Image from Sci Am '85

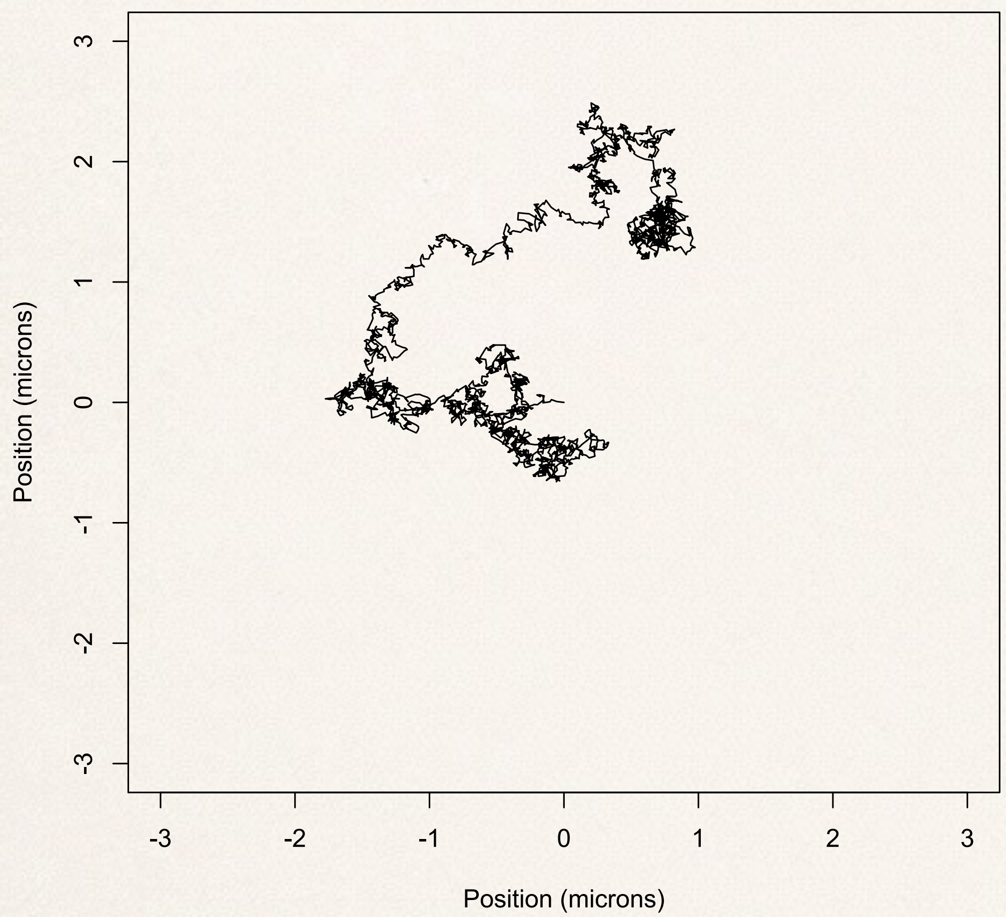
Viscoelastic Diffusion

Image from Levine & Lubensky, '03

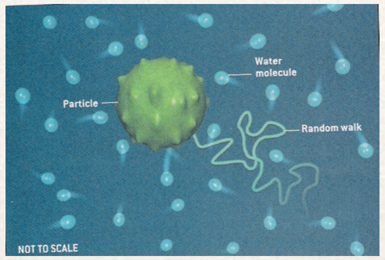


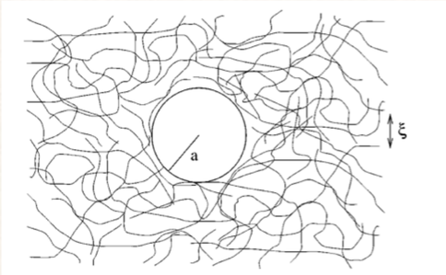
- Particles in biological fluids are semi-constrained.

Typical path for 500 nm beads in 2M concentration mucus.



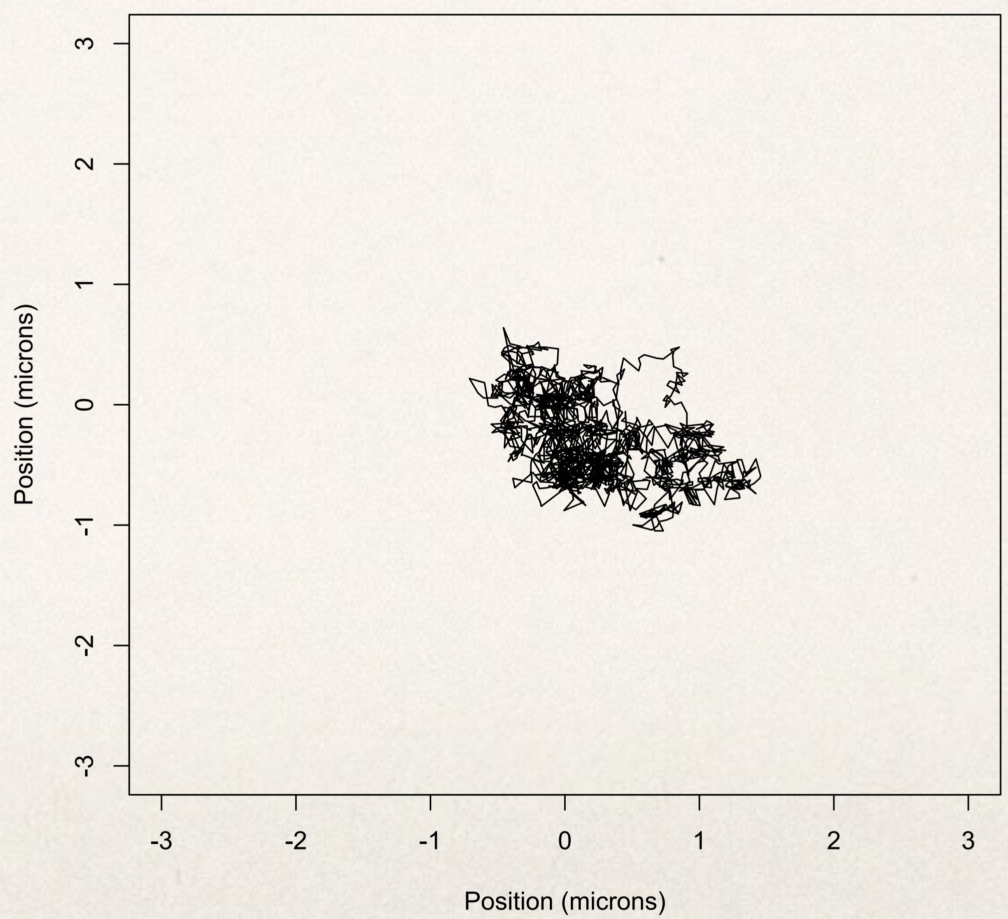
Viscous Diffusion





Viscoelastic Diffusion

Typical path for 500 nm beads in 2p5 concentration mucus.



- The driving scientific question

We have access to videos that last 10 seconds.

How do we ethically (in a mathematical sense)
extrapolate
to behavior of large populations over hours?

- Greg Forest's Innocent Question

I've got a subdiffusive process whose MSD scales like

$$\langle X^2(t) \rangle \sim t^\alpha$$

Since Brownian motion has a FPT that scales linearly

$$\mathbb{E}(\tau_L) = CL^2$$

is it true that $\mathbb{E}(\tau_L^{(\alpha)}) = CL^{2/\alpha}$?

Greg Forest's Innocent Question

I've got a subdiffusive process whose MSD scales like

$$\langle X^2(t)\rangle \sim t^\alpha$$

Since Brownian motion has a FPT that scales linearly

$$\mathbb{E}(\tau_L) = CL^2$$

is it true that $\mathbb{E}(\tau_L^{(\alpha)}) = CL^{2/\alpha}$?

Answer: For FBM, yes, but otherwise ….

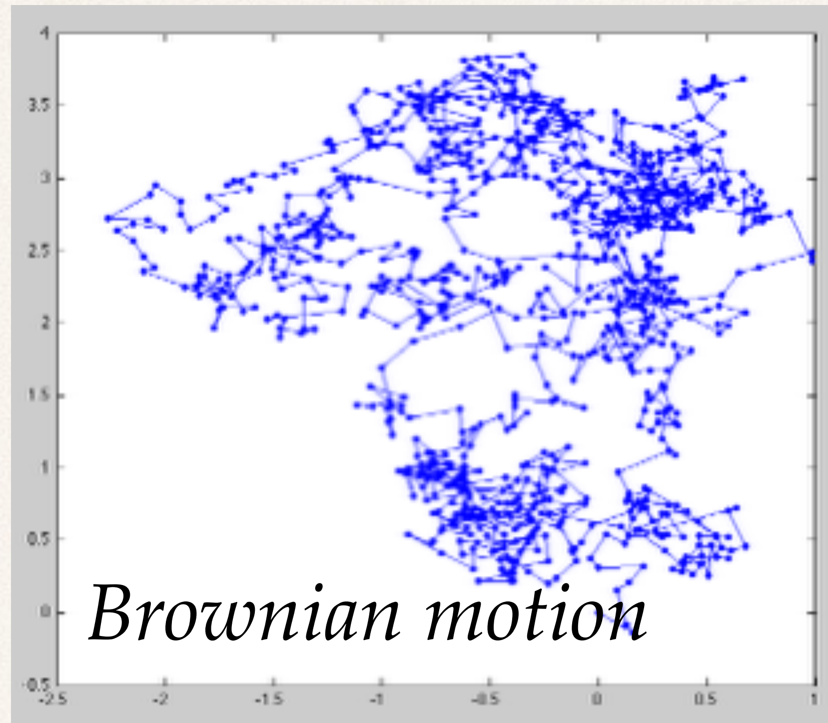
- Default statistical tool: Mean-squared Displacement (MSD)

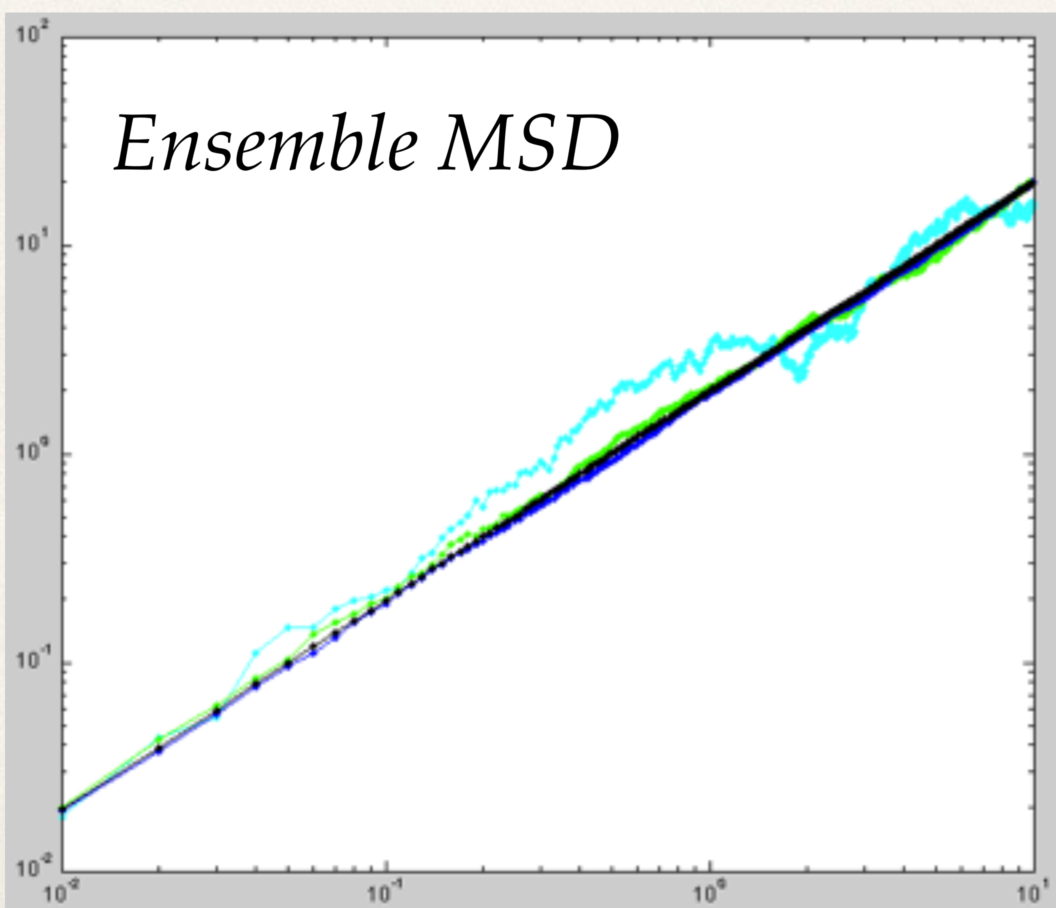
  -- *Ensemble MSD*

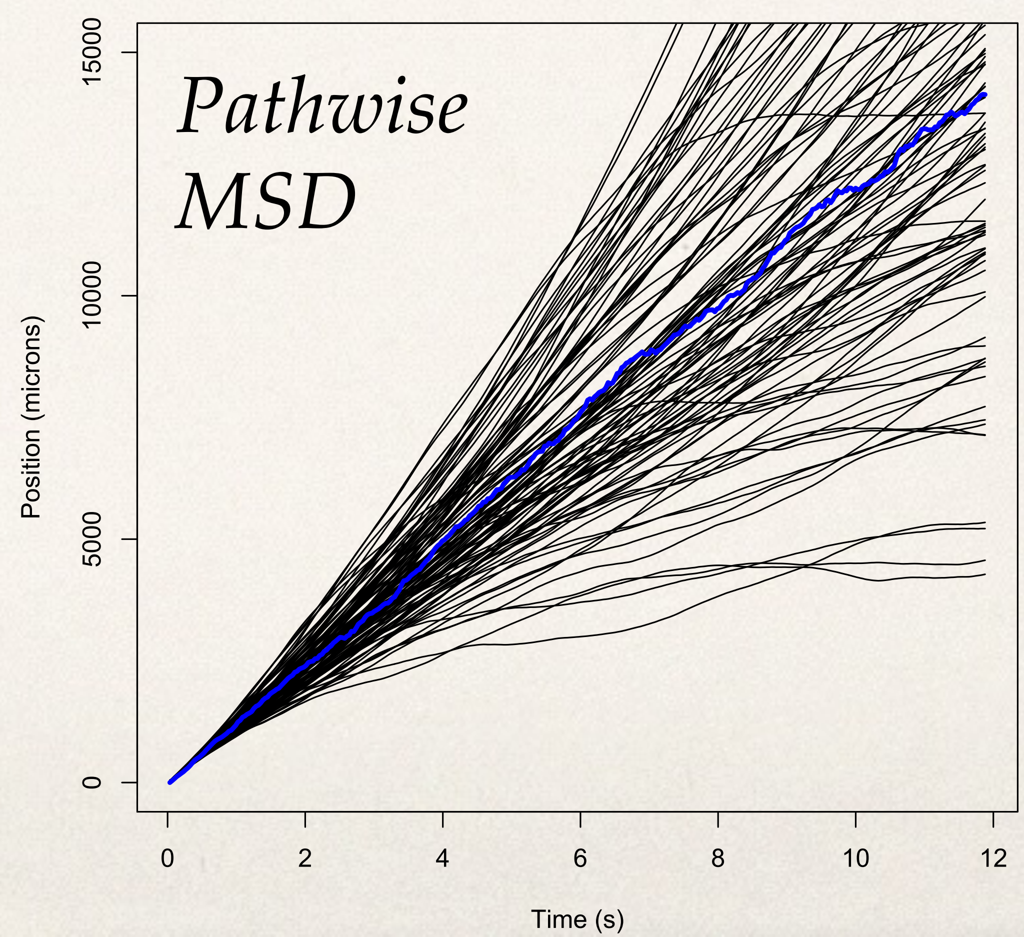
$$\langle X^2(t)\rangle = \frac{1}{M}\sum_{i=1}^{M}(X_i(t) - X_i(0))^2$$

  -- *"Pathwise" MSD*

$$\mu_i^2(t) = \frac{1}{T-t}\int_0^{T-t}(X_i(t+s) - X_i(s))^2 ds$$


Brownian motion


Ensemble MSD

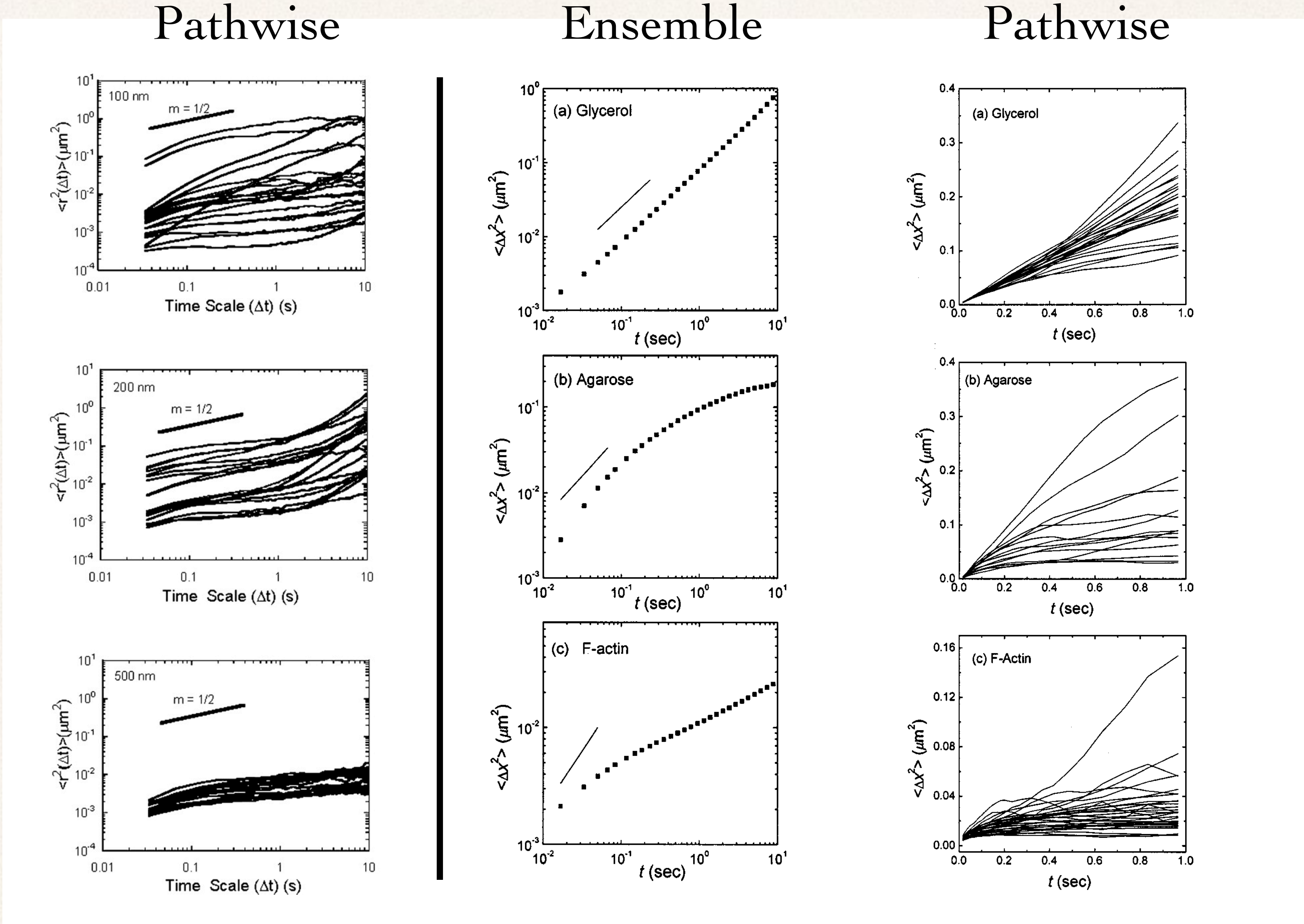

Pathwise
MSD
Position (microns)
Time (s)
0
5000
10000
15000
0
2
4
6
8
10
12

- Foreign particles in biological fluids exhibit *subdiffusive* behavior

100 nm in mucus

200 nm in mucus

500 nm in mucus



Glycerol

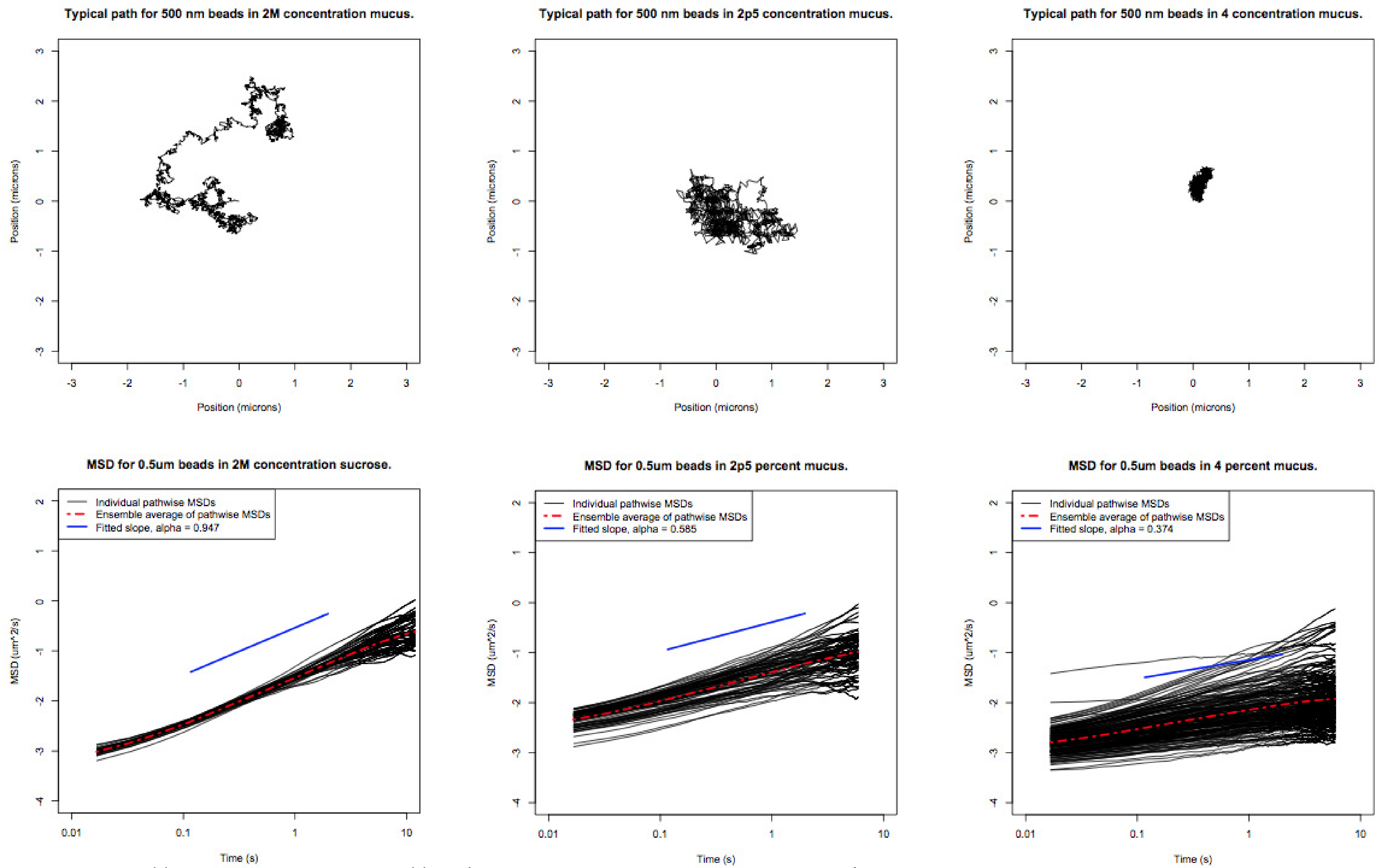
Agarose

F-actin

Left: Dawson et al 2003

Center, Right: Valentine, et al 2001

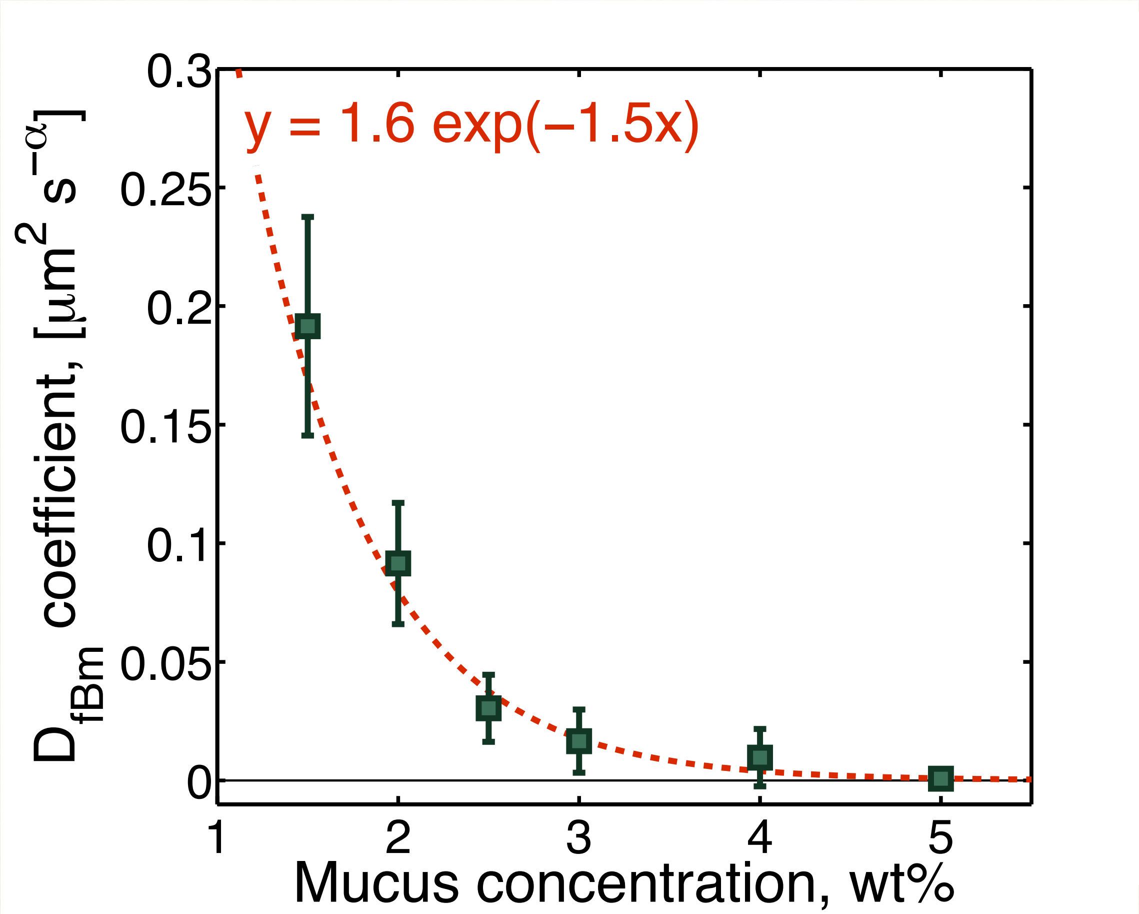
- Movement of foreign particles in mucus is sensitive to concentration



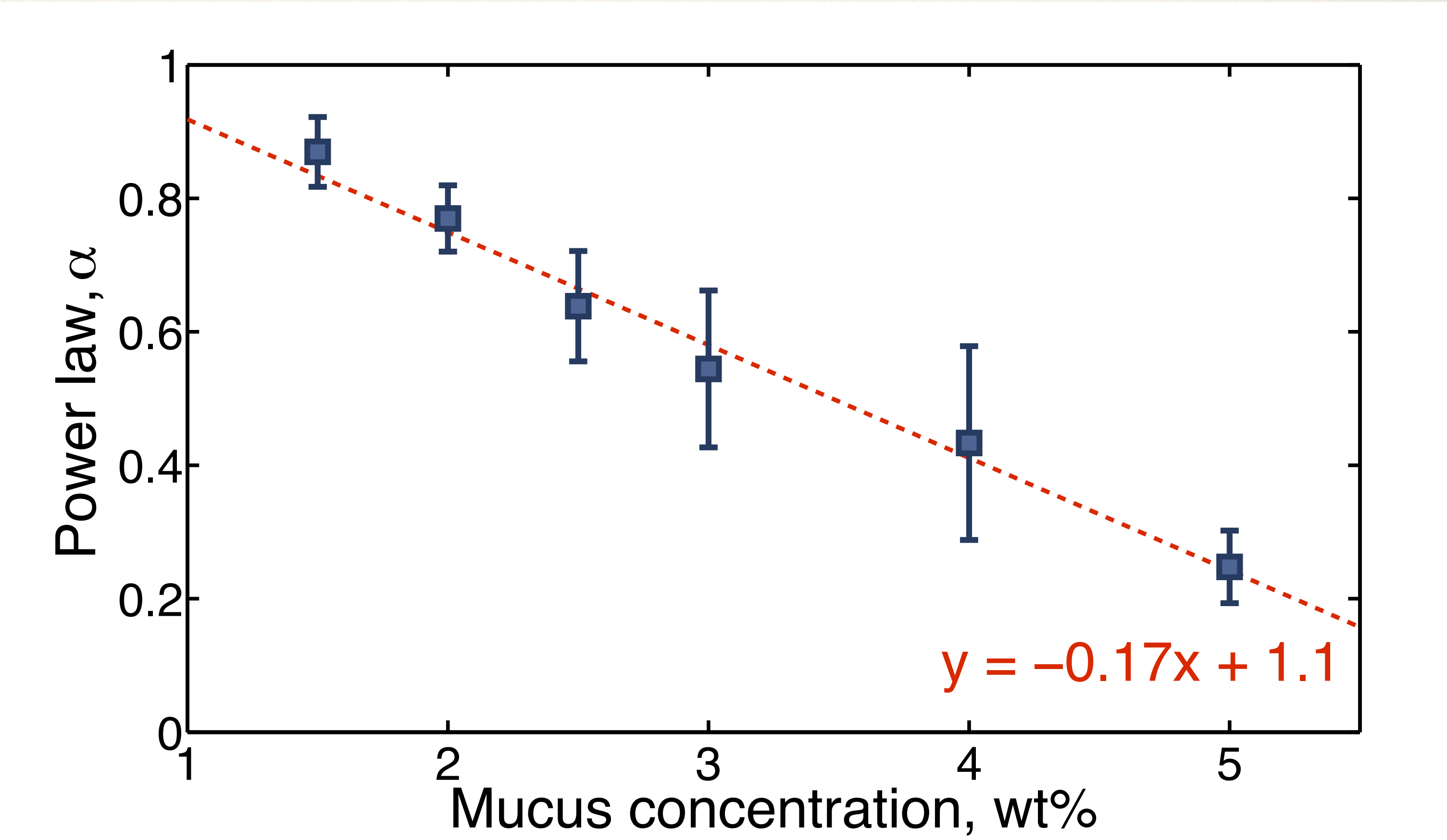
Hill, Vazquez, Mellnik, M.-, Vose, Mu, Boucher, Forest. PLOS ONE, 2014

- Movement of foreign particles in mucus is sensitive to concentration

Diffusion coefficient
vs mucin concentration



Subdiffusive exponent
vs mucin concentration



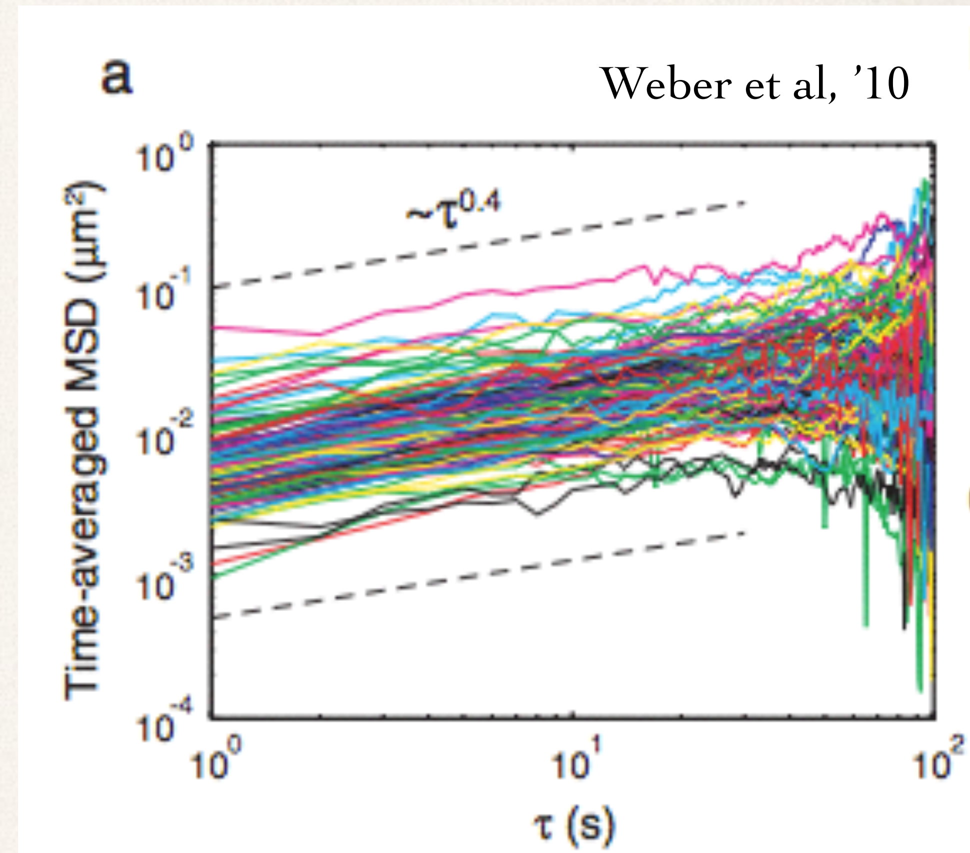
There is at present no physical theory that predicts this behavior!

- Model selection: Two regimes in the physics literature

Fractional Brownian motion in crowded fluids†

Dominique Ernst,‡[a] Marcel Hellmann,‡[b] Jürgen Köhler*[a] and Matthias Weiss*[b]

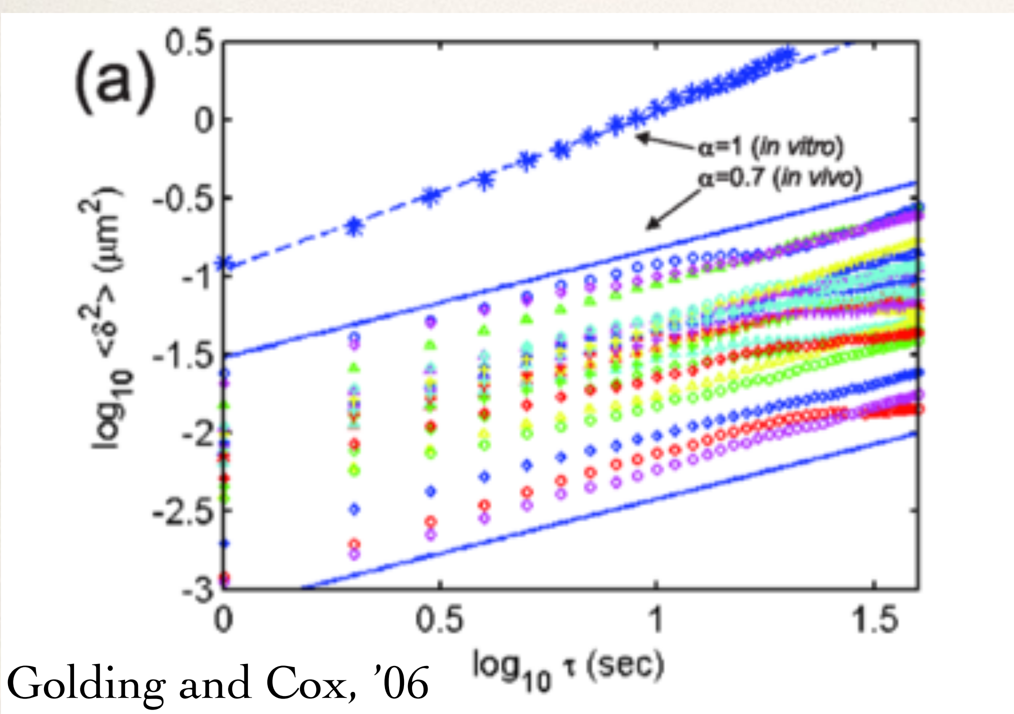
Received 30th January 2012, Accepted 22nd March 2012



Continuous-Time Random Walk

Single particle tracking in systems showing anomalous diffusion: the role of weak ergodicity breaking

Stas Burov,[a] Jae-Hyung Jeon,[b] Ralf Metzler*[b] and Eli Barkai*[a]



- Model selection: Continuous-time random walk hypothesis

-- Subdiffusion originates due to trapping and caging events

Theory: Saxton, Metzler, Klafter, Sokolov, Jeon, Lubelski ...

$$X(t) = \sum_{i=1}^{n} J_n 1_{t<\tau_n}(t)$$

Wait times $\{\tau_i\}_{i\geq 1}$ drawn from a heavy tailed distribution.

Jump sizes $\{J_i\}_{i\geq 1}$ drawn from a finite mean distribution.

Primary appeal: Non-ergodic in that an iid population will produce a wide spread of pathwise MSDs.

- Model selection: The Langevin Equation

-- *Position is the integral of a stationary velocity* $\dot{X}(t) = V(t)$

$$mdV(t) = -\gamma V(t)dt + \sqrt{2k_B T\gamma}dW(t)$$

(mass) (acceleration) thermal fluctuations

force due to drag

Physical Parameters

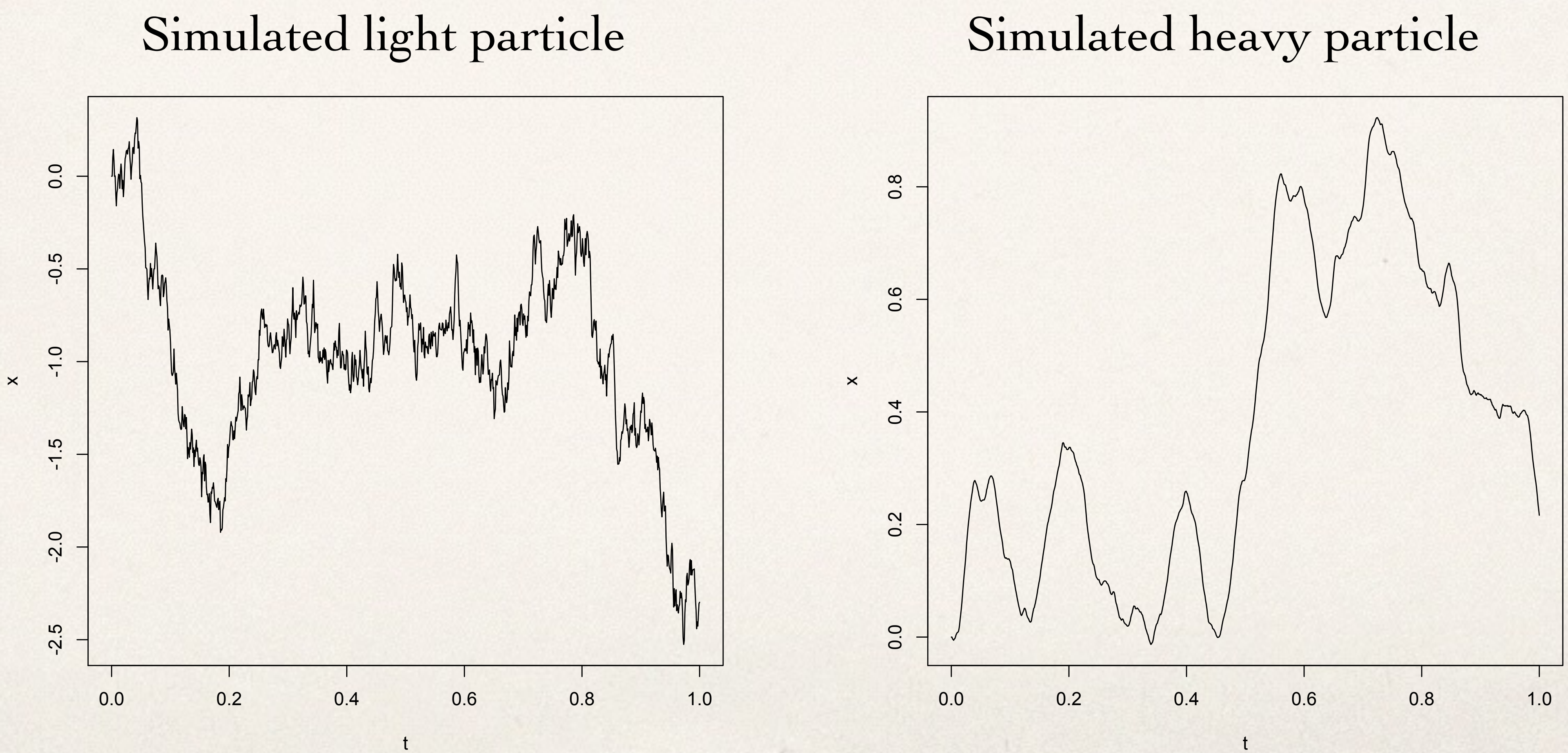
$\gamma = 6\pi r\eta$ Stokes drag for a spherical particle of radius r in a fluid with viscosity η .

$m,\ k_B,\ T$ Mass, Boltzmann Constant, Temperature

- Model selection: The Langevin Equation

  -- Position is the integral of a stationary velocity $\dot{X}(t) = V(t)$

$$mdV(t) = -\gamma V(t)dt + \sqrt{2k_BT\gamma}dW(t)$$



- Model selection: The Langevin Equation

-- Position is the integral of a stationary velocity $\dot{X}(t) = V(t)$

$$m dV(t) = -\gamma V(t) dt + \sqrt{2 k_B T \gamma}\, dW(t)$$

Zero mass limit is a mathematician's Brownian motion:

$$X(t) = \sqrt{2D}\, W(t)$$

where the diffusivity D is defined to be $D := \dfrac{k_B T}{\gamma}$

- Model selection: The Generalized Langevin Equation

  -- Position is the integral of a stationary velocity $\dot{X}(t) = V(t)$

$$mdV(t) = -\gamma V(t)dt + \sqrt{2k_B T\gamma}dW(t)$$

$$mdV(t) = -\int_0^t \Gamma(t-s)V(s)ds + F(t)$$

$$\mathbb{E}[F(t)F(s)] = k_B T\,\Gamma(|t-s|)$$

Mason & Weitz, 1995

- GLE: Construction of Solutions

Because the memory kernel Γ is a covariance function, it is non-negative definite. By Bochner's theorem, there exists a real, symmetric Radon measure μ such that

$$\Gamma(t) = \int_{\mathbb{R}} e^{it\omega} d\mu(\omega)$$

-- Sufficient Condition for Continuity of the Noise.
Suppose there exists an $a > 3$ such that μ satisfies

$$\int_0^{\infty} (\log(1+\omega)^a d\mu(\omega) < \infty$$

Then the sample paths of $F(t)$ are continuous almost surely.

See Lindgren, "Lectures on Stationary Stochastic Processes."

- GLE: Construction of Solutions

Path-by-path, the GLE is an Integral Equation that can be solved by Laplace transform.

$$mz\tilde{V}(z) = -\tilde{\Gamma}^{+}(z)\tilde{V}(z) + \tilde{F}(z)$$

where $\Gamma^{+}(t) := 1_{t>0}(t)\Gamma(t)$. It follows that

$$\boxed{V(t) = \int_{-\infty}^{t} \chi(t-s)F(s)ds}$$

where the kernel χ is defined on the Laplace side:

$$\tilde{\chi}(\omega) = \left(mz + \tilde{\Gamma}^{+}(z)\right)^{-1}$$

See Gripenberg et al, "Volterra Integral and Functional Equations."

GLE: Construction of Solutions

In order to define a broader class of solutions, we express $V(t)$ in terms of the spectral measure associated with $\rho(t) = \mathbb{E}[V(t)V(0)]$.

$$\begin{aligned}
\mathbb{E}[V(t)V(0)] &= \mathbb{E}[(\chi * F)(t)(\chi * F)(0)] \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \chi(t - t')\chi(0 - s')\mathbb{E}[F(t')F(s')]ds'dt' \\
&= \int_{\mathbb{R}} e^{i\omega t}|\hat{\chi}(\omega)|^2 d\mu(\omega)
\end{aligned}$$

which is to say, if μ has a density $\hat{\Gamma}(\omega)$, then

$$\boxed{\hat{\rho}(\omega) = \frac{k_B T\,\hat{\Gamma}(\omega)}{|mi\omega + \hat{\Gamma}^+(\omega)|^2}}$$

See K. Ito, "Stationary Random Distributions" (1954) or, more recently, S. Kou, *Ann. Appl Stat.* (2008), for similar calculations.

- Models for the memory kernel

Sum of exponentials: Heat bath models, Polymer kinetics, Muti-mode Maxwell (References included later.)

$$\Gamma(t) = \gamma\Big(\delta(t) + \sum_{k=1}^{N} C_k e^{-\lambda_k t}\Big)$$

Fractional Gaussian noise: Kou, '08; Didier, M.-, Hill, Fricks '12

$$\Gamma(t) = \gamma 2H(2H-1)|t|^{2H-2}$$

$$H \in (1/2, 1).$$

- Fractional Brownian motion from the Langevin point of view

GLE with Prony series kernel

Sum of exponentials memory kernel

$$K_p(t) := \frac{1}{N}\sum_{n=1}^{N} c_n e^{-t/\tau_n}$$

Maxwell kernel

-- Small number of modes

-- Arbitrary form

$\big\downarrow$ $N = 1, \tau_1 \to 0$

Langevin Equation

Dirac δ function memory kernel

$\big\downarrow$ $m \to 0$

Brownian motion

. . . .

Generalized Rouse Kernel

-- Large number of modes

-- Prescribed form

$$\tau_n = \tau_0 \left(\frac{N}{n}\right)^{\rho}$$

$\big\downarrow$ $N \to \infty, \tau_0 \to 0$

Fractional Langevin Equation

Power law memory kernel

$$K_H(t) := 2H(2H-1)t^{-2H}$$

$\big\downarrow$ $m \to 0$

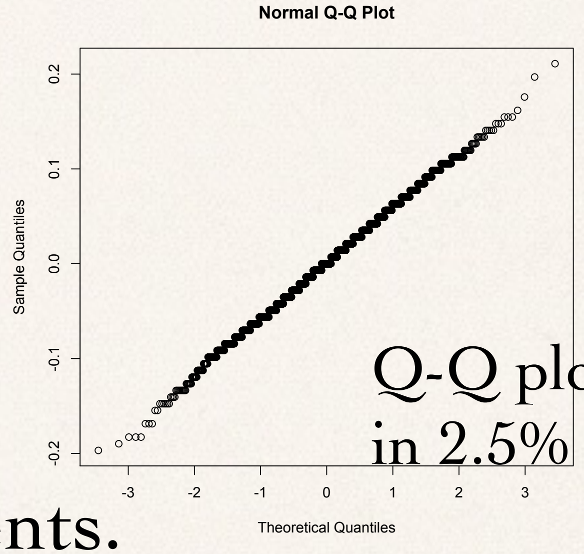
Fractional Brownian motion

- Model selection in our first paper: FBM vs CTRW

    -- *Ad hoc analysis supports Fractional Brownian Motion*

The data is has anomalous power law MSD, but also …

The data is (mostly) Gaussian.



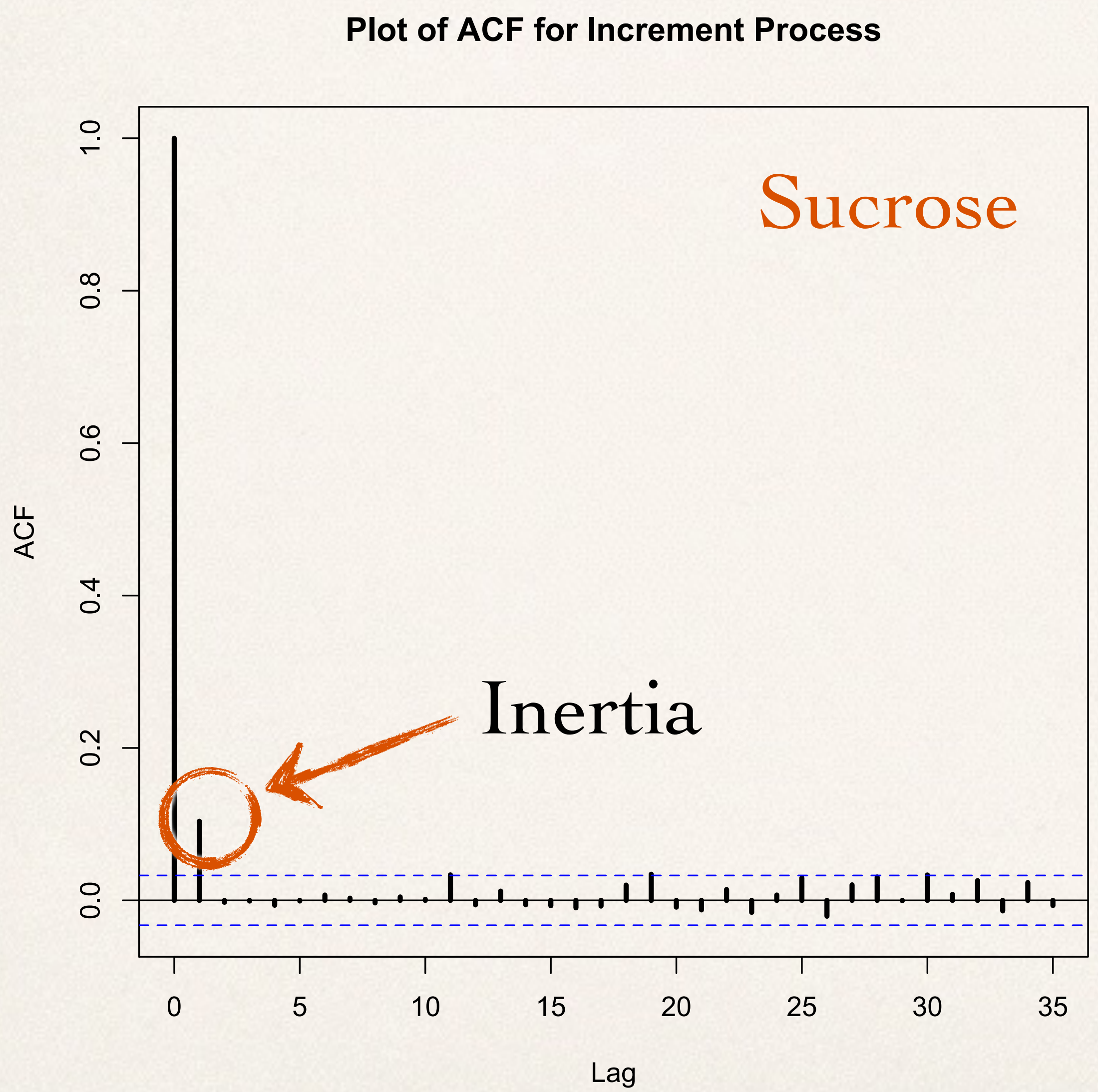
Q-Q plot of diffusion in 2.5% mucus

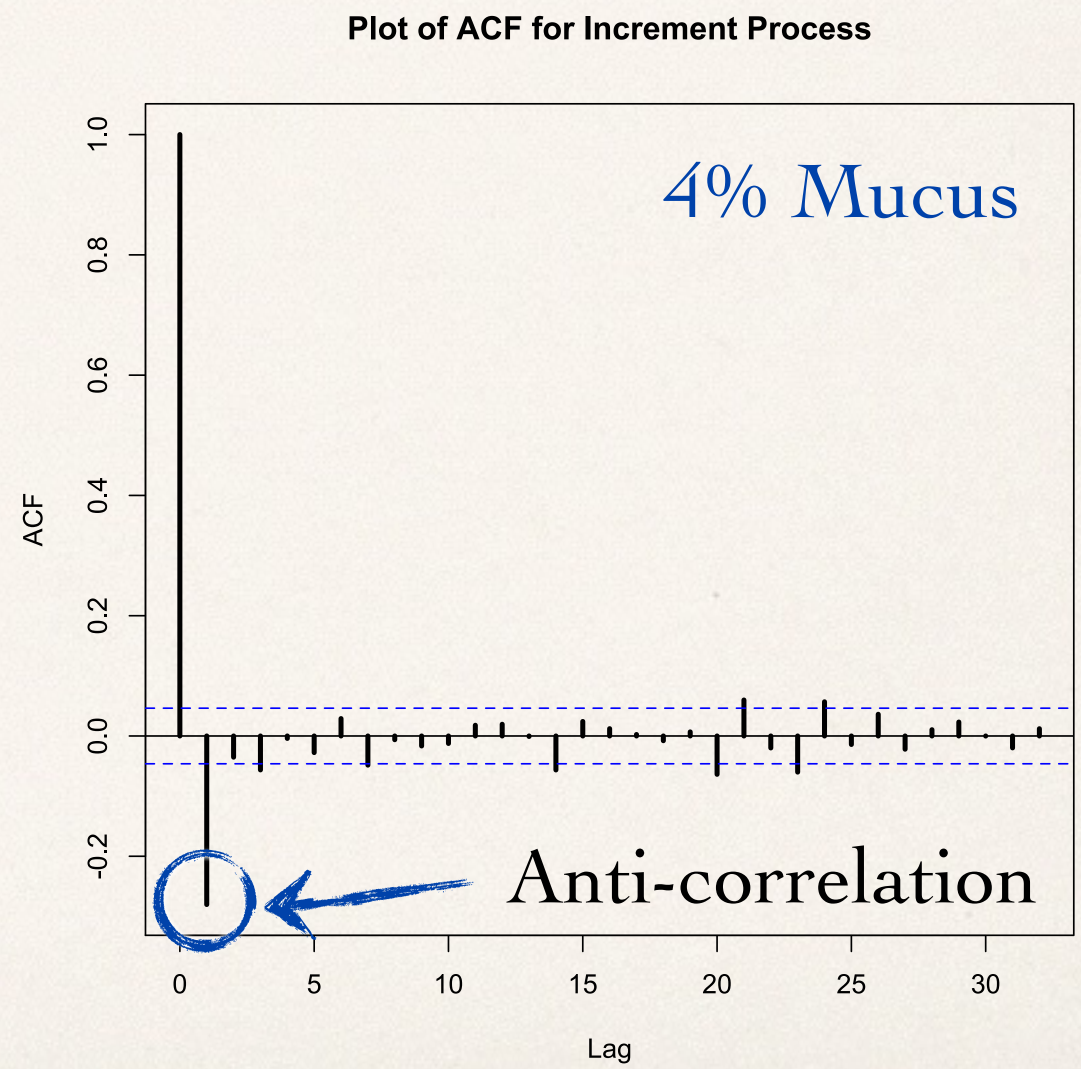
The paths have stationary increments.

The ensemble average of the autocorrelation function certainly *looks* like that of FBM!

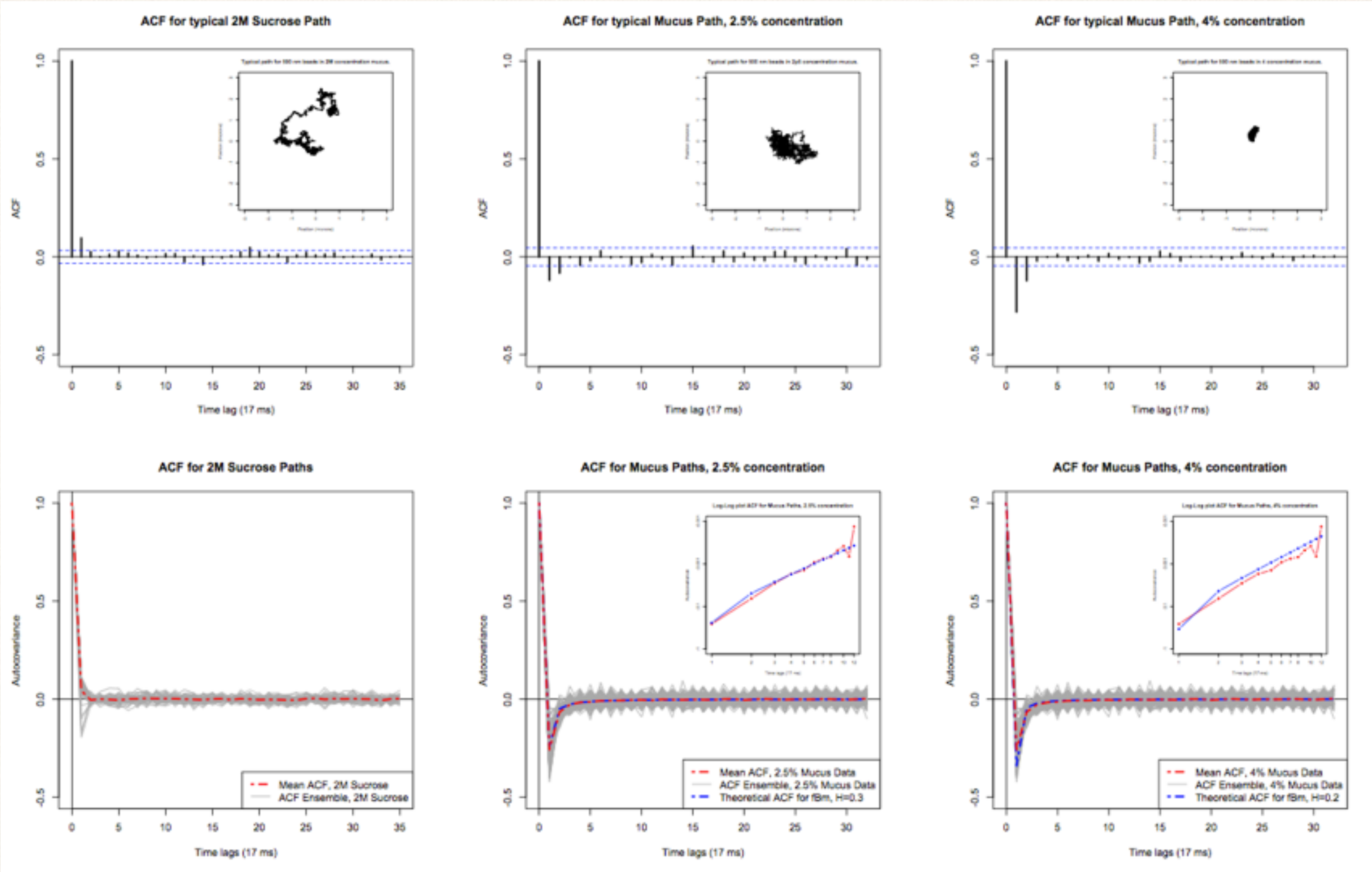
- Ad hoc method: Auto-correlation function

$$\rho(t) = \mathbb{E}[V(t+s)V(s)]$$





- ACFs for particles in mucus suggest Fractional Brownian Motion



- More rigorous approach: Bayes Factors (with Lysy and Pillai)

$$p(M_i \,|\, x) = \frac{\pi(M_i) f_i(x)}{\pi(M_1) f_1(x) + \pi(M_2) f_2(x)},$$

where $$f_i(x) = \int p(x \,|\, \theta_i, M_i)\pi(\theta_i \,|\, M_i) \mathrm{D}\theta_i$$

Proposed model for the data: M_i

The data: x Prior distributions: $f_i,\ \pi(\theta \,|\, M_i)$

Likelihood of seeing data given model and parameters:

$$p(x \,|\, \theta_i, M_i)$$

- More rigorous approach: Bayes Factors (with Lysy and Pillai)

Candidate models:

— Fractional Brownian motion

— Zero-mass GLE with generalized Rouse Spectrum with 4 modes (Mason-Weitz), 200 modes (M., Yao, Forest)

$$p(M_i \,|\, x) = \frac{\pi(M_i) f_i(x)}{\pi(M_1) f_1(x) + \pi(M_2) f_2(x)},$$

where $$f_i(x) = \int p(x \,|\, \theta_i, M_i) \pi(\theta_i \,|\, M_i) \mathrm{D}\theta_i$$

- More rigorous approach: Bayes Factors (with Lysy and Pillai)

TABLE 1

Posterior probability using the actual data for GLE-200 vs. fBM. Boldface values indicate that the model comparison favors GLE-200.

Dataset	1	2	3	4	5	6	7	8	9	10
α (%)	55.3	86.3	**62.9**	50.5	**74**	**57**	**91.2**	74.9	**59.6**	**93.3**
β (%)	19.1	2.6	10	24.2	3.7	14.6	0.6	6.5	12.2	0.2

TABLE 2

Posterior probability using the actual data for GLE-200 vs. GLE-4. Boldface values indicate that the model comparison favors the GLE-200.

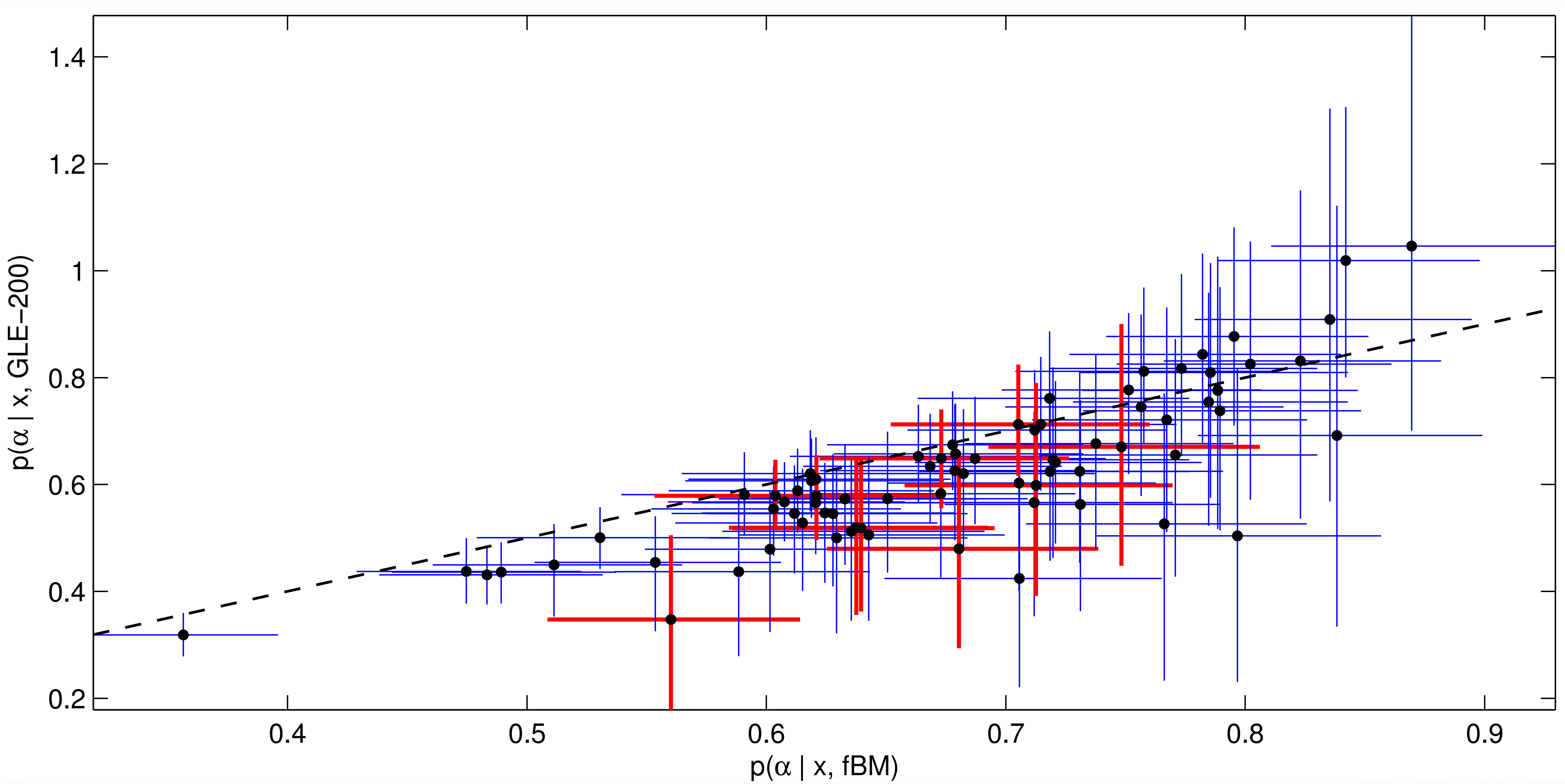
Dataset	1	2	3	4	5	6	7	8	9	10
α (%)	**74.2**	**95.8**	**69.2**	**80.8**	**86.2**	52.4	61.3	**78.8**	**76**	**50.8**
β (%)	6.8	0	9.6	3.7	1.8	26.5	16.7	4.4	5.6	25.3

TABLE 3

Posterior probability using the actual data for fBM vs. GLE-4. Boldface values indicate that the model comparison favors fBM.

Dataset	1	2	3	4	5	6	7	8	9	10
α (%)	**78**	**99.3**	**56.9**	**81.1**	**68.6**	59.3	94.2	**91.7**	**68.2**	93.1
β (%)	6.8	0	14.4	5.3	10.1	9.9	0.4	1.2	10.3	0.4

- Distribution of posterior parameter estimation 2.5% mucus.



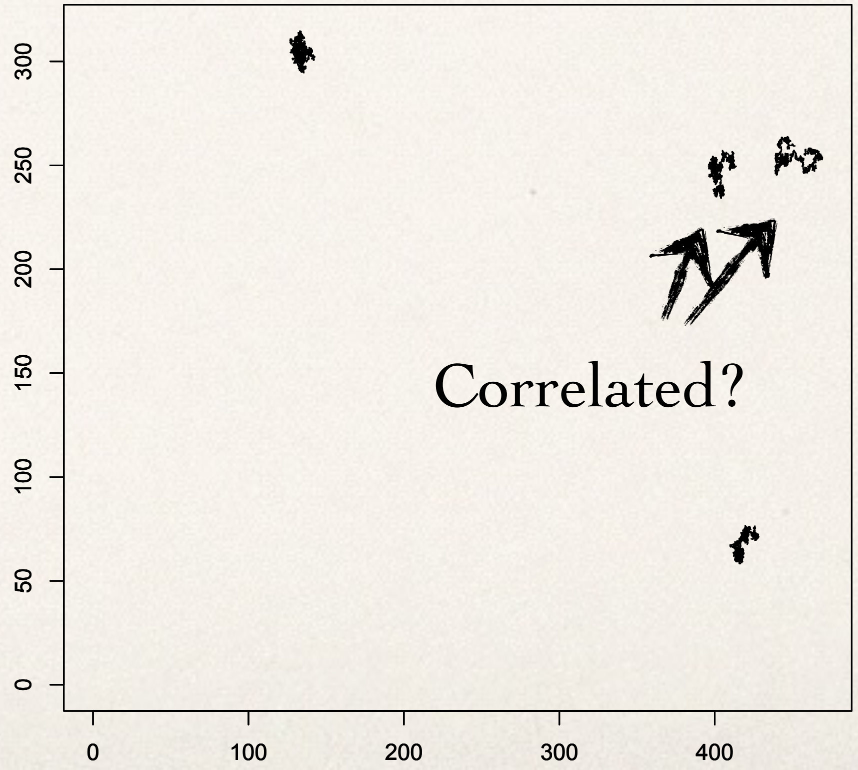
- Serious challenges that await

-- *Dealing with heterogeneity in simulation*

Hypothesis 1: Heterogeneity is particle-by-particle

Hypothesis 2: The fluid environment is heterogeneous

Hypothesis two would require a model that produces particles that exhibit a "location-dependent Hurst parameter." No such model exists!



-- *Two-point microrheology*

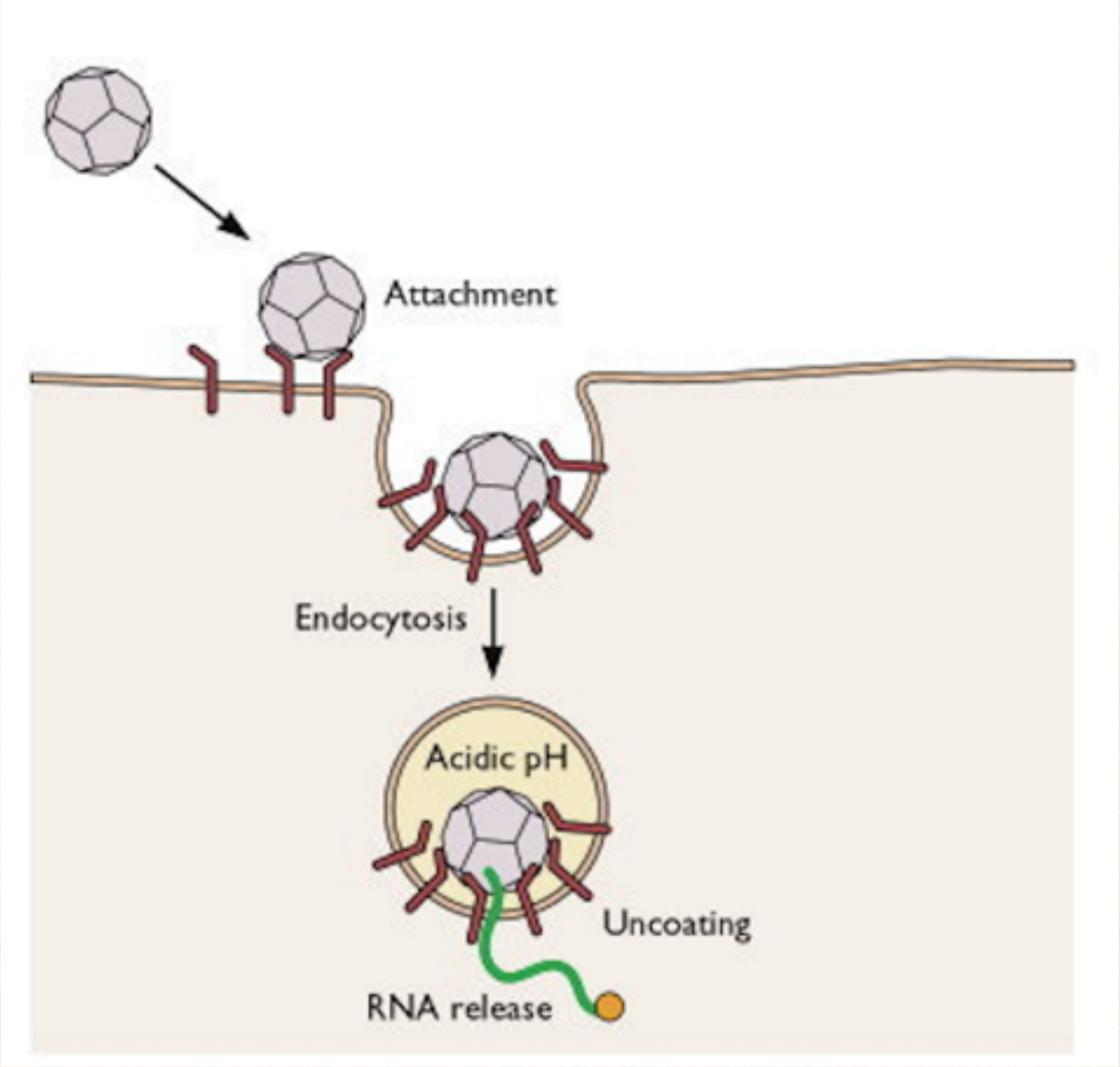
Particles advocated by an evolving random field.
(Cue: Christel Hohenegger)

- Serious challenges that await

  -- State-switching due to chemical interactions

Mucosal Immunology: Interactions between HIV and antibodies

Joint with Sam Lai,
(School of Pharmacy, UNC-CH)



- Thank you!